OPERAS-P

WP6 Innovation

D6.5 Report on the future of scholarly writing in SSH

30 April 2021
(updated 28 June 2021)

The project has received funding from European Union's Horizon 2020 research and innovation programme under grant agreement 871069

OPERAS-P

OPEN SCHOLARLY COMMUNICATION IN THE EUROPEAN RESEARCH AREA FOR SSH - PREPARATION

Deliverable 6.5

Report on the future of scholarly writing in SSH

Grant Agreement number	:	871069
Project acronym	:	OPERAS-P
Project title	:	Open Scholarly Communication in the European Research Area for SSH - Preparation
Funding Scheme	:	INFRADEV-02-2019-2020
Project's coordinator Organisation	:	CNRS-OpenEdition
E-mail address	:	pierre.mounier@openedition.org
Website	:	https://www.operas.unito.it/projects/operas-p/
WP and tasks contributing	:	WP6, Task 6.5
WP leader	:	IBL PAN
Dissemination level	:	PU
Due date	:	30/04/2021
Delivery date	:	30/04/2021

Authors	:	Maciej Maryl, Marta Błaszczyńska, Agnieszka Szulińska, Anna Buchner, Piotr Wciślik, Iva Melinščak Zlodi, Jadranka Stojanovski, Elisa Nury, Claire Clivaz, Bartłomiej Szleszyński, Kajetan Mojsak, Mateusz Franczak
Contributors	:	Interviewing and coding: Erzsébet Tóth-Czifra, Interviewing: Michael Kaiser, Valérie Schaefer, Vera Chiquet, Mate Juric, and Krešimir Zauder ; Case study analysis: Paweł Rams, Katarzyna Jarzyńska, Renata Rokicka, Izabella Ratyńska, Hanna Kwaśny, Marcin Wilkowski and Zofia Leśnik ; Other: Agata Morka, Magdalena Wnuk

Table of content

ABSTRACT

The report aims to present the result of work conducted in Task 6.5 (Future of scholarly writing in SSH) in the OPERAS-P (Open Scholarly Communication in the European Research Area for Social Sciences and Humanities – Preparation) project. The introduction puts forward the theoretical understanding of scholarly writing – especially in the context of the digital age – that underlines our research. In the literature review we address the state of the art in a reflection on the role of technology in scholarly communication. Through the interpretation of interviews with scholars, we reconstruct contemporary communication practices in SSH. The analysis of case studies provides ideas on how technology is responding to those needs. Finally, we present two conceptual prototypes (the living book and the digital scholarly edition) that could address some of the needs and could be put forward as future services for OPERAS.

EXECUTIVE SUMMARY

The report presents the results of work conducted in Task 6.5 (Future of scholarly writing in SSH) in the OPERAS-P (Open Scholarly Communication in the European Research Area for Social Sciences and Humanities – Preparation) project. The aim of the task was to explore current writing practices in SSH and thus inform future OPERAS activities on researchers' needs regarding publishing technologies, and both ongoing and upcoming transformations of scholarly communication.

In order to address its aim, the report combines various perspectives; this allows for a rich description of the phenomenon in question. It is divided into three main sections, these are dedicated to theory (introduction, position paper), empirical research (interviews, case-studies), and application of the results (prototypes). First, the **introduction** provides a theoretical exploration of the notion of writing as a means of transmitting thoughts and ideas. It is followed by a **position paper** offering our interpretation of the current trends in scholarly communication. The core of the report is its empirical section. **Interviews** with scholars and publishers provide insights into the practices and needs of the actors within scholarly communication with regard to the innovative aspects of that scholarly communication. The **case studies** complement these insights with an analysis of selected innovative tools and services. Some of the insights gained in this study were implemented in two conceptual **prototypes** for innovative services.

PART I Theory

INTRODUCTION: THE FUTURE OF SCHOLARLY WRITING

The notion of writing is explored here as the transmission of thoughts, which plays an important role in thinking processes as it allows for the verbalising of our thoughts. The theoretical exploration of the cognitive potential of communications' media in oral, chirographic, and print cultures, serves as a context for recognising how digital media influences the way we write and think.

POSITION PAPER: CURRENT VISIONS OF THE FUTURE

This section interprets recent landscape studies of scholarly communication in Social Sciences and Humanities (SSH) and the way in which writing is discussed within them. These studies concur in that the innovation in scholarly communication produced by the proliferation of open access mandates had its greatest effects on business models in the publishing industry, but has been much less transformative when it comes to harnessing the affordances of digital technologies to advance new modes of scholarly writing.

Platformisation, on the other hand, stands for a switch from a content-provision business model to one based on data analytics services that capitalise on the extraction of information about both the indexed publications and user activity on platforms that bring together several categories of customers (e.g. researchers, publishers, funders and universities).

We propose to understand the future trends in scholarly writing in terms of a spectrum of responses to this state of implementation of OA, and to the "platformisation" trend, which, as the landscape studies have forecast, will be the new frontier for commercial academic publishing. The interaction of these trends forms an ecosystem in which the directions in scholarly communication emerge.

PART II Empirical study

INTERVIEWS: SSH SCHOLARS ON INNOVATION

Key concepts. On the basis of the literature review, researchers collaborating on tasks 6.5 and 6.6 worked iteratively to define key concepts of the study that were pertinent to the future of scholarly writing. First, in order to open-up the study to various materials, we agreed to treat the **scholarly text** broadly, not only as a linguistic articulation, but rather as an expression that can utilise different media. Second, we prepared working definitions of the main concepts pertinent to the task: **Communicating** (the act of sharing a text through various formal or informal channels); **Specificity of SSH** (scholarly communication practices in Social Sciences and Humanities that are different from other fields; **Writing** (the act of generating a text, understood as the expression of an argument that may use various media,

formats, and genres); **Collaboration** (collective activities that are undertaken in writing, communicating, publishing, and peer-review); **Tools** (the services and software used in the process of writing, communicating, and publishing at various stages of the researchers' workflow); **Publishing** (the act of disseminating a text through a formal process, including intermediaries); **Innovative forms and genres** (text used by scholars to transmit their argument, beyond traditional formats of the journal article, book, report etc.); **Audiences** (the public who engage with scholarly texts and their authors); **Evaluating** (the critical assessment of products of all types of scholarly communication, i.e., writing, communicating, publishing); **Innovative forms of peer-review** (peer-review practices that go beyond the commonly accepted forms to address the perceived deficiencies of the system); **Academic prestige** (widespread respect attached to certain practices by scholarly communities); **Power structures** (dynamic systems of hierarchy and influence in scholarly communication).

Methodology. In order to gain an insight into scholars' experiences and opinions on innovation in scholarly communication, an interview scenario was prepared focusing on two types of knowledge: episodic (particular, based on one's own experience) and semantic knowledge (internalised generalisations about how things are). The interviews were conducted by a 10-person research team (representatives of IBL PAN, SIB DARIAH partner, UNIZD, DARIAH, Uni Lux, and MWS) between April and August 2020. Thirty-two interviews were transcribed and coded using MAXQDA, and 9 have been summarised and made available for further studies

Main Results:

SSH specificity:

● There are differences in scholarly communication between SSH and other disciplines, as well within the disciplines of SSH itself. These concern issues ranging from output genres and the aims of peer-review, to collaboration strategies and funding.

● The main communication genre reflects the features particular disciplines value most: in the case of the sciences, it is the timely reporting of facts, while the humanities value the depth and breadth of the interpretation. Natural and social sciences tend to communicate using journal articles, while in the humanities the monograph prevails.

● Writing in the humanities is more subjective, rooted in context, and less formally structured.

● Collaboration and authorship attribution practices in SSH differ across disciplines.

Scholarly writing

● Writing is a deeply social and technologically supported activity. We consider writing to be a different activity to publishing and communicating research. Discovery, storing, reading, and annotating research assets are treated as part of the writing process, as they influence the outcome and have already generated content that will be used in the drafting phase.

● We distinguish between digitally-enabled writing and digital writing. The former refers to writing as a textual practice supported by various digital tools, while digital writing

harnesses the full potential of digital technology by establishing different kinds of materials, like data, visualisation, or pictures, in a single output.

● Scholars appear to use many different tools in their unique workflows, deeply rooted in the individual preferences or experience, types of project, and disciplinary needs.

● When it comes to tool selection, we distinguish between two types: engineers and bricoleurs. The engineers are experts in many specialist tools, and fluidly switch between them. Bricoleurs, on the other hand, still combine digital practices with the analogue, offline ones.

● Many interviewees distinguished between the ideation and drafting phase of the writing process. *Ideation* is an activity usually conducted in plain text to keep things simple and prevent technology from interfering with the thought process.

● Interviewees reported having different experiences with collaboration, ranging from, on the one hand, negotiating the process with multiple authors, which may ultimately lead to a patchwork structure; through to a dry, "impersonal" language, on the other. Good writing collaboration requires trust and boldness, and often some level of personal, informal connection. Collaborative writing requires a lead author, who will organise the process and coordinate the team.

Choosing the publication type:

● Authors choose the form that is appropriate to the content, based on its length or on the possibility of developing a narrative or argumentation. In many cases, the choice of publication type or venue is not just an outcome of deliberate consideration, it can also be the result of taking part in a project, collaboration or peer network, or through personal acquaintances.

● Novel formats and genres are considered more appropriate for certain content for several reasons: they are liberating, communicative, interactive, and collaborative; and they enable versioning and updating. Sometimes experimenting with new forms aligns with the research agenda.

● Authors try to choose publication venues based on their expected future discoverability and visibility. They often prefer publishing in publications with good quality metadata, in high impact international journals that are indexed in international citation indexes, in the English language, and in reputable monograph series that will attract many book reviews.

● The reasons for favouring open access venues can be threefold: national or institutional mandates, personal principles or ideologies, and reaching a wider audience.

● Early career researchers are faced with stronger pressure to make the right choice, and more limited options in publishing venues.

● The speed of publication can have a role in the choice of the type of traditional publication, but it is also considered to be a major advantage of communicating through novel publishing genres (like blogs).

Traditional publication

● Books are defined by their form, but also by their status within the scholarly community. They do not easily conform to metrical systems of defining quality. For books, there are separate ways of defining and expressing impact and reputation. For

respondents, the central position in the SSH publishing ecosystem was reserved for the scholarly monograph.

● When it comes to multi-authored and edited works, it is difficult to reach a common understanding of the nature and status of such a format. There are no clear borders between special thematic journal issues, conference proceedings, or edited volumes, but their status and reputation are often perceived in different ways.

● In many of the SSH disciplines, journals are the main vehicles for communicating current research in a formal way.

● The main characteristics that make journal articles so practical for communication in academia are their brevity, structure, speed of publication, and peer review process.

Innovation in scholarly communication

● Innovation is seen as a chance to improve the sharing of ideas with audiences, thanks to technological affordances. Innovation is also understood to bring seamlessness to the process of using scholarly content: removing obstacles that are unnecessary from the vantage point of current technology. Innovation is, then, understood either in terms of form (novel means of communicating ideas), or access, i.e., accessing content or reaching new audiences.

● Innovation is mostly considered in terms of providing access to more traditional types of outputs. In this context research data and grey literature become more accessible.

● Formal innovations concern a move beyond the mere written word, that is, accepting expression in other media forms as being valid scholarly outputs. The form of a computational essay allows research and data to be linked, allowing for interaction.

● Thinking of innovation in terms of audiences means improving the communication of research findings and, thus, the perception of research in society. This is done through social media and blogs.

● Innovation is impeded by such factors as quality assessment, prestige, competencies, and a lack of established standards for referencing novel forms. The issue of how to use novel sources in a scholarly text is one of the challenges of 21st-century scholarly writing.

● These challenges push scholars toward practices of double referencing and double publication, whereby the traditional publication provides prestige for the novel form.

Openness and transparency

● Scholars unreservedly support open access as a tool for improving visibility, showing their responsibility towards the benefiting of society, changing present prestige patterns, and changing the system of scholarly communication.

● There is still a large gap between the apparent benefits of open access and the present criteria for academic career advancement, and scholars fear that publishing in open access could impair their chances of employment, diminish the value of their CV, or reduce their career prospects.

● Open access publications offer endless possibilities for improving traditional publication channels by connecting open content and taking full advantage of hypertext and other available technologies.

● The great majority of our respondents were familiar with the Author Publication Charge (APC) or Book Publication Charge (BPC) models, and they often found present

APCs and BPCs "unaffordable," "gigantic," "exploitative," "impossible to override," and "leading to a less equitable world," especially for early career researchers. Open access must support "not only equity in distribution, but equity in terms of participation in the publication".

● Many authors are aware of the “green route” to open access, and are opposed to paying APCs instead of self-archiving.

Audience

● Authors are usually pragmatic when choosing the best possible venues for reaching different audiences. They are aware of the different discovery paths and different ways of using content.

● Many respondents had expectations that new formats and communication channels would more easily find their way to new or wider audiences even though the primary addressee for scholarly works remains the community of their peers.

● Writing for popular publications can be a source of prestige, although societal impact will not always correlate with the bibliometric one.

● Popularisation activities are regarded, mainly, as being an important aspect of scholars’ work that can even become a source of prestige within academia, but concerns are raised that non-traditional and popular communication outputs would be considered inferior (primarily because they are not easily quantified and measured in various assessment systems).

● Factors that can contribute to the visibility and discoverability of scholarly outputs by potential audiences, are: online presence; open availability and affordability; coverage in citation and bibliographic databases, or search engines; quality metadata; using PIDs; and the help of social media and public online profiles.

Power structures

● Researchers themselves, and the community more broadly, are recognised as important actors in the SSH scholarly communication landscape. Depending on their approach, they can play the role of guardians of the status quo, or innovation facilitators.

● Our respondents often pointed to editors as being one of the most important groups in academic power structures, especially in the context of publishing, as they select the reviewers for scholarly texts and tend to make the final decisions. Other important groups included institutions, funders and publishers, and the “big names” (scholars recognised in their communities).

● The relative importance of the different groups is often hard to determine. The respondents referred to issues of transparency and the complexities of power structures as explanations for this difficulty.

● Innovative forms of writing could challenge traditional structures, giving more gatekeeping power to the wider readership community.

Prestige

● Many scholars admit that there is a strict hierarchy for publications, with the monograph often being mentioned as the most prestigious output in SSH.

● Factors that influence the trustworthiness of a publication in the eyes of fellow scholars include: relevance (including recommendations or the fact that the publication

is cited elsewhere), bibliometrics, authorship, the publisher or the journal, and peer review.

● Innovative forms of writing do not yet have an established position in academia. Some respondents had already expected novel solutions from their colleagues and referred to digital outputs (such as blogs or tweets) in their own work, whereas others saw them as undervalued and difficult to cite.

● There is no consensus on the relationship between open access and prestige.

CASE STUDIES: THE STATE OF INNOVATION

In the case studies analysis, we explored the current landscape of scholarly writing, focusing on the actual directions of the innovations in scholarly communication. We were interested in how the providers of tools and services defined and responded to the needs of researchers, and which innovations were considered useful. We looked at the future of scholarly writing through all stages of the research cycle (discovering, organising, annotating, analysing, publishing).

Methodology. Altogether, 56 cases were identified for analysis through snowball sampling, for which we used various sources such as literature reviews and suggestions gathered from interviews and presentations during events on innovations (e.g. Open Book Fest). We aimed for sample diversity rather than representativeness. On the basis of a literature review, a detailed study protocol was prepared in order to serve as a guide for researchers. It was attuned to the particular issues identified during the earlier stages of the project. The case studies were analysed by a team consisting of two main researchers, plus collaborators (including trainees).

Main findings

Users' needs.

● We focused on users, from two aspects. First, we looked at the reasons behind the development of a particular project and its current aims, i.e., what gap was this project trying to bridge. Second, we attempted to reconstruct intended user roles and those users' communication with the project's creators.

● We observed a progressive trend in inclusiveness in design with regard to scholars from the humanities and social sciences. On the one hand, there were projects of a more general form that sought to provide for the needs of multiple projects, data, methods, and disciplines (e.g. SSH Open Marketplace); while on the other, some addressed relatively narrow areas like the research activities around manuscripts.

Users

● In this section we look at how the platforms define their users and provide space for interaction. Identifying needs and knowing your target users can be achieved through surveys.

● Services engage with their users on many levels in order to ensure that the services address the actual gap and that the training provided fits scholarly needs.

● A method that is gaining popularity is to invite testers (people from outside the circle and network of developers) to research and open-test a tool or service, which is usually done during the development stage.

Data and technology

● We analysed the formats of the research data used in projects as well as the features and functionalities provided to handle them.

● Scholarly text in the electronic environment is often supplemented with, and well connected to, underlying data. Many services try to facilitate the connection between text and data, as it is also seen as providing a richer context for scholarly texts, visuals, graphs, and audio materials.

● When it comes to data publishing, the move beyond a mere PDF was put into practise in many of the analysed cases.

● The choice of formats for data publication corresponds with the type of project and its particular needs.

Teams and their workflows

● This section focuses on how affiliation and authorship are handled within the project. It also discusses the workflows behind the projects (the team's structure, leadership, responsibilities, and roles).

● Teams working on digital projects need to be made up of members who perform various roles (e.g. publishers, editors, engineers, developers, UX designers).

● Likewise, organising the workflow for digital publishers involves coordinating the work of diverse actors.

● The authorship of outputs like digital collections may differ across particular projects, depending on the type of work involved.

Availability and accessibility

In this section we focus on the entry requirements needed to use the analysed services. We look at whether users need to sign-up, provide affiliation, or pay. We also analyse whether tools are accessible via an existing platform or need to be installed and operated by the user's organisation. We also checked the compatibility of tools with different browsers.

● In most of the analysed projects, no login was required to access the content. Even if readers had to create an account, the content might still be open, but registration unlocked additional features.

● While a project is still in development, users are sometimes given the opportunity to apply for early access.

● The majority of analysed services declared a free of charge approach to their content.

● In the majority of cases open source and open access approaches were declared.

● Reproducibility, connected with openness, is perhaps of greater importance for STEM disciplines. Some platforms provide data on demand for reproducibility purposes.

Sustainability

The sustainability of the tools and projects related to the future of academic writing is essential not only for their creators but also for users, who need assurance that a tool will be available in the longer term.

● A grant-based model of funding prevailed among the analysed projects. In some cases a single funder provides several grants to support consecutive phases of development of a project. The funding situation is also strongly connected to the national context.
● Crowdfunding is not a common option for professional academic texts. Perhaps it would be more suitable for research popularisation.
● Many services are based on a freemium model, offering paid plans with extra features.
● Another sustainability strategy is to establish consortia for mutual support and shared services.
● Sustainability is achieved not only by financial means, but also through data standardisation with PIDs.

Impact and prestige-generating mechanisms

● This section discusses the usage of the projects we analysed. We look at examples of use and the impact these services may have on society and education. We are also looking into prestige-generating mechanisms and outreach strategies.
● Various types of statistics reported by platforms serve as a source of knowledge about their size and impact.
● Citations are still considered the basic impact measurement for scholarly outputs. Although citations aren't anything innovative, some services offer new ways of *boosting* them.
● To achieve social and educational impact, it is always worth asking which features of a given project are especially important for the broader audience. Some projects provide popular releases targeted at non-scholarly users.
● Social media remain the main channels for promoting projects. The most common strategy is to use at least one of these channels: a blog, Twitter, or Facebook.

PART III Applications

Prototypes. During the research process we identified two main innovations that were important from the vantage point of OPERAS. These were the living book and digital scholarly editions. The former was considered to be an important form for updating scholarly writing in connection with OPERAS Special Interest Groups. The latter was identified as an important form that allowed for innovative connections between text and data.

The living book – prepared in the course of the OPERAS-P project – is a prototype aimed at responding to the needs of the OPERAS community. Based on the needs assessment, the analysis of similar projects, and available technology, we prepared tailored solutions aimed at addressing the key requirements of this community, i.e., the sustainability of the solution, an automatically updated Zotero bibliography, versioning, referencing, commenting, and quality control. We will work closely with SIG (Special Interest Group) leaders during "reviewathons," observing how discussion on the white papers unfolds and how actual users interact with this format.

The digital scholarly editions prototype offers an exploration of a specific SSH genre that is dedicated to the publication of sources. A detailed analysis of the specificity, advantages, and problems of digital editions is followed by a set of concrete solutions: creating free tools for editing, providing infrastructural support to communities, creating a detailed catalogue of the tools available at any given time on various platforms, precise agreement on common standards and accordingly the adaptation of tools, and the possible integration of tools into a single research infrastructure.

1 Introduction: Essay on the future of scholarly writing

1.1 Something is coming to an end

"Does writing have a future?" (*Hat Schreiben Zukunft?*); this was of concern to Vilém Flusser, a philosopher and media theorist, in his 1987 book that examined the transition in writing from media-technological contexts towards the digital age. The volume comprises a series of essays that dissect particular components of writing in order to assess whether this fundamental practice could be replaced by other media:

> Writing, in the sense of placing letters and other marks one after another, appears to have little or no future. Information is now more effectively transmitted by codes other than those of written signs. What was once written can now be conveyed more effectively on tapes, records, films, videotapes, videodisks, or computer disks; and a great deal that could not be written until now can be noted down in these new codes. (Flusser [1987]2011: 3)

Writing is, thus, the act of transmitting certain messages. However, it also plays an important role in thinking, allowing our thoughts to be verbalised. As Flusser observes, writing "is a gesture of setting up and ordering written signs. And written signs are, directly or indirectly, signs for ideas. So writing is a gesture that aligns and arranges ideas. Anyone who writes must first have thought. And written signs are the quotation marks of right thinking" (ibid.: 6). Written signs, poetically compared to quotation marks, serve as containers of thought, putting some structure and boundaries on ideas to make them communicable.

This connection between what Flusser calls inscription technologies and human thought, is thoroughly considered in the works of such scholars as Marshall McLuhan (1962), Eric Havelock (1967), William Goody (1977), Walter Ong (1977, 1982), and Friedrich Kittler (1990, 1999), so we will limit this discussion to some basic observations. In oral cultures, thought is closely connected to the subject and the context of utterance. Handwriting allowed for the externalisation of one's experience, and introduced temporality beyond the spectrum of one's immediate experience. It made history possible by allowing events to be noted in a sequential order: "before writing was invented, nothing happened; rather things merely occurred. For something to happen, it has to be noticed and conceived as an event (process) by some consciousness" (Flusser 2011: 8). The moving type of the printing press allowed for a kind of thought that is typographic, or "typifying" as Flusser calls it (2011:53). It introduced the notion of objectivity, as facts could now be printed without the intervention of a copyist as was the case with manuscripts. Each time, the technology offered a more fluid way of thought inscription, providing newer tools to facilitate the process: "everything becomes structurally more complex, to become functionally simpler. […] After the goose quill came faster and faster writing instruments: [the]

ballpoint pen, typewriter, and word processor – faster and faster quills" (Flusser 2011: 18). Finally, the advent of digital technologies has once more reshaped the way we think and communicate:

> The informatic revolution, this production of signs and their positioning in electromagnetic fields, openly breaks with print consciousness. The new signs that appear on computer or television screens are no longer traces engraved in objects; they are no longer "typographic." The kind of thought that is producing the new information is no longer typographic, typifying kind of thought. […] It is fairly clear what will be lost in the transition from Gutenbergian to electromagnetic culture, namely everything we treasure in the Western legacy. On the other hand, we do not see what we have to gain. If we could do that, we would already have reached the first step toward the new way of thinking. Flusser [1987] 2011:52-53

"A medium is a medium is a medium," wrote Friedrich Kittler in his paraphrase of Gertrude Stein's take on a rose, meaning that a medium "cannot be translated. To transfer messages from one medium to another always involves reshaping them to conform to new standards and materials" (Kittler 1990: 264). Thus every change in communication technology brings about the reconfiguration of the broader scene, which Kittler names a "discourse network" (*Aufschreibesystem*). He distinguished two main networks: 1800 (based on print and the book) and 1900 (the breaking of the typographic monopoly thanks to audio and audiovisual media), in addition, the upcoming network of the "total media link on a digital base [which] will erase the very concept of medium" (1999:2). However, as Kittler observes, "Before the end, something is coming to an end" (1999:1). That is, we are currently living in the age between the audiovisual system and the fully digital one, whereby "[t]he general digitisation of channels and information erases the differences among individual media. Sound and image, voice and text, are reduced to surface effects, known to consumers as interface" (ibid.).

Something is coming to an end. It is very important to note that we are currently in a dynamic, transitory phase, which could by no means be considered final. It could only be compared with the age of incunabula, the half century in which the rapid development of print technology and associated practices coexisted with the most prolific period for manuscripts in history. It took some time before the characteristics of the printed codex were codified and the manuscript came to an end. In order to remember how rapidly things evolve and to what extent our thinking is confined by the features of current technology, we need to remember how the future of writing was perceived at the beginning of the digital revolution.

The 1996 volume *Future of the Book* is an interesting snapshot of the early debates on the role of writing and the word in the digital age. In his afterword, Umberto Eco remains rather optimistic about the future of writing, foregrounding the evolution and similarities with print culture rather than the ruptures. Eco perceived computers as vehicles to enhance or amplify some characteristics of print, and as a means of better diffusing printed, not digital-born, documents. Perhaps more stress was put on visual

materials and hypertextual narratives, both of which he conceived as already being prolific in non-digital culture. That Eco saw the future of writing largely in visual terms, but not audiovisual or aural, testifies to the supremacy of certain means of storage, compression, and transmission in his day, technologies that handled images better than sounds. Available technology often limits the horizon of those future innovations one can imagine. Interestingly, the institutional dimension of the change has already been noted, as "[p]eople can communicate directly without the intermediation of publishing houses" (301). However, in this context, Eco recalled Landow's remark in the same volume, that "we are entering a new samizdat era," underlining that digital circulation is conceived of as being a different means of text distribution, less formally established or controlled by gatekeepers.

Eco seems to be visionary in some places ("new technologies will render obsolete many kinds of books, like encyclopaedias and manuals" 1996:299), and anachronistic in others ("users who want to learn the program generally either print the instructions and read them as if they were in book form, or they buy a printed manual" 1996:300). However, he had already noticed that the form of the book, as a vehicle for the long argument, would prevail: "Books will remain indispensable not only for literature, but for any circumstance in which one needs to read carefully, not only to receive information but also to speculate and to reflect about it" (Eco 1996:300). It does not matter so much whether this book is printed or digital.

To sum up, we are currently experiencing a fundamental transition that is reconfiguring our approach to communication. In knowing the connection between writing and thought, however, we should anticipate how technology could affect the way we think and communicate.

> We can regard print, this alphabetic writing that has become self-aware, as the expression of Western, historical, scientific, progressive thought. The informatic revolution makes print, the alphabet, and this kind of thought superfluous. It leads to a new mode of thought that can be anticipated but not yet perceived. That sounds like an assertion, but it is really a concerned and hopeful question directed toward the future. (Flusser [1987] 2011:52-53)

1.2 Concerned questions directed toward the future

In this report, we also pose concerned and hopeful questions that are directed toward the future of scholarly communication. We try to assess its current state, directions, and future needs. By future, we obviously mean a span of upcoming years as it is impossible to infer the general direction from within the process. What complicates this further is that every now and then a unique combination of technology and social factors may contribute to a temporal rise of one form over another. For instance, the re-emergence of podcasts as a popular form of scholarly communication is likely a product of both the advancement of streaming technologies and the need to recreate less formal communication such as speech or interviewing.

In this report we assume a broad and pragmatic definition for the scholarly text, as the specific output of a research process that can be accessed, used, and referenced by other scholars. Its main aim is to deliver the scholarly argument, backed

by evidence, the exact type of which differs across disciplines from factual to speculative. Moreover, it may be directed towards various audiences, including those beyond the scholarly community. The range of such outputs differs across disciplines and depends on many factors, for instance, whether they report facts or interpretations. We note that the spectrum of what could be considered a scholarly text has broadened, and we have tried not only to catalogue these new forms, but also to describe the attitudes researchers have towards them.

We recognise that the activity of writing is not limited to the act of putting one's thoughts on paper, and that contemporary researchers perform diverse roles in scholarly communication. We look at how they use technology as authors, collaborators, reviewers, data curators, editors of collected volumes, journals, and sources. We see the role of technology as a means of supporting these practices and presenting a wide set of features that scholars can choose from. How they construct individual workflows depending on needs, competencies, research interest, and available tools, is a special focus of this study.

The work of the OPERAS-P Task T6.5 (Future of scholarly writing in SSH) started with the consideration of how technology affects writing processes. In the literature review we addressed the state of the art in reflecting on the role of technology in scholarly communication, which allowed us to formulate the main fields of investigation and associated research questions. Through the interpretation of the interviews with scholars, we reconstruct contemporary communication practices. The analysis of case studies provided ideas on how technology is responding to those needs. Finally, throughout the project we worked on two conceptual prototypes (the living book and the digital scholarly edition) that could address some of these needs and could be put forward as OPERAS' future services.

Although it is not part of the report, it should be mentioned that most of the work presented here was conducted during the COVID-19 emergency, which heavily altered the anticipated workflow. Team members had to work in lockdown conditions, coping with isolation and "Zoom fatigue" while caring for young children, or their loved ones affected by the disease. Some of the team members themselves suffered from COVID directly. We needed to alter the workflow and move all the meetings, for what was a large and distributed team, online. The February conference, planned as a 3-day working session and brainstorming event, was also significantly shortened and moved online.

On the other hand, we should also note that the general role of digital technology in scholarly communication has intensified as a result of the pandemic. Speaking about the "new normal" has already become a cliché. However, from the current position it is very difficult to assess what the long-lasting impact of the pandemic may be on scholarly communication. It will be no earlier than a year or so after all the restrictions are lifted that we may discern whether workshops, conferences, and lectures held in the digital environment still remain as attractive as they are right now in the times of restricted mobility. However, one thing is certain: the pandemic has forced scholars to broaden their digital competencies and to take up tools and services they may have been reluctant to use. This in turn may broaden the audience for innovative tools,

services, and formats, and hopefully catalyse the complex processes we have tried to capture in this report.

1.3 Bibliography

Eco, Umberto (1996) “Afterword” *The Future of the Book*. Ed. Geoffrey Nunberg, Berkeley: University of California, pp. 295-306.

Flusser, Vilém (2011[1987])) *Does Writing Have a Future?* Trans. Nancy Ann Roth, introduction Mark Poster, Electronic Mediations vol. 33, Minneapolis: University of Minnesota Press.

Goody, Jack (1977) *The domestication of the savage mind*. Cambridge: Cambridge University Press.

Havelock, Eric A. (1967) *The Muse Learns to Write. Reflections on Orality and Literacy from Antiquity to the Present*. New Haven, CT: Yale University Press.

Kittler, Friedrich A. (1990) *Discourse Networks 1800/1900*, trans. Michael Metteer, Chris Cullens. Stanford: Stanford University Press.

Kittler, Friedrich A. (1999) *Gramophone, film, typewriter*, trans. Geoffrey Winthrop-Young and Michael Wutz.

McLuhan, Marshall (1962) *The Gutenberg Galaxy: the Making of Typographic Man*. London: Routledge and Kegan.

Ong, Walter J. (1977) *Interfaces of the word: studies in the evolution of consciousness and culture*. Ithaca, N.Y.: Cornell University Press.

Ong, Walter Jackson (1982) *Orality and literacy: the technologizing of the word*. London: Methuen.

2 Literature review: current visions of the future

This section builds on the systematic literature review undertaken within the OPERAS-P project (Maryl et al. 2020), and discusses, in particular, the recent landscape studies of scholarly communication in Social Sciences and Humanities (SSH) and the role of academic writing therein. These studies concur in that the innovation in scholarly communication produced by the proliferation of OA mandates, by which we broadly mean all state-level and institutional policies encouraging or enforcing open access in Europe, had their greatest effect on business models in the publishing industry but have been much less transformative when it comes to harnessing the affordances of digital technologies to advance new modes of scholarly writing. This section proposes to understand the future trends in scholarly writing in terms of a spectrum of responses to that state of OA implementation, and to the platformisation trend that, as the landscape studies have forecast, will be the new frontier for commercial academic publishing. We discuss this spectrum of responses as a spectrum of remediation, namely the degree to which innovation in scholarly communication is adaptive or disruptive with regard to established practices – the codex form of monograph, single authorship, and the double-blind peer review – and with respect to the platformisation trend.

2.1 Disillusion with OA and the platformisation trend

For some time now, academic publishing has been regarded, both by experts (ISC 2021) and the public (Guardian 2019), as a case of market failure having given rise to publishing oligopolies, reporting profit margins in the range of 40 per cent. These profits were made possible by the notorious practice of "double-dipping," in which publicly funded research institutions sponsor the production of knowledge and its assessment of quality, while at the same time the large publishers monetise access to this record of science without facing competitive pressures, due to their oligopolistic market share.

The implementation of OA mandates – in its most popular form, a switch to an author-pays business model, i.e., from subscription to Article Processing Charges (APC) in the case of journals, and Book Processing Charges (BPC) in the case of books – failed to curb the double-dipping practice. It is yet to be seen whether the European Commission supported, Plan S – which mandates open access to publicly funded research by 2021 for the funding agencies that joined the initiative – will realise its ambition to introduce price transparency and to standardise and cap the APCs and the BPCs (point no. 5 of Plan S), as this ambition is to be realised mostly through monitoring and encouragement rather than actual mandate (cOAlition S nd).

The OA mandates coincided with a proliferation of academic governance models, often identified as being neoliberal (Holmwood 2020, Moore 2019, Bacevic and Muellerleile 2018), which rely on efficiency and excellence (Readings 1996) as well as on the quantitative methods with which they are measured – for example, the Journal Impact Factor, or the ranking of academic presses – instead of collegial, deliberative goal setting and qualitative peer self-assessment by the scholarly community. In addition to increased teaching, fundraising, and reporting duties, the neoliberal governance created pressure on academics to increase their publication outputs in the traditional formats rather than experimenting with new forms that would not be visible for the purposes of assessment measurement (Hitchcock 2016). Even though some of the surveyed authors correctly observed that the record of open governance in science before the onset of neoliberalism had not been particularly inclusive or transparent (Eve and Gray 2020), it is clear that the neoliberal governance is a self-reinforcing trend: the more the scientific community relies on the quantitative measures of efficiency assessment, the more pressured the academics are to produce; and so, the numbers of published books are increasing even as sales of individual titles and library budgets remain the same at best (Deegan 2017). The more scholars produce, the greater their need to rely on quantitative measures, as overabundance of published outputs comes together with an increased scarcity of time that could be devoted to reading, be it for the purpose of research or quality assessment. For the purposes of both assessment and exploration, one way to escape the above vicious circle would be, as Eve (2020) proposes, to open the records of science to methods of data mining, allowing scholars to design customised queries in place of preset algorithms/filters over which they have little control. But such an opening, as the author is well aware, will not happen as long as the dominance of for-profit publishers persists, since data about

science, rather than the scholarly record itself, is, as the ISC report (2021) suggests, the lucrative new frontier.

In this regard, the trend that will be decisive in configuring the landscape of the academic writing of the future is platformisation (Andrews 2020). Platformisation stands for a switch from the content-provision business model to one based on data analytics services that capitalise on the extraction of information about both the indexed publications and user activity on platforms that bring together several categories of customers (e.g. researchers, publishers, funders, and universities). In this model, the sharable academic content and the interaction built around that content might well be open (once APCs and BPCs are met), but the metadata about both the content and the interactions, and the algorithms on which the analytics products are based, are proprietary. In the academic publishing space, this trend is acutely patent in the portfolio of the Netherlands-based global corporation Elsevier, which, in addition to the citation database Scopus and indexing tool SciVal, acquired the Social Science Research Network academic networking site, the reference manager Mendeley, the institutional repository and journal publishing service bepress, and the PlumX altmetrics tracker. Similarly Clarivate Analytics, the owner of the citation database Web of Science, moved into the same market space through its acquisition of platforms providing tools for open scholarly cooperation such as the peer-review tracking service Publons and reference manager Endnote. Other examples include Holtzbrinck Publishing Group, the owner of Springer Nature, Digital Science, and Altmetric; as well as the venture capital owned Academia.edu, which offers premium services based on data about academic networking, generated at no cost by its users. Not unlike the Big Tech corporations such as Google, Amazon, and Facebook, large corporate players in the field of scholarly communication thus evolve in the direction of deriving revenue from the pro-social, sharing behaviours of academics who wish to engage in free and open knowledge exchange, and who are often ambivalent about the business models behind these platforms and tools.

“The idea that the publication and circulation of science and scholarship should not be controlled by profit-seeking corporations” Maxwell (2019) observes “has led in recent years to a recognition that profit-seeking corporations, while possibly ceding ground on OA itself, had an almost total lock on the technological infrastructure that runs scholarly communication and publishing.” The vision of future inclusive scholarly communication infrastructure for SSH developed in the OPERAS-P project clearly states that scholarly leadership over such initiatives is key to their attunement to the actual researchers' needs: “Only a scholarly-led, transparent, and researcher-oriented infrastructure will truly address the existing and emerging needs of scholars of the digital age, by basing its activities on the actual, empirically-evidenced needs of the community, not on the pursuit of commercial revenue” (Maryl et al. 2020). Discussing the history of the commercialisation of scholarly publishing in the second part of the 20th century (cf. Fyfe et al. 2017), Maryl et al. call for the seeking of a balance between the interests of the multiple stakeholder groups: “reclaiming scientific communication does not mean excluding commercial players, but rather providing a healthy balance between the commercial interests of publishers, providers, and researchers, which

would protect the interests of scholars and smaller players" (2020). The condition is, they add, a close collaboration between policy makers, funding bodies, research performing organisations, learned societies, and publishers.

The ambition of corporations such as Elsevier and Clarivate is to provide an end-to-end data analytics solutions with metrics corresponding to each phase of the research cycle (Aspesi and Band 2020). The platformisation trend foregrounds the danger of the proprietary locking-in of data about science. That locking-in would make university governance even more dependent on corporate vendors for measuring and forecasting performance and impact. That dependence would also increase in the area of support for scientific discovery, as the evidence based acquisition models proliferating in academic libraries progressively rely on measuring the attention and downloads on proprietary websites that aggregate journal and book titles, and on proprietary tools that track scholarly communication in social media and other outlets (ISC 2021, Jubb 2017).

Platformisation will have a substantial impact on the future of academic writing in two ways. First, in the way it will drive and scale-up those innovations in scholarly communication that are best aligned with the data-extractive, interaction-intensive business model, for example, forms of measuring post-publication impact, such as open peer review and alternative metrics, thus ironically capitalizing on initiatives that were meant to provide an open science alternative to the oligopolistic enclosure of both science data and data about science. It will also create pressures on academic publishers to adjust their workflows to fit the design of discovery platforms; and presses like Taylor and Francis, CUP, OUP, and Wiley already require their monograph authors to produce extensive metadata at the book-chapter level (including abstracts and references) – regardless of whether that makes sense from the perspective of the long-form genre of scholarly argument – in order to be able to make the content better visible to content aggregators (Jubb 2017).

Second, the future of scholarly writing will depend on how supporters of open science will respond to the platformisation trend; this response will be the focus of the remainder of this section.

2.2 Remediation of scholarly writing: the spectrum

To understand the contemporary changes, we will dwell on the notion of *remediation*, coined by Bolter and Grusin as the "defining characteristic of the new digital media" (2000:45). Remediation, to put it simply, is the "[r]epresentation of one medium in another" (ibid.), so that the new forms are seen as preserving the affordances of their predecessors while complementing them with features enabled by the newer medium. Thus, the innovations in scholarly writing can be understood as a spectrum of remediation, that is in terms of how digital technologies are used to emulate or disrupt the traditional format of the print codex, the related publishing workflows, and its scholarly affordances.

At one end of that spectrum, we encounter efforts that focus on creating non-profit, open source, and scholar-led infrastructural alternatives for the commercial platformisation trend. As Maxwell (2019) describes in his review, innovations in this

area are aimed at the digital modelling of OA publishing workflows to increase their sustainability, cost-effectiveness, scalability, and interoperability. In their functional scope they range from end-to-end solutions, like the Open Journal System or Libero, to offering specific functionalities, such as conversion, ingestion, the XML/HTML/Markdown rendering of text and typesetting, or managing submission and review processes. At the same end of the spectrum there is also a significant discussion on equitable business models that would foster care for community-owned infrastructure at a scale that creates an alternative to proprietary systems, and on designing transparent and inclusive governance structures, which would foster online academic networking communities that actually work for the good of the community they serve (Maxwell 2019, Penier, Eve and Grady 2020, Fitzpatrick 2020, Maryl et al. 2020).

At the same time, innovation in the area of publishing workflows, business models, and governance often does not extend to experiments in scholarly writing per se. Experimental publishing in this sense stands out as being at the other end of the remediation spectrum.

This end is purposefully diverse and disruptive. The diversity, catalogued in detail in Maxwell (2019) and in Adema, Mars, and Steiner (2021), ranges from projects such as Fulcrum and Scalar, which experiment with the integration of textual and non-textual elements (images, multimedia, graphs, maps), or text and code (Jupyter Notebooks); through collaborative writing, annotating, versioning, and post-publication review, enabled by platforms such as PubPub and Manifold; to futuristic efforts aimed at reimagining scholarly communication on linked open data principles (dokieli). Again, some of these projects offer end-to-end solutions, while others are specific interventions in a particular aspect of the scholarly workflow, such as the Zotero reference manager and Hypothes.is annotation tool, both of which also point to the fact that experimenting with the digital affordances of scholarly communication does not always entail introducing a completely new software design into the ecosystem but could be based on existing technical solutions. For example, the volume *Interacting with Print: Elements of Reading in the Era of Print Saturation*, which emerged out of an experiment in massive collaborative writing undertaken by 22 scholars from the UK, the US, and Canada (known as the Multigraph Collective), was created using wiki software (Multigraph Collective 2018). Some of these experiments are already being discussed as new genres of scholarly writing, for instance, the enhanced monograph (integrating multimedia), the living book (ongoing, collaborative writing), and the hybrid book (existing in various print and non-print versions), while others do not yet have a name in public circulation.

All that diversity brings disruption, as it challenges the tried and tested models of scholarly communication. As Adema, Mars, and Steiner (2021) affirm: “Experimental forms and practices of publishing open up and explore questions around modalities, linearity, workflow, and the relationalities of publishing; they examine established practices that we have often been taking for granted or have been repeating uncritically within conventional forms of publishing – where they have become solidified in standard print- and codex-based publishing forms and practices.”

The practices of experimental publishing stretch the forms and genres of scholarly output; for example, enhanced monographs create non-linear mash-ups of text and non-textual elements (images, graphs, maps) that question the boundaries between the long-form of scholarly writing, art books, virtual exhibits, and digital collections. Experiments in collaborative writing, the living book format, and continuous peer review, challenge not only the established conceptions of authorship, but also the core distinction of scholarly communication and the distinction between research and publication, between the process of scholarly discovery and the distinct moment when the truth claims are integrated into the record of science, thus acquiring epistemic stability and fixity and so become open to judgement. The undoing of that stability might be offset by versioning, for example; but, still, its transformative effect on scholarly communication is substantial. Relationalities in publishing workflows are also reimagined (Adema et al, 2021) at this end of the remediation spectrum. For example, enhanced monographs suppose the increased involvement of graphic designers from the very early stages of the publishing process; and post-publication peer review in monographs calls into question the role of the acquisitions editor in the entire workflow, acquisition being the most costly activity out of all those in publishing (Penier, Eve and Grady 2020).

2.3 The problem of scalability

The friction between the two sides of the remediation spectrum revolve around scale and scalability. The surveyed landscape studies that recommended that the priority was to create an open infrastructure for scholarly communication that would be competitive enough to avert the danger of platformisation, proposes concentrating innovation efforts on making publishing workflows more efficient, economically sustainable, and better integrated and governed, but are more sceptical of experimental publishing (e.g. Maryl et al. 2020).

There are some good reasons for this. The basic argument for ebooks to continue resembling the transitional codex format with fidelity (but to the detriment of innovation harnessing the digital affordances of communication) is the resilient allure of the printed book. It is the printed monograph, for which the ebook stands as the digital equivalent (although this is contestable, see below), that maintains an elevated position in the hierarchies of academic prestige; but also scholars treat it with a greater sense of ownership than is the case of journal articles, despite the fact that ebooks have, in fact, a much wider distribution (Deegan 2017, Crossick 2015).

Second, to the degree it cannot be accommodated by established workflows, experimental scholarly writing is more costly to produce and demands more effort from both publishers and academics, who are already hard-pressed enough (Deegan 2017, Jubb 2017).

Despite being more cost- and labour-intensive, the products of experimental writing are not likely to be visible to the current systems of measurement and evaluation of academic prestige because they are ill-calibrated to measure non-standard research outputs, and because the expertise needed to evaluate innovation is scarce. For example, double-blind peer review is a dubious proposition in the case of projects that

develop openly online from scratch, and, more broadly, the combination of expertise in a given discipline and the technical expertise needed to perform such a review translates into a limited pool of reviewers (Adema, Mars, Steiner 2021).

The visibility problems also concern discoverability, as library catalogues, academic publication aggregators are not geared towards cataloguing experimental, web-based projects. Making books from smaller presses, and open access books more visible in the discovery systems is enough of a challenge, while metadata formats (MARC, ONIX), originally designed for the movement of journals and books in library and bookstore supply chains, would have to adapt to better represent experimental content (Jubb 2017).

Finally, it is not certain whether publishing projects that take full advantage of digital technologies are something the average reader wants to engage with. For example, in the case of scholarly editions, it has been argued that while digital technologies provide infinitely more possibilities for the exploration of different variants of the same text, and that more witnesses can be included, this is to the detriment of the stability of the base text, which for many users remains the greatest value of the scholarly edition, and which few want to give up in favour of interacting with a vast and changing archive of variants (Deegan 2017).

Looking from the end of the remediation spectrum that is more favourably disposed towards experimental publishing, some of these arguments are contested.

First, even the sceptics of experimental publishing admit that the fidelity between the established ebook forms and the printed book is an illusion, as print has affordances such as ease of navigation and reading retention that is superior to the average ebook. In truth the ebook often degrades the digital manuscript submitted for publication, for example, by removing hyperlinks to external content added during the production process (Jubb 2017).

Second, advocates of experimental publishing argue that a truly open scholarly communication should not be bound by measurement and evaluation systems. Rather, these systems should adapt to accommodate new genres and formats, and judge them on their own merits (Adema Mars, Steiner 2021).

Third, while at both ends of the remediation spectrum there is a consensus that the governance of open, scholar-led infrastructure should be driven by ethics of care, it is not quite certain whether care and scale will seamlessly come together (Maxwell 2019, Adema et al 2021, Moore 2019, Eve and Gray 2020). To date, care has mostly driven the small communities of developers, users, and enthusiasts gathered around specific tools. It is not quite clear whether the open infrastructures of the future can foster care in a similar way.

Finally, interoperability between elements of the infrastructure of the scholarly communication of the future can be achieved in a variety of ways. Rather than through a preset stack of software operated from a single graphical user interface, which is more user friendly but can lead to “siloisation,” it could be achieved through advancements in programming skills. It is characteristic, in this regard, that while Maxwell (2019) excludes “ad-hoc toolchains,” from his review of innovations in scholarly publishing, which “do not in themselves constitute OSS projects on the scale

with which we are concerned here," the COPIM team advocates that adopters of experimental publishing should do exactly that, for example, to familiarise themselves with the basics of command line interfaces to be able to create customised toolchains for their projects (Adema et al, 2021). Similar, low-investment but computing skills-intensive, distributed, and agile solutions are advocated by the "tactical infrastructure" approach (see: Sheratt 2015).

2.4 Conclusion

Both ends of the remediation spectrum have extremes that should be avoided. Scaling up can mean consolidation around a small number of projects that may, in turn, decrease the innovation potential of the entire ecosystem. "No one wants a Soviet-style, centrally planned scholarly infrastructure" writes Maxwell (2019) "Similarly, there is considerable concern around the spectre of corporate-style consolidation. Indeed, this is the scenario that led to the idea of community-owned infrastructure in the first place." On the other hand, when boldness in experimental publishing comes with boundary work that sets stringent criteria of what counts as truly innovative and open (e.g. advanced computational competences, or certain software architecture), it becomes counterproductive to the cause of fostering more intense engagement with digital technologies in scholarly communication, a problem that applies to digital humanities more broadly (Antonijević 2015). The future of scholarly writing will be decided somewhere in between.

2.5 Bibliography

Adema, Janneke, Marcell Mars, and Tobias Steiner. 2021. *Books Contain Multitudes: Exploring Experimental Publishing*. A COPIM WP6 Research and Scoping Report. DOI: 10.5281/zenodo.4471572

Andrews, Penny C.S. 2020. "The Platformisation of Open." In *Reassembling Scholarly Communications.* Ed. Martin Paul Eve and Jonathan Gray. Cambridge, Mass: MIT Press, 2020. doi:10.7551/mitpress/11885.001.0001

Antonijević, Smiljana, 2015, *Amongst Digital Humanists*. Basingstoke and New York: Palgrave Macmillan.

Aspesi, Claudio, and Amy Brand. "In Pursuit of Open Science, Open Access Is Not Enough." *Science* 368, no. 6491 (May 8, 2020): 574–77. https://doi.org/10.1126/science.aba3763.

Bacevic, Jana, and Chris Muellerleile. "The Moral Economy of Open Access." *European Journal of Social Theory* 21, no. 2 (May 1, 2018): 169–88. https://doi.org/10.1177/1368431017717368.

Bolter, Jay David and Richard Grusin (2000) *Remediation: understanding new media*. Cambridge: MIT Press.

cOAlition S. *Plan S: Principles and Implementation*. Accessed April 20, 2021. https://www.coalition-s.org/addendum-to-the-coalition-s-guidance-on-the-implementation-of-plan-s/principles-and-implementation/.

Crossick, Geoffrey. 2015. *Monographs and Open Access: a Report to HEFCE.*

http://www.hefce.ac.uk/media/hefce/content/pubs/indirreports/2015/Monographs,and,open,access/2014_monographs.pdf.

Deegan, Marilyn. 2017. *The Academic Book and the Future Project Report: A Report to the AHRC and the British Library*, London.

Eve, Martin Paul and Jonathan Gray. 2020. Introduction to *Reassembling Scholarly Communications: Histories, Infrastructures, and Global Politics of Open Access*. Ed. Martin Paul Eve and Jonathan Gray. Cambridge, Mass: MIT Press, 2020. doi:10.7551/mitpress/11885.001.0001

Eve, Martin Paul. 2020. "Reading Scholarship digitally." In *Reassembling Scholarly Communications*. Ed. Martin Paul Eve and Jonathan Gray. Cambridge, Mass: MIT Press, 2020. doi:10.7551/mitpress/11885.001.0001

Fitzpatrick, Kathleen. 2020. "Not All Networks: Toward Open, Sustainable Research Communities." In *Reassembling Scholarly Communications*. Ed. Martin Paul Eve and Jonathan Gray. Cambridge, Mass: MIT Press, 2020. doi:10.7551/mitpress/11885.001.0001

Fyfe, A., Coate, K., Curry, S., Lawson, S., Moxham, N., Røstvik, C. M.. 2017. "Untangling Academic Publishing: A History of the Relationship between Commercial Interests, Academic Prestige and the Circulation of Research." University of St Andrews. doi: 10.5281/zenodo.546100.

The Guardian. "The Guardian View on Academic Publishing: Disastrous Capitalism." *The Guardian*, March 4, 2019. http://www.theguardian.com/commentisfree/2019/mar/04/the-guardian-view-on-academic-publishing-disastrous-capitalism.

Tim Hitchcock, 2016. *What is the Point of the Academic Book? Part Three*. Accessed 20 April 2021. https://ucldigitalpress.co.uk/BOOC/Article/1/47/

Holmwood, John. 2020. "Open Access, Publicity and Democratic Knowledge." In *Reassembling Scholarly Communications*. Ed. Martin Paul Eve and Jonathan Gray. Cambridge, Mass: MIT Press, 2020. doi:10.7551/mitpress/11885.001.0001

International Science Council. 2021. *Opening the record of science: making scholarly publishing work for science in the digital era*. Paris: International Science Council. DOI: 10.24948/2021.01

Jubb, Michael. (2017), *Academic Books and their Futures: A Report to the AHRC and the British Library*, London.

Maxwell, John W, et al. 2019. *Mind the Gap: A Landscape Analysis of Open Source Publishing Tools and Platforms*, mindthegap.pubpub.org | DOI 10.21428/6bc8b38c.2e2f6c3f

Maryl Maciej, Marta Błaszczyńska, Agnieszka Szulińska, and Paweł Rams. "The case for an inclusive scholarly communication infrastructure for social sciences and humanities" [version 1; peer review: 2 approved]. *F1000Research* 2020, 9:1265 (https://doi.org/10.12688/f1000research.26545.1)

Moore, Samuel A. 2019. *Common Struggles: Policy-based vs. scholar-led approaches to open access in the humanities*. PhD Dissertation, Kings' College London. http://dx.doi.org/10.17613/st5m-cx33

The Multigraph Collective. (2018). *Interacting with Print: Elements of Reading in the Era of Print Saturation*. University of Chicago Press. https://doi.org/10.7208/chicago/9780226469287.001.0001

Penier, I., M. P. Eve, and T. Grady. 2020. *COPIM – Revenue models for Open Access monographs 2020*. https://doi.org/10.5281/zenodo.4011836

Readings, Bill. *The University in Ruins*. Cambridge, Mass: Harvard University Press, 1996.

Sherratt, Tim (2015) “Towards a Manifesto for Tactical DH Research Infrastructure”, UC Digital Humanities, YouTube, 2 November 2015, https://www.youtube.com/watch?v=FL5pP2ysjU4

3 Interviews: SSH scholars on innovation

3.1 METHODOLOGY

3.1.1 Preparatory phase

On the basis of the early findings of the literature review, researchers collaborating on Task 6.5 and 6.6 worked iteratively to define those key concepts of the study that were pertinent to the future of scholarly writing. First, in order to open-up the study to various materials, we agreed to treat the “scholarly text” broadly, not only as a linguistic articulation, but rather as an expression that could use different media. Second, we prepared working definitions of the main concepts pertinent to the task:

- Communicating – the act of sharing a text through various formal or informal channels;
- Specificity of SSH – scholarly communication practices in Social Sciences and Humanities that are different from other fields;
- Writing – the act of generating a scholarly text, understood as the expression of an argument backed by some evidence, that may use different media formats and genres, and connect to other texts;
- Collaboration – collective activities undertaken in writing, communicating, publishing, and peer-review;
- Tools – services and software used in the process of writing, communicating, and publishing at various stages of the researchers' workflow;
- Publishing – the act of disseminating a text through a formal process, including intermediaries (publishers, reviewers);
- Innovative forms and genres – text used by scholars to transmit their argument that is beyond traditional formats of the journal article, book, report etc.;
- Audiences – different communities, academic and not, that potentially engage with scholarly texts and their authors;
- Evaluating – the critical assessment of the products of all types of scholarly communication, i.e., writing, communicating, publishing;

- Innovative forms of peer-review – peer-review practices going beyond the commonly accepted forms to address the perceived deficiencies in the system;
- Academic prestige – the widespread respect attached to certain practices by scholarly communities;
- Power structures – dynamic systems of hierarchy and influence in scholarly communication.

Simultaneously we thought about the types of stakeholders we wanted to interview. We wanted representatives from various roles in our research sample: scholars, innovators, reviewers, publishers, librarians, and editors.

3.1.2 Interview scenario

We prepared one scenario for interviews with researchers who performed various roles in scholarly communication. The final version was the result of a pilot phase in which four interviews were conducted.

The inspiration for the interview questionnaire was the methodology of episodic interviews.[1] The main premise of the method is that human knowledge is narrative and that human experience is constructed narratively. Meaning is thus negotiated both internally (internalisation of dominant narratives, reference to existing norms and values) and externally (in interaction with others, which allows the narrative to be contextualised, as others accept or dismiss it). From this the concept of episodic and semantic knowledge is derived

- **episodic knowledge** (particular, based in instances of using, e.g., digital media in conducting research)
- **semantic knowledge** (internalised knowledge, generalisations such as "Scientists should use digital media")

The sections of the questionnaire corresponded to the distinction between two types of knowledge. The full questionnaire is available in Annex 2.

3.1.3 Interviewing

3.1.3.1 Procedure

The interviews were conducted by a 10-person research team (representatives of IBL PAN, SIB, UNIZD, DARIAH, Uni Lux, and MWS) between April and August 2020. Due to the COVID-19 emergency, the initial plan to conduct the interviews in person had to be altered and most of the interviews were conducted online.

The team worked in accordance with the guidelines collected in the interviewer handbook that was shared with all members of the research team, detailing all steps of the procedure and ensuring that all interviews would be conducted in the same manner. Most of the interviews were conducted in English, some in national languages (these interviews were then translated). The full transcripts of 32 interviews (31 single and 1 double interview) were available for the analysis, and 9 further interviews were summarised and made available at the project's repository. A consent form was signed by each interviewee in compliance with GDPR. All citations are marked with an

[1] Uwe Flick (2009) An introduction to qualitative research. Fourth edition Sage, pp. 185–191.

interviewee code in brackets, e.g. (OP01). The list of codes with basic data about the interviewees is available in Annex 1.

As we focused on the content of the interviews rather than on analysing the interviewees' behavior, we adopted simple transcription conventions. Repetitions of words or phatic utterances were omitted for the sake of clarity. Interviews carried out in national languages were machine-translated into English with DeepL and manually corrected. The interviews were not proofread for English, hence some mistakes may be present both in the transcriptions of non-native speakers talking in English, and in the case of translations. We corrected some mistakes in the citations to make them understandable as they were without the context of the entire interview.

3.1.3.2 Research sample

The research sample containing 33 interviewees includes:

- Gender: Male (19), Female (14)
- Career stage: PhD candidate (5), ECR/Post-doc (up to 12 years after PhD) (11), Senior (13), Other (4)
- Countries: Bulgaria, Canada, Croatia, the Czech Republic, France, Germany, Hungary, Ireland, Latvia, Luxembourg, the Netherlands, Poland, Switzerland, the UK, and the USA
- Disciplines represented: arts and media, biblical/religious studies, cultural studies, digital humanities, education/computer sciences, English studies, history, information and communication science, linguistics, literature and literary anthropology, philosophy, psychology, science studies, sociology, other SSH.

When selecting the sample, we were also careful to include interviewees from a variety of roles, such as researchers, authors, book editors, journal editors, publishers, reviewers, librarians, data curators, and software developers. Each interviewee represented more than one role. We were also careful to include both people who were engaged in innovative scholarly communication and scholars with more traditional communication workflows.

3.1.4 **Coding and analysis stage**

The coding and analysis were performed by a six-person team (representatives of IBL PAN, UNIZD, and DARIAH). The coding and analysis team met regularly every week to discuss all important and problematic issues. The team worked with MaxQDA. The coding was divided into three stages.

In the first coding cycle we applied provisional coding to the material using codes that corresponded to the topics defined in the preparatory phase (see above). In the second iteration, transcripts were coded with a provisional coding scheme developed by the team on the basis of the interviews. Each team member was responsible for a particular topic and coded the excerpts identified in the first cycle as corresponding to this topic. This cycle also used descriptive coding to address issues not covered by the provisional coding scheme. Once this cycle of coding had been completed, IBL PAN prepared a second master file that combined all the codes added by the team

members. The third coding cycle gave more freedom to researchers working on a particular topic to recode the material and introduce subcodes pertinent to their topics. Further analysis and work on the report was conducted in previously established and assigned subject areas.

3.1.5 **Documentation**

The documentation for the project along with those interview transcripts that were approved for publication by the interviewees are available in the Nakala repository (https://operas-p.nakala.fr). The team aims to publish all approved transcripts before the end of the project.

3.2 SCHOLARLY COMMUNICATION IN SSH

3.2.1 **Specificity of scholarly communication in SSH**

As this report aims to assess the practices and needs of the SSH community, we spoke with scholars from these disciplines. Scholarly communication is, to a certain extent, always rooted in disciplinary needs and traditions, hence our interviewees often defined their activities in contrast to other fields. What is interesting is that those differences were defined both in transdisciplinary (SSH vs. natural sciences or computer science) and interdisciplinary terms (disciplines in social sciences vs. humanities, and differences between the humanities' disciplines). Interestingly, interdisciplinary research communities in the humanities like digital humanities, were defined as having unique needs: "we are a community of makers; a community in which the tools, and the dexterity and ability and mastery of how well we use these tools, and how we do things that are smart and clever, are important" (OP15); or, "in an ideal world, a scientific paper, I imagine, would include everything – the code and the data – so I could basically run the code and get the results that are published in the article […], but it's not a solution that would work in every case because – especially in SSH – many people don't do quantitative stuff" (OP16).

Hence, although in this report we tend to distinguish SSH practices, we often juxtapose the utterances of people from different disciplines, at different career stages. We are aware that SSH is not a monolith, but nevertheless we tend to address the needs and practices that are common for this field. And in cases where larger differences exist, we signal it by naming the interviewee's discipline (readers can also check this for themselves in Annex 1 at the end of this report). What follows is an attempt at interpreting the specificity of SSH scholarly communication.

One of the key differences lies in the **main genre of communication**, as the natural, or social sciences, tend to communicate using journal articles, while in the humanities the monograph prevails: "in literary studies and the arts […] you need to write lots of books and you need at least one book to even be considered for any type of academic progression. In computer science, it's conference papers and journal papers; but conference papers are almost seen as higher than journal papers because in computer science, it's cutting edge latest work and it's published fast" (OP06). In SSH, publishing a monograph is an important step in career advancement: "Let's say that it's clear that in disciplines like mine, in history and even in sociology, anthropology, [and] literature, publishing a book from your thesis, a book from your habilitation, these

are important elements in a career" (OP25). And although, sometimes, as in the following German case, one is theoretically allowed to use other genres for habilitation, the actual expectations are clear: "So if it's not a book, it could be a series of articles, let's say, but normally it is expected to be a monograph" (OP19).

As we can see, the genre corresponds to the features particular disciplines value most: in the case of sciences it is the timely reporting of facts, while the humanities value the depth and breadth of the interpretation. It should be added, though, that when it comes to the **temporal aspect,** SSH outputs may not need to be published urgently but they do tend to retain their relevance longer: "I guess in SSH it's a little bit different also, because journal articles have a much longer life, they stay relevant much longer. So an article would typically have citations, definitely [for] the first 10 years or so after it's published, which is a really long time. I mean, if you look at STEM, it's usually just one year or something" (OP26).

Writing in the humanities is more **subjective**, rooted in context, less formally structured: "in natural sciences, [...] you have a rather clear structure of discussion and research and so on and so forth. […] Work in humanities is much more subjective […] and the personalities of scholars are expressed much more in their writing" (OP01). This difference also stems from different methodologies, as in this case of juxtaposing the technology-driven approach of natural language processing with the individual knowledge of the scholar: "pure humanities hermeneutic research is more difficult, of course, because they are articles that are based on somebody's personal history of literature, of learning, reading, arguing" (OP08). These disparities are also reflected in the length of SSH outputs compared to other sciences that have more structured outputs: "in SSH, you have to have 30 pages to explain something, but in physics, there are equations and therefore you can explain it in six pages" (OP25).

This is why **peer-review** in those disciplines may serve different functions, as in the case of the sciences, reviewers engage with facts, whereas in SSH they assess the value of the argument: "I guess in the humanities, peer review provides an entirely different function to what it might do in other subject areas. And for me personally, peer review feels more like a dialogue" (OP09).

It seemed evident to our interviewees that the natural sciences and computer sciences are more advanced in introducing **open science** than the humanities (e.g. OP07). This discrepancy is largely attributed to two main issues: timing and funding. First, given the slower pace of SSH, researchers do not see additional value in paying APCs to have their articles available sooner: "So, to be honest, in the humanities, at least in my field, just wait for a year. I mean, instead of paying 3000 euros, I really don't see the point, at least in my field. I don't know" (OP02). Second, the APC model was considered to be "a STEM product" because these "authors have more money. So they are able to pay such an APC." (OP26). And, to put it bluntly, "I think that in the humanities, there's never any expectation that you will be paying to publish somebody just simply because we have no money" (OP10). This funding disparity is also rooted in the perception that natural sciences, in general, have more funding due to the public benefits of their research and their connection with industry: "the problem with the humanities is that there are no industry partners to speak of that could finance some

of these things, because people who are in creative industries see no value in financing research in the humanities" (OP01).

SSH, and the humanities in particular, are said not to be used to **teamwork** to the same extent as the natural sciences: "Even the way our institute is organised is, in a way, built around research groups but in more traditional departments and stuff... the work of individual scholars seems to me very disconnected" (OP16). This also has an impact on **collaboration and authorship attribution** practices in SSH, as they differ across disciplines and humanities: "So in computer science, it's pretty much a given that the supervisor's name goes on any work that is produced during the PhD […] whether or not they actually wrote the paper or whether or not they actually contributed ideas or whatever their involvement is […]. In Digital Humanities, I have seen that it seems to be more common that there's two or three authors normally collaborating […]. And then in, literary studies, it would be a single author – and, you know, you might acknowledge a supervisor in a footnote" (OP06). Finally, particular disciplines differ in terms of the **writing tools** they use, which is connected to general disciplinary practices and will be described in greater detail in the section on writing.

3.2.2 **Summary**

- There are differences in scholarly communication between SSH and other disciplines, as well as within the disciplines of SSH. These concern issues ranging from the genres of outputs and the aims of peer-review, to collaboration strategies and funding.
- The main communication genre reflects the features particular disciplines value most: in the case of the sciences it is the timely reporting of facts, while the humanities value the depth and breadth of the interpretation. Natural, or social sciences, tend to communicate with journal articles, while in the humanities the monograph prevails.
- Writing in the humanities is more subjective, rooted in context, and less formally structured.
- Peer-review may serve different functions in various disciplines. In the case of the sciences, reviewers engage with facts, whereas in SSH they assess the value of the argument.
- Natural sciences and computer sciences are seen as more advanced in introducing open science than the humanities. This discrepancy was largely attributed to two main issues: timing and funding.
- SSH, and the humanities in particular, are said not to be used to teamwork to the same extent as the natural sciences.
- Collaboration and authorship attribution practices in SSH differ across disciplines.

3.3 SCHOLARLY WRITING

3.3.1 **Digitally-enabled writing and digital writing**

We tend to think of writing, especially in the humanities, as the solitary practice of individual scholars wrestling with big ideas. The general evolution of research practices and widespread use of digital technologies makes this image problematic and draws our attention to how **writing is, in fact, a deeply social and technologically supported activity**. As one interviewee put it: "The broader point for a lot of these tools is they are taking us away from this traditional image of St. Jerome sitting in his study, surrounded by his books. This has become the dominant image for scholarship, a lone scholar doing his or her work. I think the digital tools and innovative scholarship allows for collaborative projects. As the tools and methods become more complex, by definition, you need collaboration because one person cannot do everything. I think this is a great role that libraries and centres for digital scholarship can play in helping scholars take their ideas and translate them into new formats" (OP24).

The image of the writing practices of SSH scholars that emerges from our study is greatly removed from the common misconception that all they really need to do their work is a word processor. It is far more complicated than that. First of all, **scholars use many tools within a workflow**: “I use [...] various tools. I like to use light tools that are not complicated or sophisticated enough to make ‘noise’ or mistakes that could be prevented otherwise” (OP29).

The main goal of writing is to codify and transmit one’s thoughts and argument: “You have all the space that you need to tell your story and convince your reader that you – that what you are trying to say is plausible” (OP29). Writing, to begin with, is understood to be a basic – but also the most time-consuming – part of the scholarly communication process (OP20). **This is an activity that takes time**, which has to be foreseen and then allocated: “I try to work in such a way that when I agree on something, I allocate time [to it], for instance, two working weeks – because, for example, that's what I need” (OP13). And, of course, these deadlines need to be updated along the way: “I wanted to finish early 2019, but I’m not there yet! In this case it is not related to the choice of publication model, but it’s only that writing takes time” (OP21).

The main issue of interest in this part of the report is how novel technologies support the process of forming an argument and delivering it. In this respect **we consider writing to be a different activity to publishing and communicating research, however, taking a closer look, the boundaries are blurred**. Moreover, this activity can be undertaken individually or collaboratively. We will discuss these issues in the following subsections, but focus here on how writing was understood by our interviewees.

In general, the responses were rather down-to-earth without much theorising on how new media support thought-forming and transmission. Instead, we had a deep insight into the more practical side of things, namely, how scholars harness existing technologies to achieve their goals and where their individual preferences for certain tools stem from.

We may distinguish here between **digitally-enabled writing** and **digital writing.** The former refers to writing as a textual practice supported by various digital tools like note-taking apps, word processors, CMS, etc.; while digital writing harnesses the full potential of digital technology by integrating different kinds of materials like data, visualisation, or pictures in a single output. The relationship between scholars and their tools will be described in greater detail in the Tools section. Here we will just briefly discuss the role of the tools in the digitally-enabled writing process, unpacking the complicated relationship between writing and technology.

One intriguing insight from our interviews is that scholars appear to use **many different tools in their unique workflows, which are deeply rooted in individual preferences and experience, type of project, and disciplinary needs**. Scholars tend to switch between tools throughout the writing process (e.g. OP14). This choice depends on the scholar's competencies and purpose, which could be to write individually or in a team (to be covered extensively later on). Also, **the length of the text and complexity of the project play a role**: "as soon as the text becomes a bit complex, if you have lots of references or an index, or even if you're working on a bigger document like a dissertation, well, the software at some point will choke, will [start] misplacing figures, the formatting will get lost, etc., etc." (OP17). Finally, **some particular disciplinary needs may call for the use of specific software** as in the case of a psychology researcher who needed a tool to model graphs (OP29) and a writing tool allowing for the automatic application of the particular styles demanded by publishers (OP29).

Digital writing, on the other hand, refers to new modes of expression that use other materials and formats beyond text, like data, software code, or dynamic visualisations. In other words, digital writing allows various media forms and different outputs to be incorporated into the writing practice: "what changes a lot, for me, with digital writing, is also the capacity to provide access to primary data, whatever the discipline, which have a completely immeasurable capacity [compared] with traditional scientific writing. And here I think there's really a key, because it brings together epistemological issues, data verification issues, open science issues, and writing issues" (OP25). This could be done not only through simple access to the output but also by providing interaction with data throughout the writing process: "I tend to go back and forth between the text editor and all the other materials. So maybe I'm using software to produce a network visualisation. […] So I will produce the image and then go back to the text and integrate it" (OP17). Tools support writing activities but may also slow down innovation: "in library and information science […] the relation between data and writing is still a bit conflictual because people write in Word and there's no way to integrate your statistics or your lines of code nicely, or to have good synchronisation between the data and the text you're writing, or to provide interaction between the text and the reader" (OP17).

The difference between digitally-supported writing and digital writing is increasingly blurred as scholars explore new modes of communication. Digital writing is very much connected with the innovative genres we describe in the next chapter.

3.3.2 **Tools**

3.3.2.1 Typology of digital tools used by SSH scholars

This typology[2] is based on the tools mentioned by researchers in the interviews. The names of the tools were extracted using Named Entity Recognition software, delivered by CLARIN-PL[3], and analysed with the support of MaxQDA.

Digital tools used by SSH researchers in their work:

- writing tools – Microsoft Word, Atom, Open Office, Google Docs, Scrivener, LibreOffice, StackEdit, Jupyter Notebook, iA Writer, Overleaf
- data storage and sharing tools – Google Drive, Dropbox, Microsoft SharePoint, Next Cloud, OneDrive
- tools for presenting work results – Microsoft PowerPoint, Adobe Connect, yEd Graph Editor, Adobe Acrobat
- tools for developing, presenting and sharing concepts and ideas – Miro, Padlet, FreeMind, Microsoft MindManager
- reference and citation management tools – Zotero, EndNote, JabRef, Citavi
- tools for planning and monitoring work – Asana, Trello, Google Calendar
- tools for managing notes – Evernote, OneNote, Reminders
- tools for managing and sharing scientific papers/database research – Mendeley, Google Scholar, ResearchGate, Academia.edu, Scopus, Web of Science, SciHub
- typesetting tools – LaTex, Markdown, Pandoc
- data processing and analysis tools – SPSS, Microsoft Excel, nVivo, MaxQda, Atlas.ti, Stylo, CLARIN tools for NLP
- communication tools – Zoom, Skype, Microsoft Teams, Google Meets, WhatsApp
- programming tools – GitHub, Python, RStudio, JavaScript, Eclipse, Visual Studio Code
- online translation tools – Google Translate, Dictionarycambridge.org

3.3.2.2 The phenomenon of Google tools

Among the tools for collaboration reported by the interviewees, Google tools were especially popular. They are available, easy to manage, enable people to work together in real time, and most importantly, they are familiar to all writers – those with more advanced digital competences and those with less, and those who know how to use more advanced digital tools and those who do not. For this reason, even though they are not flawless, these tools are the first choice when it comes to writing a text together. They are best described by the phrase "[s]ufficient, not efficient,"

[2] The typology presented here is qualitative. *Innovations in Scholarly Communication* report (https://101innovations.wordpress.com/) provides an overview of the tools used by scientists, based on data from the global survey of scholarly communication tool usage.

[3] Maryl M., Piasecki M., i T. Walkowiak (2018) "Literary Exploration Machine A Web-Based Application for Textual Scholars," in Selected papers from the CLARIN Annual Conference 2017. Ed. Maciej Piasecki, Linköping University Electronic Press, pp. 128–44.

mentioned by Antonijevic and Stern Cahoy as the guiding motto for scientists who choose digital working tools (Antonijevic, Stern Cahoy 2018).

> **Google, in some sense, is evil, we all know it**, you know it: their tracking system etc. But what they do get right is a sort of a sweet spot between functionality and overkill of functionality, that they're really good at, seeing how users are actually using their stuff; and stuff that doesn't get used is just thrown out in the next version. (OP08)
>
> It actually depends on whether I'm writing it by myself or within a team. [...] And we usually actually use Google Docs, Dropbox sometimes, but we have a few problems with synchronization. And **Google Docs just works quite well in the workflow**. (OP02)
>
> **For collaborative writing**, if it's a very intense interaction among people and we are going to be very fast, then we typically use Google Docs and Writer, because it's easier to [...] collaborate in those versions. (OP30)
>
> More and more people use Google Docs because the **commenting and collaborative stuff goes much better**. (OP05)

3.3.2.3 Digital tools for workflow planning and monitoring

We specifically asked the respondents about their use of tools for planning and monitoring work.

- **Trello and Asana** are used by people who work on projects with many co-workers (usually on a few projects simultaneously).

> I use Trello, and I've been using Trello for four or five years. I'm really happy with it. I've just discovered the to-do text file workflow, which is basically – you have just one file, which is a simple plain text file, and you'll put each task on one line, and [then] some programs can handle that file and provide a convenient interface for, maybe, setting a due date, or things like that. (OP17)
>
> I use a tool, an on-line platform called Asana in order to manage the whole project, so everything becomes like a task or a sub-task in Asana. So I set up a task for this publication, and many, many sub-tasks, in order to manage anything from analysing [to] interviews with these people in these two archaeological projects. (OP15)

Overall, digital tools specifically dedicated to facilitating the division of labour were rather seldom used; instead, the work was organised and planned.

- **In email exchanges and while making arrangements via online communication tools.** "We use certain conferencing tools, anything from Adobe Connect to Zoom or Skype […] And e-mails are, of course, the basic means of communication" (OP30).
- **Through a shared calendar.** "I think calendars are a good keyword. There are also tools, for example, for organizing work, especially if you do it collaboratively, where you can set deadlines in the tool, where you can also assign tasks, for example, where employee 'A' receives an e-mail: Okay, I have to scan this book or something by tomorrow" (OP20).

- **With tools not dedicated to this particular task**, for example, Microsoft Excel, Microsoft Word, yEd Graph Editor, or Reminders. "Log book, a dumb Word file" (OP27).

For some, planning and monitoring work was the opposite of the creative process.

> It's kind of very unsystematic. And that's sort of how I work. I don't like to work in a way that is, sort of, that planned because I feel like a lot of my ideas come from happenstance and from, kind of, my engagement with the text in a sort of – just sort of immersing myself in it I suppose, and surrounding myself with it, which means that I don't have a particularly good note-keeping system. (OP10)

Others reported they needed to invest a lot of time to master the skills required to use the planning and work monitoring tools.

> I was thinking about this, especially in the context of my research projects. Something like a planning tool, organizing tool for teamwork, for researching. I did try one, which I can't remember the name of, a freeware [tool], but then I discovered that for me it would just be too much. OP31

3.3.2.4 Bricoleur and the Engineer as models for using digital tools

In order to better capture the relationship between researchers and tools we will use the figures of the engineer and bricoleur introduced by Claude Lévi-Strauss and elaborated for the DH context by Smiljana Antonijevic and Ellysa Stern Cahoy[4].

> The "bricoleur" is adept at performing a large number of diverse tasks; but, unlike the engineer, he does not subordinate each of them to the availability of raw materials and tools conceived and procured for the purpose of the project. His universe of instruments is closed and the rules of his game are always to make do with "whatever is at hand" [...] His (engineer) means, power, knowledge are never unlimited and that in this negative form he meets resistance with which he has to come to terms. (Levi-Strauss 1966, 17–18)

Among our interviewees there were engineers as well as bricoleurs. The engineers are experts in many specialist tools, which they fluidly switch between depending on whether they work alone or in a team, and on what their teammates' competences are. They strive for perfection, and gladly use more advanced tools. Thus, flexibility with regard to choice of tools could be considered an extra competence.

> When many people write a text, it depends on the group I work with. When they're humanists, I use Google Docs, and that's where we work. Sometimes, unfortunately, I still work on files exchanged as attachments. But even then, I

[4] Lévi-Strauss, Claude. The savage mind, Chicago 1966, pp. 17-18. Cf. Antonijevic, Smiljana and Ellysa Stern Cahoy: "Researcher as Bricoleur: Contextualizing humanists' digital workflows". DHQ. 12(3) (2018).

> try to use Zotero if these are people from my research group. I also use Overleaf to write texts in LaTeX. We also have a paid version. And we also work with part of the team, because I have an eight-person team and different people have different backgrounds and different competences. I try to adapt – so when I have a PhD student who is writing his PhD in philosophy, I don't wear him out with Overleaf. But if I have a PhD student who is writing his PhD in Computer Science – Overleaf is obligatory. So I'd bring it down to Word, Zotero, Google Documents and Overleaf; plus a basic PDF reader. (OP13)

Bricoleurs, on the other hand, still combine digital practices with offline ones. They gladly conceptualise their work with a pencil on paper, and they use digital tools mostly because, nowadays, it's absolutely necessary. Sometimes they don't have full access to the sources they need, but they can deal with this as well.

> I'm basically using the things I have at hand. So for example, if I suddenly just have a spark and I have a great idea, I just write it down in a notebook on my phone if I'm not at the computer – because this notebook will synchronise with my computer, so I have it everywhere right away. Besides this, I make a lot of notes just in Word. And I try to keep an eye on these documents so that I know what is a draft version, and what is not. But actually, I do the whole phase, or almost the whole phase, of inventing an idea on paper, because I think a little bit when writing by hand. And I just do it that way. So all of them are just some brilliant – not only sparks, but also ways to work out a topic, to give it a structure, to come up with a structure, to come up with where it should go next in terms of substance – I make it up with a pen, definitely – but then on the computer. If I'm looking for some texts, I do a bit of a patchwork, because I just download some things right away, if they're PDFs. And I categorise them immediately in catalogues of, let's say, essentials. [...] There are also things I can't download, because I can only read them online [...] It depends on the fact of having access to it or not. If I have access to it, I just write it down in these online book collections in the libraries. If I have access to it by preview only – well, like everyone else, I'm very capable of browsing almost a whole book using Amazon or Google Books preview. (OP12)

3.3.2.5 Sources of knowledge about digital tools

We investigated where knowledge about digital tools comes from, how they are mastered, and what obstacles accompany the process. Interviewees reported the following sources of knowledge about digital tools**:**

- **Self-learning, motivation to improve one's work, knowledge mostly from the internet.**

> I learned by [...] reading stuff on the Internet. That's my main learning platform. (OP29)

> Most of the tools, when I need something, usually come from Googling and just trying to define my problem, and then searching for a tool that answers the

> need that I have, so I compare recommendations and try to find something that fits best with my workflow. (OP16)
>
> So I just started playing with it and then checking the forums for any questions I had on it. (OP06)

- **Research team colleagues, or a specific unit within the institution**: there is always a person within the institutions with broad technological competences, who can suggest solutions and a choice of tools – and who helps with learning them.

> The different stuff like Mendeley or Dropbox I just learned by myself, and sometimes there are younger colleagues who are better at this and then I ask for some tips on how to solve specific problems. There are always people in the research team I can ask for advice or guidelines if I need to. (OP31)

- **Sometimes, in collaborative work, people whose tool expertise is very extensive have to "level down," and adjust to their colleagues who have a lower level of competence.**

> Well, yes, but unfortunately I did not use the tools that I like, or the tools I know how to use. I used the tools that other people know how to use, which is mainly Google Drive and stuff like that. Which is unfortunate, but yeah. (OP29)

- **Knowledge acquired in the course of studying.**

> Basically, almost all the tools [I use], I discovered during my master's degree. It was – I needed to translate the wording – but basically it was a degree in library and information science with a focus on technical and scientific information. And so I got into a lot of tools to organise the work and produce formats and curate content and all that stuff. So basically, I learned about all these tools during my studies. (OP17)

- **Training organised by universities or other organisations.**

> This was training organised by the Ministry, or various foundations. And [in] this way I also instruct doctoral students on courses. There has been training on Scopus, recently the Web of Science has also advertised at the University. (IBL 05)

The digital tools that facilitate work aren't for everyone; the learning process is difficult, especially for older people. Sometimes force of habit prevails over the need to facilitate the work process.

> Yes, some people will use more innovative things and some people will not. I'm surprised in my own department at the faculty […] Despite my introduction of Zotero, despite it becoming pretty much universal, and students using it,

> many of my colleagues don't. They do things by hand […] and it is very simple to use, and extremely useful, and time saving […]. (OP30)

Changing one's habits and learning how to use a tool effectively is time-consuming.

> Because a tool is a tool; and it requires some time to gain the skills to use it efficiently – and I'm not very keen on learning new tools unless they provide something really, really important. So, basically, I will be going for a new tool when I have a significant lack of functionality, and then I'm seeking for [one] and asking around. (OP30)

The selection depends on two key issues: whether a tool is free (or available through licences purchased by the institutions where scholars work) and the understanding of the features provided. The motivation to explore new tools, even if they are better and more effective, drops significantly if it's yet another tool one has to become familiar with. Force of habit, the bricoleur's perspective of what is at hand, familiar, and tested, is very strong. **That is why the means of communication about tools and mastering them seems so important. It should be as suited to the current workflow of SSH scientists as possible, it should be subject to the creative process.** What is interesting is that the work of American human sciences researchers has led to the same conclusion.

> It is through this analytical lens that we can understand [the] "inefficient" and "unruly" bricolage practices observed in this study. Like bricoleurs in Lévi-Strauss's original account, humanists are constructing their digital workflow with an orientation on creativity and interpretation rather than on efficiency. Embracing humanities researchers as bricoleurs will thus enable us to prevent the quest for computational efficiency from overshadowing the quest for humanistic understanding." (Antonijevic, Stern Cahoy 2018)

3.3.3 **Components of the writing process**

3.3.3.1 Where does writing begin and end?

When we discussed the writing process with interviewees, we noticed how difficult it was to make sharp delimitations between writing and other elements of the research process, as all the components interact throughout.

> I don't really see reading as different from writing within that research process. It's all one kind of flow through, and all the components are kind of glued together. So when you said writing, you know, I immediately went to the word processor and Zotero, but then it quickly expanded out, I thought, actually, you know, and there are loads of things I use that are about this process. (OP03)

Hence, we treat the discovery, storing, reading, and annotating of research assets as part of the writing process, as they influence the writing and have already generated the content that will be used in the drafting phase.

Not only is it difficult to assess where writing begins but also – where it ends. Writing nowadays seems to be a rather open-ended process, allowing thoughts to be released earlier through social media or blogs, collecting feedback, and introducing changes.

> I keep this kind of open notebook in which I'm just sort of sharing my thoughts regularly about my research; and eventually the book that I'm writing will be based on all of these different snippets. But none of the actual blog posts will be in the book. I quite like just releasing my thoughts as I have them. (OP10)

There is a broader question here: which part of the communication process that leads to publication should be considered the end of the writing process? This also concerns the myriad of responsibilities assigned to scholars. This uptake of various roles in the publishing cycle was described by one interviewee in terms of the professionalisation and differentiation of responsibilities.

> I do see it, critically, that science is becoming more and more involved in the publishing business – as in the spirit of the Gutenberg era, where the author does everything else as well. Well, that was more extreme, there is also a process of professionalisation and differentiation, but I don't think it is actually the task of university scientists to take care of publication formats and platforms. Sure, on the content side, no question, but we are not publishers. (OP20)

Moreover, **the writing process appears to be non-linear**. We initially thought about these components as phases of a linear workflow, however, what struck us was that this is an iterative process in which all the components coincide or follow each other in different sequences, dictated by the progress of writing. Of course we may attempt to establish a Popperian type protocol for this activity, starting with selecting, reading, and annotating content, followed by hypothesis formation, evidence collection, and concluding with drafting and editing. Yet, the actual practices are more chaotic and idiosyncratic, hence, we will be describing these components separately. Another reason to treat writing practices as a set of intertwining activities is that the actual order of activities may differ depending on the discipline or actual research project.

> You're talking about [this kind of] research: "We have an idea, you design, you collect data, you analyse, you publish." But most of the things we do in the technical field are not that way. We have an idea, we design, we create things, we test it, and then we see if we need a redesign, and then we redesign and we build and test it, and then we redesign again, and it's endless. And somewhere in the middle, I start talking to you. And you influence my redesign, and then you help me in endless things, and you join the discussion and change things. I think you contributed significantly to the end result. (OP30)

3.3.3.2 Discovery

The discovery part of the writing process depends on the context and the particular needs. **Scholars use various online services to find relevant sources,**

and it is often a mix of search engines, Google scholar, article databases, and university catalogues.

> Well, obviously, like everyone else, I start collecting material using Google Search in general. In coronavirus time, it is difficult for a person to get to a library. (OP05)

> This of course depends on the subject I'm dealing with at the time. But I often start with the Web of Science if there is a lot of material. (OP11)

Google Scholar was also named as being an important source because it allows for the validation of sources through the citations they receive.

> [...] as I look at Google Scholar […] Obviously, I would see, initially, in the first pages of the results, publications that have more citations. However, as I say, I would also look several pages further – several pages down. I wouldn't mind reading a paper if I read the two sentences under the Google Scholar window that look[ed] interesting to me" (OP15).

What is also interesting in this process is the open attitude towards sources: accessing them and checking along the way during the discovery phase.

Discovery is also achieved through a practice one interviewee called "**citation hopping**": "I would read something, and then I would have a reference to something else that I found interesting, so that [then] I would search to find this other interesting piece" (OP15). Often this approach provides an entry point for engagement with scholarly search engines: "What I'll do is I'll begin with, like, four to five papers, and then [go] from their bibliographies. And that's how I work" (OP02).

New content is also discovered thanks to "**invisible colleges**," networks of scholarly contacts referencing or recommending relevant works. It is noteworthy that technology seems to facilitate this very process through various services.

> [...] for me it's going to be Researchgate, and for someone [else] it's going to be, more, Twitter, I don't know – someone is using Academia.edu. And the moment I start following it – if these people post and update their statuses on what they do scientifically – then I follow it. (OP11)

Informal recommendations also serve a **content filtering** function, similar to the one performed by journals (OP13). The very fact that a source is recommended through such channels is often taken as a sign of quality: "if I see an article shared fifty times by my colleagues, it's also a form of peer-review for me" (OP21). Similarly, if the work is indexed in a database it is perceived as credible (OP13).

Another entry point for discovery is **university catalogues**, which are valued mostly because they are connected with large full-text platforms, allowing seamless discovery and access.

> I'll use the Web to search my library catalogue, and to search under other discovery services and to find the things that I need to read in order to deepen my knowledge, in order to write the thing that I want to write. (OP04)

> Our library search system is really good, [there is] the simplest integration between finding something or just getting a PDF or just looking at it. So we get all that. (OP15)

Some scholars (e.g. OP14) use library catalogues to access offline resources too.

However, the seamless download of relevant articles, praised by employees of research institutions with robust search catalogues and paid subscriptions, may be the biggest obstacle in the discovery phase for scholars who lack such access. And this is even more true for some countries than others.

> Here in Europe or in the US you can say that it's not a very big deal if books are very expensive because they are available through the library anyway, and maybe the library is a bit far-away but there is online access. But this is not the reality for everyone, and there are countries where even the libraries don't get enough money for our books. (OP23)

In order to overcome these obstacles, scholars engage with "**shadow libraries**": "so mostly people get information from PDFs, from legal and illegal sources, for example, SciHub or LibGen, everyone uses LibGen" (OP05). One scholar even mastered the art of using free previews on Amazon and Google Books to navigate the copyrighted content for free (OP12). It seems that **access is an integral part of discovery**, as one scholar put it: "[w]hat we need as researchers is free access to literature. Databases are probably the bottleneck of this whole system" (OP32).

Another aspect worth mentioning in the context of this report is the **difficulty in discovering innovative scholarly outputs**, as they are often not included in the sources described above: "making sure that these new kinds of work are included in those catalogues and those databases so that they can be found in the places where scholars actually go looking for work, is going to be key" (OP04). This situation calls for a change of approach in libraries, and the evolution of metadata schemas from monograph-oriented to more flexible ones that are open to new inputs (OP24).

3.3.3.3 Storing and annotating

Another component of the writing process concerns storing and annotating resources. These activities may be discussed separately, however they seem to be closely connected in the workflows reported by our interviewees. PDF versions of texts were usually reported as being a vehicle for storage: "the PDF format still remains important to me because it's nicely encapsulated and I can store a copy offline" (OP03). However, **various tools and services are used for this purpose: bibliography managers, cloud storage, and note-taking applications**. Interviewees store resources in Zotero, Mendeley, and Endnote, using their tagging features for better retrieval.

> I usually pull the results of those searches into Zotero so that I have all of the articles and the citations and everything all in one place. (OP04)
>
> I use Mendeley for all of my research. (OP06)
>
> I then throw the full text in Mendeley, because that's where I'm most comfortable in taking notes. (OP11)

Bibliography managers are also connected with the discovery phase, serving as a repository for discovered content (even if it's not immediately needed): "I'm downloading like crazy on illegal platforms – like crazy. I'm also downloading a lot of things that I know I won't use now. But for later, everything is in my database and when I'm beginning with a new topic, I usually begin with Zotero rather than with Google. I have ten thousand references now" (OP02).

Scholars reported various cloud services they use to store content: "I use for data sharing or – for sharing or for my personal needs – I use OneCloud or Google Drive" (OP16). In one particular example, a sociology scholar used an Amazon basket as personal storage and reference space: "I throw the books into the basket in Amazon, for example, and then I remember that the books that I needed for something, but didn't have access to, are in that basket in Amazon" (OP12). In this way, technology becomes somewhat appropriated for particular uses. Similarly, one scholar reported emailing files to their own address for safe keeping (OP15).

Note-taking apps are usually used for both storage and annotation:

> I use OneNote a lot actually. It's sort of a dump for various memory stuff. I used to do Evernote, but then they made some severe restrictions on that. (OP16)
>
> I take the notes in Bear, the same note taking app that I use. And I usually create Bear lists to organise files – not in folders or notes, not in folders, but by hashtags (OP04).
>
> I used Tropy for archive storage and annotation. (OP28)

For content annotation, scholars use PDF readers:

> I use basic PDF readers, just on a Mac. I have Adobe Acrobat, but I use it to edit PDFs, not to read. So if I read PDFs, it's in an ordinary reader and I underline. (OP13)
>
> I use some tools like Hypothesis to annotate text digitally, and I use the Acrobat program to annotate PDFs. (OP24)

3.3.3.4 Ideation and drafting

This component concerns the "writing proper," namely the very act of "putting the ideas on paper," whether that is meant literally or figuratively. The analysis showed that **many interviewees distinguished between the ideation and drafting phase**, using different tools for each.

Ideation is an activity usually conducted in plain text to keep things simple so it doesn't allow technology to interfere with the thought process. Paradoxically, researchers try to avoid any robust features and use the most basic tools to collect and write down ideas.

> In the ideation phase, I always start with writing in text editors. I always start with pure text editors, where I write the structure of the paper. In combination with the text editor I always use a reference manager. (OP32)

One interviewee reported that the complexity of dedicated writing software did not match their creative process:

> Atom is a plain text editor, or at least you can use it as a plain text editor and nothing gets in your way. I completely grew out of using anything like Word, either on Mac or PC, or things like – what's the Mac tool – Scrivener. I liked the idea behind Scrivener, but it didn't work for me. I just needed a plain empty box to tick in, apparently. And that's just for dumping my ideas, as it were, and developing those ideas further. (OP08)

Hence **the simplest technology seems to be better to think with.**

> [...] when developing ideas, one usually starts from a broad picture of the problem, and that lends itself best to technology that isn't limiting or complex in any way, so having something that is simple helps. I prefer, for initial stuff, – writing down thoughts – I use Notepad++ because it's just a very simple tool that has a lot of powerful options down the line, but basically just to write words without any thought about the format or anything else. (OP29)

This sociology professor used a note-taking app or simple piece of paper and pen to write down and explore ideas:

> If I suddenly just have a spark and I have a great idea, I just write it down in a notebook on my phone if I'm not at the computer – because this notebook will synchronise with my computer, so I have it everywhere right away. Besides [this], I make a lot of notes just in Word. (OP12)

One interviewee reported a telling interaction with tools, whereby the intended functionality was overridden by the scholars' habits. When the publisher released a writing environment for a big collective online encyclopaedia work, our interviewee found its features very helpful, but still did all their writing in another tool and pasted it gradually into the online service:

> I really liked […] how it worked, how you had this encyclopaedia format, […] where you were controlled in your writing, so to speak, in the structuring, in the bibliography, and so on, and everything was calibrated to the final product. […]; of course, I didn't write it online, but I gradually transferred it there. (OP20)

The ideation phase could also be supported by a visualisation, which could be hand-drawn or prepared in a dedicated tool.

> [...] sometimes I would just write down notes manually, like on a piece of paper; or connect words, like a little diagram or something like that. I like to experiment with tools, so sometimes I would use a tool in order to do that, anything as simple as, let's say, creating an outline using Microsoft Word, – which can be one solution – to using some kind of an idea or concept mapping tool like Compendium, in order to structure ideas. (OP15)

> I use software to draw diagrams that I used to make some workflow […] sketches, so it's yEd graph, […]. It's great for setting the basic structure of how things progress from ideas because you just use squares and you write things in those squares and you connect them with arrows and then you have a nice workflow, flow chart, and whatever you need. (OP29)

> [W]e use mind mapping tools to organise ideas, to organise research, and to build taxonomies. (OP30)

3.3.3.5 Formatting

Another step in the writing process is formatting. In this step **researchers rearrange and structure their ideas**.

> So at the point when I move into Microsoft Word, I have a tendency to start introducing the formatting, and introducing footnotes and citations and other things like that. (OP04)

> In the next step, when I want to make text look like something that has structure, I use Markdown, which allows me to edit – very simply – simple bits of the structure. But [the] idea basically – the basic idea is just words, and words don't need any form I think. (OP29)

Some researchers report using only one word processing tool for writing.

3.3.3.6 Referencing

Many interviewees reported using a bibliography manager to support the writing workflow at the discovery, storage, and referencing stages.

> I actually tried Zotero in the beginning. So I found Zotero and then I struggled with getting over the learning curve for that one. And also I found it wasn't as usable as Mendeley, for how I would use the tool. So I chose Mendeley and I just pretty much did the usual, like read on the, like, the tech support [on] how to use it. (OP06)

> If we are talking about decent reference managers, I recommend EndNote, as a commercial tool, as a very good, reliable and straightforward tool. We have Mendeley, Zotero, etc. (OP32)

An important feature of bibliography managers is their **integration with word processors**, which is an added value for the entire writing process.

> I would say that the two primary tools that I use most on a day to day basis are Libreoffice and Zotero, and the same Zotero in a web browser. Those three interlinked components are kind of the core environment within which I do my work. I use the web browser, and the Zotero plug-in to ensure that I have a good set of metadata for things that I'm reading and work that I'm going to cite. And then the plug-in framework for Libreoffice ensures that the bibliography is generated accurately from those metadata. (OP03)

3.3.3.7 Copy editing and proofreading

Scholars struggle to improve their own writing, especially when they are non-native speakers writing in English. Some researchers report hiring proofreaders at additional cost.

> The costs I bear – and I bear them individually – well, I write in English, but I still have someone to proofread it for me. (OP11)

> Of course, when I write in English, I have to take into account that I need a good editor. All my costs of language editing – it depends on the project – but I have also funded some articles myself. (OP18)

Some writers report using online grammar-checking and translating tools.

> I use something called Grammarly, it's a page where I can check my syntax. Dictionarycambridge.org: this is a site where I often search for different synonyms. […] Sometimes I just use Google Translator, because they have better and better suggestions. (OP11)

> I still use Grammarly to check my structure, my style. Grammarly is better than the MS Word spellchecker because there you can define the style you want to write in and it catches things that don't fit the style. Of course, you have to control it. (OP14)

3.3.4 **Collaborative writing**

3.3.4.1 Specificity of collaborative writing

We asked our interviewees about the difference between writing individually and in collaboration, and also about the particular tools they may wish to use in both contexts. As mentioned earlier in the section on disciplinary specificities, **individual writing seems to be more the domain of SSH and the humanities in particular**. Most of the discussion on writing, so far, referred to this type. What can be added is that individual writing seems to be a less organised activity, as it does not depend on cooperation with others.

> [...] when I'm working on my own, I have quite a, sort of, primitive set up, which is just, like, throw as much information into a Word document as possible, sift through it, and eventually that turns [into a text]. That's how my PhD kind of was written, [it] was just a load of different notes, which eventually, kind of, turned into a page. So it's really haphazard. It's, kind of, very unsystematic. And that's sort of how I work. I don't like to work in a way that is, sort of, that planned, because I feel like a lot of my ideas come from happenstance and from, kind of, my engagement with the text in a, sort of – just sort of – immersing myself in it, I suppose, and surrounding myself with it. (OP10)

Interviewees reported **different experiences with collaboration and negotiating a process with multiple authors. For some, it ultimately led to the output having a patchwork structure, whereas others complained about a dry, "impersonal," homogenous language.** A postdoc in information studies gave an account of a writing process in which the tool (Google Docs) allowed for unrestrained, creative collaboration, leading to uneven chapters instead of a coherent, structured book:

> [...] every time we've written together, we just seem to kind of get together for like a couple of weeks and then just, kind of, blitz it. Now, this was mainly led by [researcher's name], who was the main orchestrator of it all. But there's this really fun way of working, which is just, sort of, go to a Google Doc and just see what happens. It's completely unplanned. […] And so what's come out of

> it is this kind of weird, like, book, sort of three massive chapters. It doesn't really fit any other kind of style of publication. It's not like a traditional format. I'm not even sure it's the kind of thing that [name of university press] were expecting or they want to publish. My sense is they [would have] preferred a book that is five or six very short chapters. So we've just given them this kind of weird Frankensteinian monograph. (OP10)

A psychology postdoc described his experiences with collaboration, stressing the negotiation of the language:

> if you have an amount of people collaborating, oftentimes the language is dry because they need to have dry language so they can all connect to that base – they can all connect to the single manuscript they are trying to produce. They need dry language because if everybody was trying to write their metaphors, and be colourful in their expressions, then it would be a mess because everybody would be colourful in their [own] way. (OP29)

3.3.4.2 Reasons to write collaboratively

Writing in collaboration stems from the dynamic of project work, as the same psychology postdoc asserted:

> I write in collaboration because it's something [...], that's how modern scientific projects work, at least in my field. You have a team of people working on a project, and then of course the publications are going to be a team effort. (OP29)

The biggest advantage is the ability to share the workload with others.

> [...] when it comes down to how much you have to work as an individual it lessens the burden significantly. So if you have four people writing on a project, even if you have one or two people tagging along, you still have some work – so can still have some work done by them; and [if] you have another person who is really devoted to the project, they can carry some [of the] weight. (OP29)

It may seem that collaborative writing is, in a way, scholarly communication in a nutshell, where research ideas and arguments are already being debated during the writing phase. This is how a professor of Information Studies compared individual and collaborative writing, underlining the need for communication early on in the process:

> There are some things in research where you need to be like a philosopher, just sit down and think about it. But from my perspective, in the end, I need some people to talk to me about it. And if you talk to someone too much and he's helping you – is he the co-author of the paper, or is he just your sparring-partner? (OP30)

Scholars tend to work collaboratively, and also extend their competencies beyond their field and provide more depth to their research.

> [...] you write about something and then someone reads [it] and says, "I have a data set that can lead us to another step, we can improve this paper by

> adding a new layer of empirical information that could lead to something new." (OP31)

This could also mean **providing an international context**, as for this author who collaborated with scholars from other countries in order to put his own results in a comparative context:

> I think that in this field, which I deal with, it's very important to not only present the Polish perspective. […] In my opinion, these studies are more interesting when it is possible to compare these research results with different countries and draw conclusions from them. (OP11)

Yet in collaboration, there are sometimes **challenges that need to be addressed**, such as reconciling cultural differences among scholars from different countries, disciplines, or working cultures.

> I'm now working on a project where I have co-authors from eight countries – but these are co-authors from such countries as Italy, Turkey, China. So these cultural relations are definitely much more difficult. So negotiating, for example, co-authorship, negotiating who does what, or negotiating recommendations, conclusions [...] (OP13)

There is a clear **difference between the humanities and social sciences in terms of how often work is co-authored**. In the former, collaboration is widespread. A postdoc in psychology asserted that "[e]verything was a collaboration basically. Even the journal papers were a collaboration so nothing, none of what I published, was my own" (OP29). A sociology professor claimed that in the last "three years there were, maybe, just one or two papers that I wrote just as a sole author" (OP31). Conversely, a professor of literary studies admitted: "Most of my work is single authored. Because English literature, my field, often works in that mode" (OP03), while a PhD student in History wrote: "All individually" (OP28).

3.3.4.3 Organisation of work

The crucial element of successful collaboration seems to be the organisation of work, as the process requires team effort, linked to project cooperation: "we discuss not only how to describe things but sometimes we find out that we need to change something in our research, or we get an idea for further research" (OP30). Our interviewees stressed **the need for a feedback culture that requires trust and boldness**:

> [...] collaborative writing in this regard, to me, is very much a quite personal process. I'm kind of lucky to have found [name of the collaborator Y], we write together really well, she's a kind of a fierce critic of my work, which [she] is very good at, kind of, knocking it into shape with other people. […] we can just sort of tear each other's work apart and we're kind of more equals in that regard. (OP10)

A professor of literature invoked the Wikipedia-like principle **of bold collaborative editing as a good way of managing the interaction between co-authors**:

> [...] we had a principle that anybody could just write over the top of what someone else had done and delete their work or change it as they saw fit. And nobody would take offense at that. And if someone really objected to something, we'd go and have a discussion around it. But the idea was, that if you're not bold in your editing, a kind of Wikipedia like principle, then collaboration becomes really, this quite tricky thing, where everyone's just being nice and tiptoeing around each other. When really what you want are people's confident, bold assertions to come through and get the strength of everyone's writing. (OP03)

Trust building, then, is a prerequisite for such collaboration, requiring some level of personal, informal connection. **Collaboration at a personal level may facilitate cooperation or make it impossible** as in the case of this history professor:

> You need to have very strong trust with the person, as we discuss a lot anyway. I also have two articles that I never wrote because I couldn't stand the criticisms that others made and we got angry. But that didn't discourage me and I prefer to write an article in collaboration with someone else. (OP27)

Our interviewees expressed a preference for establishing such relationships in person: “Zoom is cool, MS Teams is cool. But the level of trust that needs to be built up – you can only build up over beer, over dinner, on trips together. MS Teams and Zoom are used to maintain this relationship” (OP13). Establishing cooperation through offline workshops (OP18, Pos.1) or Bookathons (OP22) was another way in which collaboration has been fostered.

3.3.4.4 Lead Author as coordinator

Collaborative writing requires a lead author, who will organise the process and coordinate the team: “We usually agree that one person should be at the front, and communicate with the journal [editor] and such” (OP30). Sometimes, the cooperation hierarchy depends on the writing genre: “In articles, there was often one lead author. For project writing it is more evenly split between co-authors” (OP21). **Distribution of labour is crucial at the early stages**, which seems to be the responsibility of “the first author who was very upfront about what needs to be done and who needs to do what, and I think that's a good thing because you have a very clear picture of what you need to do and [...] Because if everybody does everything, again, things get lost” (OP29).

In some teams, this is the person who provides the first draft, a structure for others to fill out as in the case of this philosophy professor:

> [...] if you want to do something, there has to be someone who will, of course, gather the team, but at the same time declare that he will deliver the first draft. Because this is such a key moment, beyond which people can get involved. (OP13)

This can also happen at a slightly later stage, when a lead author emerges to impose order on chaos, as in the case of another collaboration that started with “anarchy.” However, “the lead author really took over and [...] he wrote most of it. Like, we came up with the, sort of, the structure and the plan” (OP10).

3.3.4.5 Collaborative writing tools

There is an interesting feedback loop between the writing tools and collaborative writing practices. **It would not be an exaggeration to say that all of our interviewees who wrote collaboratively have used Google docs for this process**, which has already been described in the section on tools. Interestingly, it seems that this platform, due to its accessibility and ease of use, may have enabled widespread collaborative writing in the form described above, i.e., as a dynamic group process as opposed to a more static exchange of ideas or drafts by email. A postdoc in psychology elaborated on this notion of what actual collaborative writing is:

> I don't think that before Google Docs we really did a collaborative effort here. We did collaborative writing in the sense that [it] was more of a serial effort at that time, because I would gather text from my colleagues with a bunch of changes, then I would change it and send it back, and it was back and forth. (OP29)

The choice of the tool in a collaborative setting is a trade-off between the needs, functionalities, and competences of the team members. The importance of the right choice of technology becomes apparent when difficulties arise or when glitches occur.

> When writing in a collaborative way, sooner or later, we get to the problem of the tool. If the document gets too voluminous or too big later on – you have a problem. For instance, when using Google Drive, we all know how many problems we get later on with formatting the paper. All those tools at the moment have some kind of a bottleneck. Either in the purely technological sense or in the sense of authors who are still not used to using different tools or different approaches. (OP32)

The selection of tools and the competences in using them is connected to disciplinary differences, and to what tool is commonly used in particular communities; so it could be Google Docs for humanities scholars and the more specialised Overleaf for collaboration between computer scientists (OP13). This digital humanities researcher neatly described the differences between the most popular tools discussed by our interviewees:

> [...] some people still like to work mostly in something like Word, with the back and forth with this Word document, which is the most annoying way of working; but if other people like it, I usually sort of, you know, comply with that. And usually when people are at different locations we tend to use Google Docs, actually, as the major tool to make the bulk of the publication. With several other people, I also use Overleaf, so that's when a paper or an article has a LaTex kind of template. (OP08)

Google Docs use, as mentioned above, is widespread due to its ease of use and its ability to track progress and individual contributions (OP01), although it seems to fare better with shorter texts than hundred-page long documents (OP05, Pos.1). Open-source alternatives (like Next cloud) were also reported (OP10).

Sometimes, the **lack of competences of some scholars pushes other team members to abandon more specialised tools**, as in the case of this humanities scholar who

> could not think of any other tool except Word. He was the main author, in a way, and the most responsible for the monograph. The rest of us were keen on testing different solutions, or working in a collaborative environment, but in the end, we stuck to the first idea of that humanist scholar who forced us, in a way, to write in Word and to share the chapters and the writings via email, which was pretty time consuming, but in the end still interesting. (OP32)

A lack of competences among some team members can be compensated for by overall team effort, for instance, with one person doing Zotero references for the entire group (OP02).

Emailing a draft to collaborators may also be considered part of the ideation phase, as in the case of the back and forth with a literary studies professor from another town: “So I send it to him for some kind of study. Then, he sends his ideas [back]. We add; I change what he wrote, and so on. Then I send it again, he sends it back again. It doesn't take too long” (OP14). **To support ideation and have discussions with collaborators, our interviewees used various communication channels,** such as email, videoconferencing, and instant messaging.

> In most cases we just use email; we exchange the manuscript with the track changes, and then we use comments to sort of communicate about the changes that we made. [...] this is usually related to a paper that I'm just planning (in the planning phase), so I have an idea and I want to check how future collaborators see it, and then we may use Skype or Zoom to sort of talk a little bit about it and to discuss the structure, the idea etc. (OP31)

3.3.5 **Summary**

- Writing is a deeply social and technologically supported activity. We consider writing to be a different activity to publishing and communicating research. Discovery, storing, reading, and annotating research assets are treated as part of the writing process, as they influence the outcome and have already generated content that will be used in the drafting phase.
- We distinguish between digitally-enabled writing and digital writing. The former refers to writing as a textual practice supported by various digital tools, while digital writing harnesses the full potential of digital technology by establishing different kinds of materials like data, visualisation, or pictures, in a single output.
- Scholars appeared to use many different tools in their unique workflows, which are deeply rooted in individual preferences or experience, type of project, and disciplinary needs.

- When it comes to tool selection, we distinguish between engineers and bricoleurs. The engineers are experts in many specialist tools, and they switch fluidly between them. Bricoleurs, on the other hand, still combine digital practices with those that are analogue and offline.
- Interviewees discovered their sources through online search, catalogues, “citation hopping,” and “invisible colleges.”
- Storing and annotating are closely connected, and are oftentimes done with the PDF versions of texts. Various tools and services are used for this purpose: bibliography managers, cloud storage, and note-taking applications.
- Many interviewees distinguished between the ideation and drafting phase of the writing process. Ideation is an activity usually conducted in plain text to keep things simple so that it does not allow technology to interfere with the thought process.
- Scholars struggle to improve their own writing, especially when they are non-native speakers writing in English.
- Interviewees reported different experiences with collaboration, ranging from negotiating the process with multiple authors, which may ultimately lead to a patchwork structure on the one hand, or to dry, “impersonal” language on the other. Good writing collaboration requires trust and boldness, and often some level of personal, informal connection. Collaborative writing requires a lead author, who organises the process and coordinates the team.
- The choice of tool in a collaborative setting is a trade-off between the needs, functionalities, and competences of the team members. It would not be an exaggeration to say that all of our interviewees who write collaboratively have used Google docs for this process.
- Interviewees used various communication channels to support ideation and have discussions with collaborators: email, videoconferencing, and instant messaging.

3.4 CHOOSING THE PUBLICATION TYPE

The interviews provided insight into the publication practices of SSH scholars, but also allowed us to recognise the main motives that guide them in the process of choosing a publication type, venue, or publisher for their works. We were able to identify a number of elements that guide such choices.

3.4.1 Appropriateness of the form to the content

When considering a publication type for their future research results, and before examining any other factors that could influence their decision, most respondents were, understandably, looking for a form that would accommodate their content in the

most appropriate way. Certain characteristics of format or genre can make them especially appropriate for specific content or purposes.

Length. Many times, the simple characteristic of a text's length was decisive in terms of format choice. For a “voluminous piece of work, [of] around 280 pages” it would be a “natural decision to publish it as a book” (OP32), while in other cases a journal article would be “the best format in terms of its brevity and structure” (OP31).

Oftentimes, the decision about publication type is not made until the research is over. In one case, only when the research was finished did it become obvious that there was too much material for a journal article, and that a short book was a more appropriate form:

> And it very soon became clear that we could have written, you know, three or four journal articles. But actually, it was nicer to package it together within the single narrative space of a book. (OP03)

Developing a narrative or an argumentation. For many authors, especially in the humanities, developing a proper argumentation and theoretical approach is only possible in books.

> I like to write scientific articles very much, but I suffer here because I am not able to make any real arguments. […] And I need, sometimes – despite the fact that I try, and, according to others, I am clear in expressing my thoughts in writing – but I need a diversion. So I write a book. (OP13)

We were especially interested in the reasons why authors might find novel formats and genres more appropriate for their content, and there were several reasons mentioned. These, and other features of novel forms, are discussed more extensively in the upcoming section on innovation.

New formats are liberating.

> [...] they allow you to express yourself in ways that standard formats cannot allow you to express yourself. (OP29)

> [to a virtual exhibition you] can add videos, images, you can even add archives, not only a quote; it's less fixed. There is no final date for publication, you can update the content, add documents regularly – it's more dynamic. (OP27-34)

New formats are communicative.

> [The blog is] a great tool for gaining contacts, for building cooperation, for getting someone interested in the subject. (OP13)

New formats are interactive and collaborative.

In one case described, the author argued for the use of computers and software in the study of biblical texts. A digital publication (e-book with accompanying data for re-use by readers) would be the most appropriate.

> Because of the specific content of that argument, you can't just talk about that; you have to actually do some of it […] because that affords the reader […] more functionality in terms of understanding the data. (OP24-22)

> [The living book] was an exploration of how to use multimedia in a publication. The links could be articles, of course, but also videos, sound files, anything that's already online. (OP22)

New formats enable versioning and updating.

> The peer-review will come at the end of the process, after the [webbook] is published online. And the [webbook] will be modified following the peer-review comments, and thanks to versioning, the modifications will be visible. (OP21)

Experimenting with new forms aligns with the research agenda.

> [Concerning a podcast] we were analysing media histories, and it was really important to us, from the start of the project, to use different media to explore what we were talking about. (OP28)

> And because we're working in publishing studies anyway […] people kind of expect you to do something a bit different, I guess, with your publishing. (OP10)

3.4.2 **Community (and thematic) relevance and status**

There are publishing outlets (journals, book series etc.) that are perceived as being the most relevant form for certain topics within their respective scholarly communities.

> [Book series] have the corresponding reputation, because they have the corresponding editors, and of course you can position yourself clearly with that. And when I was thinking about how I wanted to, or should, publish, this question was, of course, at the top of my list. (OP20)

Of course, in such a setting, the interdisciplinary nature of a text will present itself as an additional challenge in finding an appropriate publication venue.

> For my work, choosing the publication venue was one of the most difficult challenges because my work is transdisciplinary […], actually just choosing which home discipline I want to frame the work under is a challenge to start with. […] So finding interdisciplinary journals – and ones that are open minded for this kind of research – is usually my hardest task. (OP06-50)

Such tacit understanding of reputation and the relevance of publishing venues does not always coincide with the bibliometric indicators of impact. For instance, an edited volume with no metric value can have huge significance in a field.

> But that volume is of such importance and has been referred to and is used so much that it turns out to have a much more lasting impact than an average journal article might. (OP04)

Some important publications with reputation in their discipline are not indexed in the relevant databases (and therefore, are not counted in formal assessment procedures).

> [...] what I'm looking for is that these are places, [for] journals or books, that I want to publish [in] because they either have a long history, they're interesting to me, their editorial board is good and so on. But as a researcher in a particular institution, and educational and research system, I had to take care about that. […] sometimes I sacrifice myself and also publish in journals that are very respected in my field but they haven't been included in these databases. (OP18)

Even the formats themselves (irrespective of publishers) can have an associated status and reputation. For instance, there is a difference in the perceived status of edited volumes, conference proceedings, and journals.

> [...] there is a difference between conference proceedings and just an edited volume. So there are a lot of shenanigans surrounding whether you publish this thing as conference proceedings. (OP01)

3.4.3 **Expected future of discoverability and visibility**

Discoverability and visibility will determine the size of the potential audience and readership; therefore, authors try to choose publication venues that they hope will have wider reach and visibility. Based on these criteria, our respondents indicated that they preferred publishing in the following venues.

Publications with good quality metadata.

> I think that they are major drivers of accessibility. And if it's search engine optimisation, I think that that's something that more journals do nowadays. (OP01)

High impact journals indexed in international citation indexes.

> Why did I publish in […] journals in the Journal Citations Report? Because these are the best journals. There are so many journals and publications nowadays that I have the impression that to be read by someone more than reviewers, you have to choose the best journals. (OP11)

Reputable monograph series that will attract many book reviews, which serves as an important informal evaluation practice in the humanities.

> The leading series of a discipline are at least very reliably bought by the big libraries, so that's how you get a wide reach. I think you can also tell from the reviews that those are received very widely. (OP20)

Up-to-date formats. Some formats may become obsolete and make access to the work difficult, as in the case of the following PhD dissertation that was published on microfilm, which was later perceived as being a bad choice:

> I think I've been really misled because I published on microfilm, which is not a publication [format] for me. I mean, because it's not available. So for me, it's basically like printing your thing and putting it in your own drawer […] nobody reads it. (OP02)

International journals and languages.

> [T]hese are English language journals – so I hope to reach a much larger audience with my research than if I published it in Polish. (OP11)

3.4.4 **The economy of publishing**

A set of incentives that governs publication choices is related to the financial aspects, which can have an impact on the different actors in the publication production chain – publishers ("sometimes presses don't believe they're going to earn enough revenue to make it worth publishing an edited volume." OP04), and authors ("It's definitely a problem if you want to publish open access and you don't have a grant, and you don't have funding in the grant that is specified for [the] article processing charge, then of course your choices are limited." OP16)

3.4.5 **Bibliometric indicators**

Bibliometric indicators such as citation counts or simply being indexed in key citation indexes proved to be a matter of consideration for many of our interviewees. We asked interviewees about behaviours related to bibliometrics, but also about their attitudes toward a system in which metric criteria are often imposed by formal assessment and evaluation procedures. It has been reported in many countries that changes have been made to evaluation criteria that have even made humanities conform to the natural sciences' evaluation model.

> [...] increasingly, for the past couple of years, those metrics which have been used for the natural sciences have been adopted by, and adapted to, the humanities and social sciences. So now it's very important to publish in a journal with impact factor in the first or second quartile of, you know, Scopus – a journal that is indexed in Scopus or Web of Science. (OP19)

Many respondents reported that they had adapted to the system and were playing by the rules of the "economy of scholarly prestige":

> I care about getting grants. And to get grants, you need extra points for quotations. (OP11)

> I will submit first to the one with the highest impact factor, which I did. If my article is not accepted there, then I will submit it to the next one. (OP19)

Some interviewees had a reasonably positive view of bibliometrics, and suggested that impact metrics and the quality peer review are associated.

> Well with all the limitations of bibliometrics I still find it probably the best orientational tool in a sense: okay, is this journal really good or it's not that good. And since I have strong reasons to believe that many of my colleagues, many other scholars, apply the same logic, it means, you know, the journals with high

> impact factor would probably receive better manuscripts and would have better reviewers. (OP31)

It must be noted that bibliometric impact indicators do not always coincide with a perception of prestige and scientific relevance (covered in more detail in the prestige section of this report).

> So bibliometrics is important, but I would say, at a certain stage, what counts more is the position measured by non-parametric prestige, the resonance in the environment – where it is really worth being, where it is worth publishing. (OP13)

And there are certain types of publications that have relatively bad scores in bibliometric assessments, but are still perceived to be relevant to the community.

> I think there are some disadvantages in terms of bibliometrics, in terms of citations and things like that. I think that the rewards for publishing book chapters are not terribly high. (OP04)

3.4.6 **Reputation of publishers and editors**

The reputation of publishers and editors can play an important role in deciding where to publish. And this reputation can stem from track records, tradition, and editorial policies and practices. There is a tacit understanding about what is reputable, and the ways in which reputation is related to visibility and quality.

> I care about which publishing house [I use]. Because I know that this book won't really have many readers. But it will have readers in different countries if a famous publisher publishes it. (OP13)

From several of the responses, it seems that editors and reviewers are the key to the reputation of journals or series.

> I like trying to publish papers in reasonably good journals, which means […] you get, more often than not, really high-quality reviews. At least one in three reviews will have something really important and interesting to say and this is what makes the paper better. (OP31)

> [...] the series are quite important. And they have the corresponding reputation, because they have the corresponding editors […] And when I was thinking about how I wanted to, or should, publish, this question was, of course, at the top of my list. (OP20)

Looking from the publishers' perspective, it can take a lot of time and effort to build reputation, which can be a problem for small, new OA publishers:

> So, this is a point that also has to do with open access or the establishment of new formats. This standard has to be achieved first. (OP20)

3.4.7 **Open access**

Many of the interviewed scholars mentioned open access as a decisive criterion in choosing a publication venue.

> So when I'm the one in charge, open access is a very big thing. (OP02)

The reasons for favouring OA venues can be threefold:

National or institutional mandate.

> [...] it's actually a non-enforced directive of [my institution] that we should try to publish as much open access as possible, […] So, therefore, if I write, I try to aim for open access. (OP08)
> [...] for many authors, because there are national incentives and reward systems, for example, in Finland, Norway, and Belgium; this open access is an absolute priority. (OP13)

Personal principles or ideology.

> I was reading about open science and the issues of reproducibility and replicability. That's when I thought, this is an idea that I support and if I want to produce research that somebody else can verify or build upon, all the data has to be out there – so in a way it was a very conscious decision based on my belief. (OP16)

Reaching wider audience for some topics, even if it requires paying APCs.

> We want this to be as open as possible, so this is why we thought that it's important. It's not an open access journal, this particular one. So we had to – we decided that we wanted to raise the money; and we got the money from the project. (OP15)

However, authors sometimes have to choose between open access opportunities and such factors as the thematic relevance of the venue, or the quality and reputation of the publishers, publications, or editors; because some quality editions are only available through subscription or by purchase.

> And now we submitted our special issue to another journal and we got accepted. And it's open access. But I think, thematically, the journal is a bit less of a good match. (OP02)

3.4.8 **Invitations by editors**

A common situation in Humanities and Social Sciences is publishing upon invitation from editors, whether it is an article in a special thematic issue or a book chapter in a collective volume.

> This book chapter is coming out in a book that is being co-edited by a couple of scholars […] They approached me and told me about the book they were

> editing and asked if I wanted to submit something for it […] So it wasn't so much a matter [of] like having an idea for a book chapter in mind and then looking for a publisher – they came to me with an idea. (OP04)

Understandably, this happens more often for senior scholars than for PhD students or early career researchers (ECRs).

3.4.9 **The different situations of early career and senior scholars**

The arguments that influence the decisions about publication path are not the same throughout scholars' careers. Scholars just starting out are faced with greater pressure to make the right choice – and their options are limited.

> [...] early-career scholars are more exposed, more exposed and vulnerable, but are also, in many cases, given bad advice by some more senior scholars who tell them to be conservative and not take chances, and only do the most prestigious things. (OP04)

3.4.10 **Peer networks and acquaintances**

In many cases, the choice of publication type or venue is not just the outcome of deliberate consideration or reasonable argumentation, it can also be the result of taking part in a project, collaboration, or through a peer network or personal acquaintances:

> I was asked by a research colleague of mine, who I like and respect scientifically, to participate in a conference on a panel that she co-chaired, and to come up with a text for that panel. I didn't have any special idea, [though] there was something in my head. And she led me very coolly to a book that somehow opened my eyes to the whole field. So I wrote an ordered text. […] It was very nice. Because it was mostly about that "you're smart, you do interesting things, you achieved something." And when someone you like tells you this, you feel somehow motivated, committed. (OP12)

3.4.11 **Speed of publication**

The speed of publication can play a role in choosing among the traditional publication types, but it is also considered a major advantage of communicating through novel publishing genres, like blogs.

> I do blog a lot. […] I like to share my notes as quickly as possible. It feels like my area – because I work in open access publishing, which is such a fast moving environment that I'd rather just, kind of, let everyone know what my thoughts are about it via my blog, and then the formal publications will catch up with it. (OP10)

3.4.12 **Past experiences**

A positive, prior experience with certain publishers or editors can reinforce any future decision to publish in the same venue: "And when you chose, I mean in this case, it was a journal with which you collaborate regularly" (OP19).

3.4.13 Language (in relation to the intended audience)

For scholars from non-English speaking countries, one of the first decisions has to be made about the language of the publication, which will then influence the choice of publication venue.

> I try to publish in English. […] Because even if it concerns [national] issues, I want it to have some sort of overtone and to reach beyond [national] borders. (OP11)

The choice of publication language is often dependent on the evaluation criteria, as well as on the targeted audience.

> Then some diversification began, with the articles or book chapters in foreign languages having a higher score than the ones in Bulgarian, which is perhaps understandable because those in foreign languages normally have a larger audience than those in Bulgarian. (OP19)

Other options regarding the language of publication, apart from publishing in one's mother tongue or in one of the widely spoken languages, are multilingual editions or translations.

3.4.14 Issues of copyright and intellectual property rights

Issues concerning intellectual property rights (IPR) can impact the choice between print and digital formats, because online publishing can bring about more difficulties with copyright clearance, and this can be more complicated and expensive than is the case for print; and this includes translating or any use of third party copyrighted content, such as images.

> But this was a bit exhausting […] when it came to acquiring the image rights, and [to] the costs. And there were actually institutions that charged twice. […] but, of course, it produces a lot of communication when you ask and beg for it to be waived and so on, but that is, of course, one of the difficulties you have. (OP20)

3.4.15 Summary

- Authors choose the form that is appropriate to the content, based on its length, or on the possibility for developing a narrative or an argumentation.
- Novel formats and genres are considered more appropriate for certain content, for several reasons: they are liberating, communicative, interactive and collaborative, and they enable versioning and updating. Sometimes experimenting with new forms aligns with the research agenda.
- There is a tacit understanding of the reputation and relevance of some publishing venues for certain topics within the respective scholarly

communities, which does not always coincide with bibliometric indicators of impact.

- Authors try to choose publication venues based on their expected future discoverability and visibility. They often prefer publishing in, publications with good quality metadata, high impact international journals indexed in international citation indexes, in the English language, and in reputable monograph series that would attract many book reviews.
- Bibliometric indicators proved to be a matter of consideration for many of our interviewees.
- The reputation of publishers and editors can play an important role in decisions about where to publish. This can be the result of historical track record, tradition, or editorial policies and practice.
- The reasons for favouring open access venues can be threefold: national or institutional mandates, personal principles or ideologies, or reaching a wider audience.
- A common situation in SSH is publishing upon editors' invitation.
- Early career researchers are faced with greater pressure to make the right choice, but more limited options concerning publishing venues.
- In many cases, the choice of a publication type or venue is not just an outcome of deliberate consideration, it can also be the result of taking part in a project, collaboration, or through a peer network or personal acquaintances.
- The speed of publication can have a role in choosing among the traditional publication types, but it is also considered a major advantage in communicating through novel publishing genres (such as blogs).
- For scholars from non-English speaking countries, one of the first decisions to be made must be about the language of the publication, which will then influence the choice of publication venue. The choice of publication language is itself dependent on the evaluation criteria, as well as on the target audience.

3.5 TRADITIONAL PUBLICATION

Even though the focus of our research was on innovative forms and genres, the interviews also revealed a lot of information about the landscape of traditional publishing in SSH. They demonstrated the diversity of the field. Although it seems that the humanities (and to a lesser extent, also the social sciences) are inclined to experiment with forms and genres, the format that received the most attention by our respondents was the book, along with a range of its variants.

3.5.1 **Books**

Throughout our interviews, respondents used concepts familiar to them to denote certain formats and genres, however, the main goal of this research was not a clarification of terminology or finding exact definitions. Many of the interviewees used the terms “books” and “monographs” interchangeably. For a better understanding, in this paragraph we will use the concept of the “book” in a broader sense, which encompasses long-form scholarly writing, such as monographs, edited volumes, textbooks, critical editions, trade publications, and, possibly, other formats, irrespective of their media (print or online).

Books are defined by their form but also by their status in the scholarly community. They do not easily conform to the systems of metrics for defining quality. For books, there are separate ways to define and express impact and reputation. For instance, book reviews are an important type of evaluation in the humanities, and book presentation events are important for initiating discussions that can be gratifying for both authors and publishers.

3.5.1.1 What is a monograph?

According to many of our respondents, the central position in the SSH publishing ecosystem is reserved for the scholarly monograph. For many of them, the monograph was the first thing that came to mind when they were asked what a scholarly text was for them:

> the humanities are distinguished by the fact that the most interesting things are in books. (OP14)

From the answers gathered, it is possible to identify what is commonly understood to be a monograph. One decisive characteristic is the volume.

> I still think we, as scholars, are sort of ingrained with the belief that the monograph is 80 thousand – essentially an 80 to 100 thousand word document that’s been peer-reviewed and [that] it’s often written by one author. (OP10)

More substantially, the idea of a monograph was regularly associated with the concepts of “linear argument” (OP10), “single narrative space,” “a nice story” (OP03), or “a holistic […] picture of a research problem” (OP11). It was regarded as an appropriate format for “a theoretical approach” and conceptualisation (OP13). The special status of the monograph with respect to its prestige or perceived impact is elaborated in another section of this report.

3.5.1.2 The different roles of edited volumes

When we arrived at the concept of a multi-authored and edited work, it was even more difficult to reach a common understanding of the nature and status of this format. We can even understand it as a sort of continuum, where there are no clear borders between special thematic journal issues, conference proceedings, or edited volumes.

For instance, it is very common in SSH for a conference presentation to become a journal article.

> I first made the presentation for a conference. And then it was somehow only natural that the conference paper would develop into an article. And then there was this possibility to publish it in a thematic issue of an online journal. (OP19)

The nuances of the perceived status and prestige of edited volumes (as opposed to conference proceedings or journal articles) are puzzling to any newcomer to the scholarly communication landscape; this is further complicated because these nuances vary across fields and specialisations. The origins of this differentiation relate to the role of editors, and to the strictness of selection (or invitation) and peer review procedures, and is often reinforced by evaluation and assessment criteria.

> [...] for a journal article, we get higher scores than a collective volume. (OP19)

> [...the] conference publication […] is damned devalued, so no one wants to publish in a conference volume because it's worth only a few points in our evaluation system. (OP05, P. 4: 1729)

Nevertheless, "edited volumes frequently have a really important impact on their fields" (OP04).

3.5.1.3 Importance of the book series

In many SSH disciplines, prestige and reputation are associated with monograph series rather than with publishers: "the series is very central, especially in history" (OP20).

They offer a certain guarantee of visibility and availability since they often guide selection procedures in library purchases: "the leading series of a discipline are at least very reliably bought by the big libraries, so that's how you get a wide reach" (OP20). Editors are key to the series' reputation: "the series are quite important. And they have the corresponding reputation, because they have the corresponding editors" (OP20).

3.5.1.4 Are e-books not books as well?

It might seem strange to even pose the question as to whether the digital medium is an important vehicle for scholarly content nowadays, as it is certainly widespread in journals. All our respondents implied that the online editions of journals were their primary editions. But, with books, the question of digital as opposed to print editions arose more than once during the interviews. For some respondents, having a digital edition was highly desirable, while for others it was perceived only as something secondary and that it was not so important for the book to be immediately available online. Interestingly, one interviewee even forgot to mention that their book had been published as an ebook along with the print edition. For this interviewee, the print edition was important, and the ebook was only a side product. Another interviewee mentioned the problems of smaller presses and e-books – despite the author's desire to publish an ebook, the publisher was not able to manage it.

Still, the questions raised around digital media for books are more related to practical issues like cost, availability, distribution, and discoverability; and rather did not concern changes in content. "So far, the form is not influencing the content so much, […] it's still the same text. […] an e-book is not automatically different from an

ordinary book" (OP23). In the next section we will discuss the innovative use of digital carriers as a predominant means of reaching new audiences and making the work accessible. In cases where they described innovations in content, interviewees referred to them under different labels (for instance "web books" or "living books," which are addressed in greater detail in the section on innovation).

3.5.2 **Journals**

Despite the unique position of the monograph, journals remain the main vehicle for communicating research in many SSH disciplines ("the most important things happen in journals, i.e. in scientific articles," OP14). Moreover, several respondents identified journal articles as being the most prominent type of scholarly text.

> A scholarly text, for me, is the first association I have with a journal article and all the relevant stuff around it, a traditional structure with literature review, method, [and] some bibliography at the end of the article. (OP16)

The main characteristics of a journal article that make it so practical for communication in academia are its brevity, structure, speed of publication, and the process of peer review.

> [...] they tend to be characterised by the shortest amount of time needed for an idea to reach other people, and then you can have a discussion, then other things can be built upon this first paper. (OP31)

It should be noted that the journal article was also understood as being a format undergoing a slow, gradual evolution.

> These models evolve over several decades […], but it is an unconscious, collective, and a very slow evolution that is not subject to specific deliberation. (OP25)

In the future, we can expect more and more journal articles to include code and data for better functionality: "so I could basically run the code and get the results that are published in the article." Such an article would consist of "the text itself, with some supplements: this would be the data set, this could include licenses, and this would include the code to process the data with" (OP16).

3.5.3 **Theses**

Besides books and journals, PhD theses were another type of publication mentioned several times. These can be seen as an important source of information in their own right.

> [...] it's interesting to go back and read their full thesis, because their thesis was much more detailed and they had time and space to fully show all the results and everything. (OP06)

But in many disciplines, "publishing a book from your thesis, a book from your habilitation, these are important elements in a career" (OP25), and this is considered to be a standard procedure. Sometimes, the theses can be published in some other

format, like microfilm (which was described as being the wrong choice, due to the unavailability of that format). Finding the appropriate format for publishing a thesis, and experimenting with the dissertation form can be a challenge for an early-career scholar.

3.5.4 **Summary**

- Books are defined by their form but also by their status in the scholarly community. They do not easily conform to metrical systems for defining quality. For books, there are separate ways of defining and expressing impact and reputation.
- For many of our respondents, the central position in the SSH publishing ecosystem is reserved for the scholarly monograph.
- According to common understanding, one decisive characteristic of the monograph is its volume, while another is its capability for presenting a “linear argument.”
- In many SSH disciplines, prestige and reputation were associated more with monograph series, and not so much with publishers.
- For a multi-authored and edited work, it is difficult to reach a common understanding of the nature and status of the format. There are no clear borders between special thematic journal issues, conference proceedings, or edited volumes, but their status and reputation are often perceived differently.
- Questions raised around digital media for books are more related to practical issues and less to changes in content. In cases where innovations in content were described, interviewees referred to them under different labels (“web books” or “living books”).
- In many SSH disciplines, journals are the main vehicle for communicating current research in a formal way.
- The main characteristics of the journal article, which make it so practical for communication in academia, are its brevity, structure, speed of publication, and the process of peer review.
- In the future we can expect more and more journal articles to include supplemental content (data sets, code, etc.).
- Theses and dissertations are an important source of information in themselves, and in many disciplines, publishing a book from a thesis is considered to be standard procedure.

3.6 INNOVATION IN SCHOLARLY COMMUNICATION

3.6.1 **What is innovation in scholarly communication?**

When asked to define innovation in scholarly communication, our interviewees pictured it as the activity of experimenting in order to find a better way of doing something.

> [T]he point of an innovative publication is that it's not been done before, so there are no guidelines! There's a trial and error aspect. (OP21)
>
> Innovation means trying something new, and sometimes it might not work. (OP28)

Innovation may be also seen as an **additional, popular version of science**, not a new means of communicating with researchers (e.g. OP27). "I think innovation is something that sort of unsettles the way that we have always done things" (OP10).

Not only does innovation unsettle the way things have been, but it also provides much needed room for improvement and novelty.

> Innovation can be disruptive. All the tools that I'm using and promoting can be very challenging to use for some people who are not used at all to that system, and who see them as a threat to the efficiency of their process. (OP17)

In general, innovation is seen as a chance to improve the sharing of ideas with audiences thanks to novel technology.

> I think, these days it's the changes that have to do with the Internet being the main platform for our communication; so it's much easier to share things now and so I think that innovation basically means catching up with opportunities that technology offers. (OP16)

Interestingly, this respondent saw innovation as a means to reconnect with the roots of scholarly communication, as current norms and traditions of scholarly communication tend to be incompatible with what is currently possible due to technology:

> I think now it's clear that we should change the norms and change the traditions to catch the original intent of scholarly communication, which I think is to publish your results and share the results of your work. (OP16)

In a similar spirit, another interviewee reported his turning towards innovation, namely, publishing a blog, because of a dissatisfaction with how his writing was displayed on publishing platforms (OP17). Hence, he chose innovation, because publishers did not support features he considered better for communication.

Innovation was also understood to bring seamlessness to the process of using scholarly content, removing obstacles that are unnecessary from the vantage point of current technology. A philosophy professor valued easy access to articles online, hence, he rather used Sci-Hub than his library access so he didn't have to "think about which window, where to click, which database to connect to" (OP13). Innovation is also understood to provide alignment between scholarly publishing and modern communication practices, creating an environment to capture readers'

attention (OP14). Perhaps this interviewee captures the general attitude most accurately:

> I think that innovation comes in a number of ways. One is innovation and access, so moving beyond the model of the paywall or moving beyond the model of subscriptions to get scholarship out there […]. Two, there is innovation in terms of modes of scholarly output, incorporating images incorporating websites, etc., into scholarly output. (OP24)

Innovation, then, is understood either in terms of form (novel means of communicating ideas), or access, i.e., accessing content or reaching new audiences.

One could risk the hypothesis that **what scholars consider innovative depends on the horizon of possibilities they see thanks to their experience, needs, and the types of sources they deal with in their research**. Hence, researchers who are more engaged with digital methods tend to consider innovation in formal terms, while the others focus on access and reaching audiences.

3.6.1.1 Access

First of all, innovation is considered in terms of providing access to more traditional types of outputs. According to interviewees this seems to be the most tangible form of innovation, perhaps because it responds to a more basic scholarly need for accessing content regardless of its form or features.

> I see that most innovation has been done in the area of the distribution of scholarly work and sharing scholarly work, either between people or between machines. That part is actually pretty innovative compared to previous phases or stages of scholarly communication. (OP32)

Interestingly, this form of innovation is **usually described in "negative" terms, i.e., as removing some of the obstacles rather than providing new value**: "moving beyond the model of the paywall, or moving beyond the model of subscriptions to get scholarship out there. I think that's innovative" (OP24). The innovation lies in the platforms providing seamless, non-paywalled access to scholarly content, be it ResearchGate (OP11, Pos.66), Sci-hub (OP13, Pos.69), or a repository: "You can just browse the journal and look at those articles. And this is not the most visited one. So there are a couple of thousand, a few thousand visits, a few thousand readers. And this is the enormous advantage of open access online scholarly publications – that they can find readers" (OP19). The ability to use citation metrics and usage statistics is also considered innovative, as they allow for quality assessment as well as the measurement impact of the scholar's own work.

In this context it is worth mentioning some forms that seem traditional at the first sight, but have been given new life and meaning due to digital distribution, namely **grey literature.** This is a pretty broad term that describes the various outputs of academic work beyond the usual, traditional genres of academic publishing, i.e., reports, policy briefs, presentations, working papers, and drafts. Despite not having the seal of approval of traditional publishing outlets, these materials are published and read by the community, as in the case of our interviewee's research reports which weren't

published: “They are just circulated among some particular community or given to the ministry” (OP18).

One of these genres, the Powerpoint presentation, is viewed as a simple form of visual genre, allowing researchers to get their argument across visually, or to provide some interaction between text and images without advanced technology. These formats have become more popular thanks to the advancement of online repositories and a rising awareness about self-archiving among scholars. Many conferences nowadays encourage the depositing of presentations for further use, thus allowing access to more ephemeral products of research. Despite all this, one interviewee claimed that “Nobody publishes Powerpoints; there is no format of publication like a Powerpoint, or based on images” (OP22), which points to the fact that sometimes these formats are not recognised as publications, perhaps because they lack quality assurance.

Equally important in this context is the issue of **research data.**[5] Recently, there has been an increase in discussions about FAIR principles (which emphasise that “data should be Findable, Accessible, Interoperable and Reusable to the greatest extent possible” [Directorate-General for Research and Innovation (European Commission) 2018]), data storage, and sharing, and many interviewees are aware of the issues surrounding supplementary resources in SSH. We asked the respondents directly about their opinion about publishing the entire material from a given study, including complete interviews, annotated texts, research protocols, and the data collected during the research process. There were different opinions on the topic, including doubts related to the time-consuming aspect of academic work, which already requires researchers to read, write, and peer review articles.

> It doesn’t make any sense. I already don’t have time to read all the articles I want to read. I understand it intellectually, but given the time I have, I don’t think I would take the time to get into the underground area below the article. (OP27)

> In History, we are already happy if one person takes time to read what we write! Who is going to read research notes on a subject for which the final monograph or publication will already be read by too few scholars? (OP22)

> I’m very much in favour of there actually being digital data repositories that allow as much data as possible to be accessed by people who are interested. I think that [...] the accessibility part of the data should be increased online [...]. The problem is that the research data is only relevant to a very small portion

[5] This section uses several excerpts and findings from Elisa Nury and Claire Clivaz, with Marta Błaszczyńska, Michael Kaiser, Agata Morka, Valérie Schaefer, Jadranka Stojanovski and Erzsébet Tóth-Czifra, “Open Research Data and Innovative Scholarly Writing: OPERAS Highlights”, Proceedings of the Swiss Data Research Day 2020, Makhlouf Shabou Basma et al. (eds.), forthcoming. The article will be published under the licence CC BY SA.

> of the readers. That is to say that, in fact, it's like footnotes, footnotes are very important for the epistemological and ethical guarantee of the work. (OP25)

Research data are a very important part of guaranteeing the validity of research processes, but they are usually used only by a minority of scholars, as opposed to publications, which means that data sharing represents a large investment in time for what might seem like little return. However, the availability of research data is crucial for the accuracy and reliability of peer review. Interviewees also stressed the importance of transparency, with caveats about privacy, copyright, and the reuse of data.

> It is even necessary, or it is becoming more and more mandatory in certain cases. Today, transparency is very important. (OP21)

> Publishing the data and publishing what you did with that data, so publishing some form of code that you used to get from the data to conclusions, and to create visualisations, and tables, and stuff like that – I think that would be very beneficial […]. Also, it would make the whole process much more transparent and while it would not eliminate it, it would reduce the margin for foul play. (OP29)

> One of the things that we come up against is that, culturally, people expect transparency. This becomes dangerous because then you can violate things like privacy. [...] But if I put that stuff out there, scholar X is going to take that data and write that next book that I'm not going to [write] right now. Because the incentives of scholarship are what they are, you still have to be careful about what full publication would look like. (OP24)

Ideally, the research data would be published with the same standards of rigor as traditional academic publications, however the peer review of data would raise an enormous challenge in terms of the workload it would impose on reviewers, who may already be short on time. Even researchers who agree that research data should be peer-reviewed admit that realistically, we will never be able to do it at scale.

> And really, the labour involved in evaluating these things just goes through the roof. And I just don't think people are going to have time to do that kind of evaluation for every piece of digital scholarship that emerges in the next few years. So, I think there's a looming crisis for the labour of peer review. (OP03)

Thus, in SSH there is an interest in the need to publish research data, although there are limitations and obstacles, such as this scarcity of time often referred to when discussing all aspects of scholarly communication. Moreover, as a result of the reward structure, there is a strong incentive for researchers to publish traditional scholarship in prestigious venues for their field of study in order to advance their careers (see Power structures and Task 6.6 report).

In this context, and within the time constraints of research projects, for scholars this situation creates a tension between the need for traditional publications and desire for innovative practices, as highlighted during our interviews: "publishing data is time-

consuming, which is a disadvantage; the work often has to be done twice, once for preparing and depositing the digital output, and once for a more traditional publication" (OP21). These problems are not new and have been repeatedly highlighted in the field of digital humanities: "Digital humanists find, time and time again, that they are expected to perform twice the labour of traditional scholars; once for the work itself and once again for its evaluation" (Eve, 2020; see also Baillot, 2016 and Fitzpatrick, 2011).

A similar issue, related to code, was highlighted by one of the interviewees. He emphasised that in the case of a workshop publication (published online), although many members of the audience would mostly be interested in the presented code, which had its own DOI and was "on GitHub [...and] on Zenodo" (OP05), it needed to be accompanied by a more traditional narrative. Thus, even when the code is the key output, it needs to be accompanied by an academic text that was quickly produced just for the sake of it: "[W]ithout a scholarly article that gives context to it, we cannot publish the code, which was the important thing to do" (OP05).

There is also concern about how to link together the various outputs of a project: the data, the articles, the code, the source materials etc. The common practice now is to use persistent identifiers such as DOIs. This also has implications for publishers and libraries: publishers will have to deal with projects that have multiple outputs (OP24). How do these outputs hold together as a unified, complex entity? How do librarians catalogue and provide access to a publication made of multiple parts? Are PIDs sufficient? Can they be used to keep track of citations for the data? There seems to be no effective information management system in place for this. Moreover, while certain writing tools allow for a greater integration of the data into scholarly text, often only the minority of researchers use them. One of the interviewees felt that in their field

> [t]he relation between data and writing is still a bit conflictual because people write in Word and there's no way to integrate your statistics or your lines of code nicely, and to have good synchronisation between the data and the text you're writing, or to provide interaction between the text and the reader. (OP17)

3.6.1.2 Formal innovations

Apart from innovation in providing access, another kind is the innovation of form. It should be stressed at the very beginning that what is meant here is more advanced forms, not simply digital recreations of traditional genres: "an e-book is not automatically different from an ordinary book. Or the database handbook is not different from the handbook itself – content-wise" (OP23). Innovation is thus part of the general process of the slow evolution of communication forms. One historian stressed that all forms evolve over time:

> Each field has its article model, and these models evolve over several decades. That is to say that today, in the humanities and social sciences, it is not quite the same texts as 50 years ago, 100 years ago, but it is an unconscious, collective, and very slow evolution that is not the subject of specific deliberation. (OP25)

Importantly, this is viewed as an **evolutionary process that does not replace older forms**, but opens new "niches" (OP30). So, every change needs to address the values and traditions of the field in order to become accepted, "and [to] show that these can still be met in this new format" (OP24).

Innovation, in fact, does not need new technologies as it can be played out in traditional formats. One of the interviewees gave Punctum publications as an example of such works; these are written in conventional form and published as PDFs but retain innovative potential: "Often it's still a contained book format, but it can still be radical" (OP03). However, in most cases our interviewees referred to the technological aspects of innovation.

Innovation allows for new types of interaction with the text. But first and foremost it leads to a change in our perception of what constitutes a text. In this case it is not only about the features of scholarly writing (i.e. what makes writing scholarly), but rather a more general understanding of what the text could be. Creating hypertextual connections to other texts and materials seem to be a basic innovative feature that is recognised by researchers, turning the text into a gateway to other materials: "your text could actually be a kind of reading guide across the digital space on the issue you were addressing" (OP25). However, the digital medium has a greater potential for the radical disruption of this understanding, which was a conclusion shared by many other interviewees, namely, that **contemporary scholarly text goes beyond simple verbal expression, incorporating different, new types of content**. The following quote is longish, but very important in terms of conveying the gradual sense of innovation in writing, and it starts with linking the text and data and then goes on to suggest an even more radical mode in which the text becomes executable, allowing for dynamic interaction with its content.

> But let's say that most people regard what you see on the screen or what you read on page as the text, you know, it's those characters in those sentences. That's the text. But now imagine if we can convince people that something like code, programming code, is also a text. On a philosophical level, people never have any problems with acknowledging that. Yes, that looks like a text and it's sort of the same thing as text. So, yes, it's a text. [if this is agreed, can we then read a text that is code?]. Then you get two types of people: you get the ones that say, no, that's not a scholarly article; [it] is simply not because it doesn't have the form and format that we as scholars expect as the hallmark of how we do things – how this whole scholarly process works and how we report about it. So that's not a scholarly text. And then on the other hand, there are the people, obviously like me, that say no, that's an interesting innovation of how text could also be a mechanism of reporting your research. So why not accept a text that can actually execute itself as a scholarly text? And if you go in that direction, we haven't even produced anything innovative because all the things that we produce until now are still basically those things that we just read, they are on a screen. And sometimes they are supported by some data repository or a code repository, but we don't have anything that executes, you

> know, that creates itself by you executing or running the text, as it were. (OP08)

Based on this perspective we may distinguish three kinds of formal innovation emerging from our interviews. First, there is a basic **move beyond the mere written word**, i.e., accepting expression in other media forms as valid scholarly outputs.

> [E]verything that moves away from text, other media – it is innovative to consider them as a possible way to transmit a scholarly reflection. For instance video, sound, podcast […] Everything that is moving away from traditional writing processes. (OP21)

This may also entail the very loose understanding of scientific text as a transmedia practice for delivering content through a range of various utterances in different media. Greta Thunberg was an example used in this context:

> She is innovative because she is multi-channel. And she communicates through many channels. But she wants to acknowledge the scientific truth, right? I mean, her main message is: you don't listen to me and you don't listen to scientists. (OP14).

Second, the **text can be linked to data that allows access to the source material of a given study, be it data or code.**

> [In our team] we are discussing exactly that: how do we publish something that is telling a story so there is a narrative, but then also include how the researchers got to the story, the analytic part, and then what dataset they used to do that? So that's a three-way approach to the whole thing and that's not easy. (OP23)

This means acknowledging that **scholarly writing should allow access to the underlying content for validation, replication, or further interaction**.

> I think the vast majority of scholarly texts in my field are still text, right; they're still sort of paper shaped, they come out in PDF or they still pretend to be print on paper even when they're not. But I think there are more and more options and more ways in which publications and the kinds of scholarly texts that I rely on are starting to break those boundaries. And to think about, you know, ways that deliver a text through HTML on the Web allows you to create links instead of citations and thinking about the embedding of images or of charts or of other kinds of media forms within the frame of that text. (OP04)

Finally, the third dimension could be treated as an enhancement of the previous one. **If we connect text and data we should also think about providing a novel level of interaction, which is impossible in static texts**. As this archaeology scholar mentioned:

> I'd like to see more powerful and intelligent ways of connecting research findings and research claims with evidence, […] [allowing] people to construct research artefacts, on-line publications that are more dynamic. (OP15)

This may include novel outputs ranging from dynamic visualisations to the generative text envisaged earlier in this section.

Such innovations also change the way we think about scholarly argument and its authorship. One interviewee discussed how experimentation undermines

> the history of the last 50 years of [a] kind of liberal, humanist, Western thought, which is you write a text that's yours. […] And obviously that's bad for a number of reasons, that it ingrains certain ways of thinking, certain kinds of linear, rational ways of thinking [about] that kind of work against the other ways. (OP10)

If the knowledge becomes generative and interconnected, how should we measure contribution: "if you post something on-line and I write something that relates to what you said but creates a new idea, how can I claim this as part of my scholarly output? I cannot connect it with other things that I produced" (OP15).

Our interviewees often spoke about more advanced forms of actually linking data and text within a publication, going beyond the mere depositing discussed in the earlier section on access. This idea opens up the concept of the book by allowing text to be linked to other outputs like data, code, or supplemental materials. A website seems to be by far the main format mentioned by our interviewees in the context of innovation, probably due to its flexibility of handling different genres. Here is how a post-doc in biblical studies describes her involvement in a web book creation:

> It was thought of as a book publication, but only for [the] Internet. It is not like an e-book, which can be both paper and e-book, and its purpose is not to imitate a printed book, but only to have HTML pages. […] The idea is to keep it light and easily manageable. It can be a sub-type of a website. But the idea is still to make a book, to keep [it as] a long text. It was our conviction that we should still be able to carry out long-term research and reflection. It's an added value in humanities research compared to other scholarly texts. (OP21)

Thus a web book seems to remediate the book by preserving the long scholarly argument on the one hand, and opening it to use by different media on the other. Some interviewees discussed a similar concept, which we distil here as a **computational essay**, an article, or a book that focuses on linking the text with underlying data:

> So you've written some research in a programming notebook and not only have you done that, but you provide it in a format that also leverages that functionality So, for example, people can see that there's a parameter in an experiment that's been used to produce a graph and they have a little checkbox that they can use to make the parameter vary and see the graph update. That sort of thing for me is innovative, not in terms of technology, because it's quite old, actually […]. It's just that publishing systems don't use it. (OP17)

It is important to note that it is not only about providing the data but rather linking them with the outputs in a dynamic, interactive way, allowing readers to engage with

the scholarship at a deeper level. This archaeology professor compares such a publication to discovery research "in which the relationship between the claims that are made and the warrants for these claims – typically data used as evidence – would be clearer. […] Instead of just being given a diagram, I might be given a pivot table that I can sort of play with and see how they came to that conclusion" (OP15). These functionalities provide readers with new means of interacting with the content, offering new ways of understanding the data: "It is one thing for me to write a paragraph that talks about the conclusions of the data, but it's quite another thing for the reader to actually get into the data. So I think the digital format allows that greater flexibility" (OP24).

The computational essay also leverages the web format to establish links with external materials and sources: "delivering a text through HTML on the Web allows you to create links instead of citations, and think about the embedding of images, or of charts, or of other kinds of media forms within the frame of that text. So I think that the text, the notion of the scholarly text is starting to open up a bit" (OP04).

A scholarly edition may be considered closely connected to the computational essay, as it aims to facilitate interaction with a specific type of research data, i.e., textual sources. A professor of biblical studies describes a platform for such editions: "[A]s a librarian, I love this because it allows us to take rare books, to digitise them, and then to put an apparatus of commentary around the text" (OP24). Editions will be addressed in the prototype section of this report in greater detail.

3.6.1.3 Audience

Finally, thinking of innovation in terms of audiences, bridges both aspects of form and access, which we discussed earlier. Innovation may improve the communication of research findings and, thus, the perception of research in society.

> And I do think that [...] scientific publishing should go in the direction of using more blog-like things and that we should be publishing and speaking of our research ideas, our research progress, research intermediate results, and our research failures. We should speak more frequently, timely and openly, in order to speed up [and] improve scientific work worldwide in any way. We don't exchange enough information and not well enough, and that's wrong. (OP30)

Unconventional formats allow us to reach new audiences and help reconnect research and society, showing the importance of the work being done in academia and how the taxpayers' money is being spent: "the more people understand that academics are not in an ivory tower, then the more likely you are going to see funding for the humanities, funding for research is not a waste of time and money" (OP01). Thus, innovation allows us to communicate with audiences in an attractive way, attuned to the contemporary media landscape. In other words, thanks to innovation, research speaks the same language as the public.

> These are non-conventional academic texts that are being read by the public and they're much more accessible than your standard research paper in that

> they do not require as much learning to be able to get the gist of, like, where this particular piece of research is going. (OP01)

This history scholar, working on the innovative dissemination of her project results through an interactive website, describes the issue of navigating between the level of scholarly detail and accessibility for wider audiences. Hence, the aim "is to speak in a general way that the public can understand, but also [in a way that] academic people will be interested in it. But it's not too dumbed down for the academics, but not too, kind, of highbrow for the general audience" (OP28).

Reaching new audiences may also mean making scholarly content available to countries and communities in which traditional forms of scholarly communication, closely bound to the market, are inaccessible for economic reasons.

> When you start to take the book out of the marketplace or take scholarship out of the marketplace, then you realise that the audience can be whoever you want it to be. And that's simply because you're no longer writing for [a] financial kind of gain, or for the publishers to gain financially, or for the book to look like it [is], sort of, a commodity. (OP10)

Hence, new modes of publishing may increase the readership and societal impact of scholarly outputs.

Blogs are frequently considered to be an innovation that allows ideas to reach wider audiences, as they do not try to remediate scientific articles or monographs, but rather serve as a vehicle for lighter and shorter texts. They may be used as an entry point to research or, also, to other disciplines.

> I also like very much reading blog posts and not so much from my narrow field of research but from other fields that are not too familiar to me, which explain things to me in a bit more [of a] popular way; so for me to see if it works for me or not, or what directions I should go to find some connection with my research and so on. (OP18)

Blogs, in this context, almost serve as popular abstracts of more complicated works. As a psychology postdoc put it plainly:

> That's why you have things like blogs and portals and scientific outlets […] They take the scientific [paper] which has twenty pages of tables and graphs and data and stuff like that, and they boil it down to two. (OP29)

Apart from making research more accessible, blogs may serve as a place to communicate early thoughts and to work on ideas. A postdoc in information studies, interestingly, treated his blog as a humanities equivalent of an open notebook: "I keep this kind of open notebook, in which I'm just sort of sharing my thoughts regularly about my research and eventually the book that I'm writing will be based on all of these different snippets. But none of the actual blog posts will be in the book. I quite like just releasing my thoughts as I have them" (OP10).

Audio and audiovisual materials perform a similar function. They also seem to be treated as lighter versions of traditional scholarship, but require a certain talent and competence.

> I look with interest at such forms as short podcasts, short video forms, which are terribly difficult for scientists. Because scientists generally don't know how to express themselves in such an engaging, relatively light way – that is a rare talent. […] This alone would also require investment on the part of the institutions, and not just on the shoulders of the scientists themselves – as usual – to learn these different speaking techniques, just as politicians can be taught. (OP12)

Videos were also mentioned as providing an addendum to one's work. For instance, a PhD candidate in digital humanities produced a documentary based on her research (OP06). A professor of information processing mentioned, in this context, short talks on one's own research that may serve as TED-talk-like trailers: "Twenty or thirty seconds video, like a commercial, as a marketing tool, explaining to you what you could find in this paper, may be something to be considered. […] a real person can make you interested in a paper much more than abstract" (OP30).

Podcasts are generally thought to serve a similar role, as this sociology professor remarked: " Presenting it in such a concise way, a cool way, if it's just for a short podcast, but a really short one, a quarter of an hour at the most. In fact it's probably five minutes, as a teaser [...]" (OP12). The production of scholarly podcasts may have intensified during the pandemic. As a French PhD student in information science noted, many of his colleagues had "started to record not only lectures, but sometimes a review of an article or a book. One of my colleagues started – it was just, like, a side project. And he's actually in his 12th or 13th episode" (OP17). This format is described as particularly engaging because it does not require one's full attention, or much screen time, and allows for other activities in the meantime: "I probably would also use podcasts, just listening and not even looking at the picture, but just listening to the voice" (OP18).

Finally, the use of social media was reported in this context. Not only do scholars inform others about their work on social media, but they often use it to communicate their talks, which they later turn into blogs or into articles. So, **we see an interesting communication loop here, in which the thought is discussed and elaborated in a continuous discussion with peers and a wider audiences**. A postdoc in linguistics described her use of these channels, pointing out that they often allowed her to reach different audiences:

> I'm trying to make blog posts out of my Twitter threads. Sometimes referring to the tweet, but [...] I feel like a lot of content gets lost and I really like Twitter. And I also know that people who actually follow me on Twitter and read my stuff don't go on my blog so they don't read my blog. (OP02)

Another innovation regarding engagement with audiences is the **living book,** which allows for the fluidity of the text – allowing for versioning and user interaction. As this postdoc in biblical studies pointed out:

> The idea is that we keep track of variations and make this information visible. As for the difference between a book and a web book [WB – Author's note], the idea is to keep a regular publication rhythm. I publish as soon as I have

> written a chapter. The peer-review will come at the end of the process after the WB is published online. And the WB will be modified following the peer-review comments; and thanks to the versioning the modifications will be visible. (OP21)

This changes the approach to publication from something finished and closed, to something that makes the changes transparent and accessible to readers. Living books spark community discussion, allowing for comments and replies. This form will be discussed in greater detail in the Prototype section.

3.6.2 **Challenges and obstacles**

The disrupting potential of innovation opens up new possibilities, but also appears challenging on many levels. **The actual uptake of novel communication forms is impeded by various factors, among which, quality assessment, prestige, competencies, and the lack of established standards for referencing novel forms seem to be key**.

3.6.2.1 Main challenges

The main problem with innovation is that **we have novel services for communication but not so many of the quality-assessment mechanisms built upon them**. Traditional forms of assessment often prove to be incompatible with the needs and challenges of innovative outputs. This is covered extensively in the report on quality assessment, so we'll limit the present discussion to the most important remarks.

The lack of recognition of innovative forms as scholarly texts impedes innovation. "If you're doing something so new and different, there is, by definition, no audience to say: 'yes, this is a good thing to do,' or 'no, this is not a good thing to do'" (OP24). So the question boils down to the ways of assessing whether a publication is scholarly or not. As one interviewee put it:

> [T]he barrier comes with the question: what is recognised as scholarly writing in academia, and lets you obtain a position? Until recently (but maybe it is changing) the digital, and especially what is not peer-reviewed, does not count as scholarly writing, at least not for career advancement. (OP22)

Hence, it is often pointed out that innovative forms need to have a scholarly apparatus in order to correspond with the established conventions of scholarly writing.

> [...] footnotes, references, data, which are, as far as possible the most accessible, so that one can dive into the text. (OP25)

> What matters is to find a way to keep the prerequisite[s] of scholarly texts (citing your peers, knowing the state of the art), and to integrate a form of scientific validation. (OP21)

> My innovative publications will be taken seriously only if they are accompanied by a traditional bibliography. (OP21)

It is also the broader issue of engaging with the scholarship of others: "The scholarly text is part of a continuum of scholarly texts: it must cite these texts, and take a position on them" (OP22).

The prestige attached to traditional forms tends to have a cooling effect on innovation, as this postdoc puts it: “our reward structures are so embedded in us that I have to write a book that looks like a book” (OP10). A professor of English studies adds: “many scholars tend to be conservative in going where they see prestige, and, so, that reliance on prestige is still of enormous influence” (OP04). **This leads to a strange situation in which the format of the work influences the assessment of the quality of its content**. As one interviewee put it: “it could be cutting edge work with amazing results, amazing data, and it's completely relevant. But they might not be cited because it's just a thesis” (OP06).

That, in turn, creates a sort of vicious circle in which early career researchers receive advice that blocks innovation.

> I think that early career scholars are more exposed and vulnerable, but are also, in many cases, given bad advice by some more senior scholars who tell them to be conservative and not take chances and only do the most prestigious things, when, in fact, many early career researchers who do take chances and publish in a new way and insist on open access can actually have a real impact both on the field and on getting their work [recognised]. (OP04).

Hence, **scholars are afraid to experiment because they want to publish in prestigious venues, which in turn results in a lower number of innovative works and low prestige**. These issues are covered in greater detail in the section on innovation.

There is also the issue of competencies:

> I look with interest at such forms as short podcasts, short video forms, which are terribly difficult for scientists. Because scientists generally don't know how to express themselves in such an engaging, relatively light way – that is a rare talent. (OP12)

Some scholars are reluctant to invest the time in such activities: “I'm not a great fan of this because I do not have time to learn these new forms and so on” (OP18). There is also the matter of senior scholars, who already have their own established ways of publishing and are reluctant to change: “I don't think of myself as a particularly digitally literate person. And I think I know, and I am familiar, and I work with things that are probably already quite established” (OP19). On the other hand, **researchers invested in innovation cannot understand why some scholars refuse to use innovations that could facilitate their work**: “Despite my presentation of Zotero, despite it becoming pretty much universal and students using it, many of my colleagues don’t. They do things by hand […]” (OP30). The issue of competencies is closely linked to the **need for infrastructure, as innovation may be blocked by a lack of relevant technology**.

> [T]his is a huge challenge also for universities, not only for us, because sometimes universities are hosting it themselves. It really needs an updated website, it needs a good technical specialist and this is really not easy. (OP23)

Finally, our interviewees reported **problems with referencing novel sources**, as there are no established standards one can follow. Generally speaking, **the issue**

of how novel sources should be included in a scholarly text is one of the challenges of 21st-century scholarly writing.

> I cite tweets in that book, as well as citing blog posts, as well as citing formal journal articles and scholarly monographs. So the things I cite are not format independent, and I'd say I try and bring the same scrutiny to whichever media form I'm citing. (OP03)

> So I would consider blogs, scholarly texts as well. And also YouTube talks like lectures on YouTube and interviews as well, because there's been interviews that I've referenced either from YouTube or even, like, a blog style interview where it's published in some informal format, and then also unpublished scholarly work. (OP06)

Our **interviewees distinguished between citing novel forms as sources and scholarly references**.

> Well if you are referring to a specific data set – um – why not cite data sets. I don't have a problem with that. Because then you would not be citing somebody's conclusion about somebody's analysis of some data, you would be citing the data directly for example, or you would be citing the blog directly. (OP29)

> I understand that in some cases one needs to make distinctions of a kind, for instance, is something that you cite data, or is it some scholarly output? And we need to make this distinction if we work with evidence. (OP15)

While our interviewees seemed to agree that blog posts, especially if "written more less like an academic paper" (OP18), are citable, they were more reluctant to quote social media posts. In both cases the issue of ephemerality was a noted factor. In one case our interviewee was asked by his supervisor to turn a Tweet with an interesting visualisation into a blog post, a more serious form, so it could be referenced in a paper (OP17). Other examples included Wikipedia, which is considered a good source but problematic for scientific writing: "I think Wikipedia is amazing and I love it! But then my students always ask if they can cite it in their papers and I tend to say no. The only reason is because I know that my colleagues won't like that" (OP24). One interviewee referenced presentation slides in her writing, quoting the researcher's website where they were posted (OP11).

Finally, **interviewees pointed to the need for referencing software**, "to give recognition to the people who actually work on the tools but also just in terms of methodology, it is important for the research to be replicable, reproducible" (OP16). This is also considered to be a strategic issue in crediting infrastructure crucial to the work: "there are funding issues for tools, in part, because, actually, nobody's citing the tools that they have online" (OP24).

3.6.2.2 Stopgap practices

The challenges of novel forms push scholars toward some practices that allow them to have their cake and eat it too; that is, to **take the advantage of innovation while retaining some signs of prestige. We call these practices double referencing and**

double publication; i.e. using a traditional format in place of an innovative one for the very reason of retaining the prestige of the traditional form. These can be considered either as harmful for innovation, or, on a more positive note, as supporting the transition – as stopgap practices during the transformation phase.

In cases of double referencing, there is pressure to find and use traditional forms of publication for referencing: "If you cite something innovative (a video, a recording of a talk), we still feel that we have to cite another traditional publication. But I try not to erase the innovative publication" (OP21). The same interviewee prepared a webbook that was a digital edition of their thesis, because it had to be presented in a traditional form (OP21). One advantage of double publication is that authors have the best of both worlds – the prestige of the publication, but also faster delivery: "If you are a junior scholar you don't have to wait for your publication at the end of the project in maybe 5–6 years" (OP23).

3.6.3 **Summary**

- Innovation is seen as a chance to improve the sharing of ideas with audiences thanks to novel technological affordances. Innovation is also understood to bring seamlessness to the process of using scholarly content, removing obstacles that are unnecessary from the vantage point of current technology. Innovation, then, is understood either in terms of form (novel means of communicating ideas), or access, i.e. accessing content or reaching new audiences.
- Innovation is mostly considered in terms of providing access to more traditional types of outputs. In this context research data and grey literature become more accessible.
- Formal innovations concern moving beyond the mere written word, i.e., accepting expression in other media forms as valid scholarly outputs. A form of a computational essay allows research and data to be linked, allowing for interaction.
- Thinking of innovation in terms of audiences means improving the communication of research findings and, thus, the perception of research in society. This is done through social media and blogs.
- Innovation is impeded by such factors as quality assessment, prestige, competencies, and a lack of established standards for referencing novel forms. The issue of how to use novel sources in a scholarly text is one of the challenges of 21st-century scholarly writing.
- These challenges push scholars toward practices of double referencing and double publication, whereby the traditional publication provides prestige for the novel form.

3.7 OPENNESS AND TRANSPARENCY

By openness and transparency we mean the practices that are aimed at improving access to scholarly outputs and the reproducibility of results supported by new research methods and information technologies (McLaughlin, 2017). So, although the concept of open scholarship has been developed and found its first applications in STEM-focused fields, it is undoubtedly relevant in the social sciences as well, and especially in the humanities. Openness and transparency are certainly key factors in judging the reliability, accuracy, and relevance of the scientific research results reported through scientific publications across all disciplines

3.7.1 **Attitudes toward open access**

Almost all of our respondents unreservedly supported open access to scientific publications. The researchers highlighted a number of benefits that included the ease of finding publications, free access, savings for the institution, an improved time frame for the publication, and improved visibility, readability and citability. They have supported the changes that have occurred due to the pandemic, during which they have had significantly more open content available, but also expressed concerns about going back to the old "normal," i.e., paywalls and closed access.

Arts and humanities' research can have robust public value, "beyond the academy, through interactions with social partners, ultimately expanding and enhancing wider societal capacities and capabilities" (Benneworth, Gulbrandsen, Hazelkorn, and Gibson, 2016). Our respondents were also aware of the importance of openness, showing the researchers' responsibility towards benefiting society.

> You know, that's the world I dream of: [it] is just one where there's, you know, some piece of scholarly research and I can just get it without it being a problem, without having to encounter paywalls, without having to go through a billion and one hoops to get my university to purchase it. (OP03)
>
> It definitely helps to move to open access practices. I know it first-hand, so I do believe in it. And I hope it also helps with the time frame of publications, because we don't have the luxury of waiting so long anymore. (OP07)
>
> So there are a couple of thousand, a few thousand visits, a few thousand readers. And this is the enormous advantage of open access online scholarly publications, that they can find readers. (OP19)
>
> [Open access] should be standard nowadays. It is a commodity; we as researchers expect to have access to research publication, so that I wouldn't connect it with prestige – it's a prerequisite. (OP32)
>
> Open access is considered to be the most important thing, which should improve the current scholarly communication system: That everything would be accessible for free. (OP27)

Some respondents highlighted the important role of scholars in changing the system of scholarly communication, as well as the fact that only scholars have the power to influence change. Examples of institutional initiatives show how scientists, in

collaboration with funders and editors, can build a modern publishing platform to avoid paying the high open access fees to large publishing houses. Still, some scholars are very conservative when it comes to prestige, which has a huge impact on publishing choices. Although researchers have the ability to shape the publishing landscape, "it is still controlled by the publishers" (OP06). Scholars' conservative perception of prestige and the reliance on prestige still has enormous influence, but this is not what is moving publishing forward (OP04).

Many authors will prefer a closed approach or even a printed version of their publication if they use a prestigious book publisher or a prestigious journal. The long history that publishers have is perceived as a confirmation of quality and the basis for undeniable reputation within the scientific community, despite the often very conservative ways of publishing, which does not take advantage of digital technologies or ensure effective distribution of content. So while new publishers are emerging in the field of scientific publishing, with modern approaches to publishing high quality content and innovative business models that ensure low prices or free open access publishing, such publishers are often not considered prestigious enough. Even when scholars want to publish with such publishers, they fear that this could impair their chances of employment, diminish the value of their CV, or reduce their career prospects.

> I mean, for me personally, I think open access is absolutely crucial. But I recognise that there are scholars out there who still believe that open access publications can't have the same prestige [...] But because of those ingrained ideas, I think for many scholars, the prestige of open access publications is still lower than that of the closed access, traditional journal that's been around for one hundred years. (OP04)

Open access publications offer endless possibilities for connecting open content, and taking full advantage of hypertext and web technologies. However, there is a problem with citing content that, at some point, ceases to be publicly available.

> In fact, it was to switch from a system where the reference was traditionally the footnote referring to a book in a library, to a system where you could only use hypertext references, so only existing online and open-access publications. But on the other hand, you could hypertextualise almost one word out of three, meaning that your text could actually be a kind of reading guide across the digital space on the issue you were addressing. In the last few months I've been working on a chapter called "History and Social Sciences." I used a number of open access resources on the history of sociology, classical texts, which give students access [...] For me that is very interesting, and it highlights the well-known effect of a research practice that is 100% online. On the one hand it increases the visibility of what exists, but on the other hand there is the problem of "invisibilisation": what does not exist online does not exist anymore […], they are disappearing out of sight. (OP25)

Nevertheless, one respondent expressed reservations about an open approach as a solution to all problems in scientific communication: "These people think everyone should just [use] open access out of the blue, and then the whole world is cheerful and

beautiful and liberal. That makes me nervous" (OP05). Interestingly, even with such views, the importance of open access is not disputed, but concerns about a possible loss of control have been highlighted. The position that advocates the current state of scientific publishing, which is dominated by large publishers, while "ensuring" access through possibly illegal or other time-consuming activities, nevertheless lacks good arguments.

> But if I'm looking at this from the personal side, I'm saying I'm not damned interested in which way I can get access to an article. Either it's up on SciHub, or if not, I'll write to the author, or it'll be a preprint anyway. So it's not true that it would hold me back. (OP05, P. 12: 1402)

Some views on open access had a more observational character, with the respondent registering some features of the (un)professionalism of the open access journal.

> I find that just purely by the design and layout of the publication and the small number of people on the review board, it makes it look like it's not reputable. So it's not about the open access aspect of it. I think it's the amount of professionalism that's put into it.. (OP06)

A resentful objection to the large gap between the obvious benefits of open access and the present criteria for academic advancement that needs to be fulfilled was expressed by one respondent:

> I would very much like if all my work was openly available and I would very much like to be able to prefer and submit only to such journals, but there are not many in my fields that are recognised as very valuable journals, blah blah […] So this is my problem. And I'm publishing because I have to. So I have to achieve my primary goal. I have to get a point to survive as a scientist. (OP30)

3.7.2 **The role of policies**

3.7.2.1 National and institutional OA policies

The key factors contributing to the increase in open access publications are certainly the open access policies at different levels. Although some authors argue that open access policies in the humanities do not have such a big influence, and that "[t]he demand for openness in the humanities seems to be rooted in scholarship itself" (Knöchelmann, 2019), open access policies definitely impact the pace of implementing open practices in SSH. According to the SPARC Europe report *An Analysis of Open Science Policies in Europe*, v6 from August 2020, there are 14 European states with national OA policies in place, while thirteen further EU states do not yet have active policies in place, "but are known to be developing national approaches" (Alston, Proudman, Sveinsdottir, and Davidson, 2020).

The approaches to the policies themselves are different; some policies oblige the researcher to take an open approach, while others only recommend it. The content of the policies also implies different models, so that in some countries or institutions, researchers have funds available to cover the costs of open access publishing, while in the absence of such funding, they resort to publishing in the so-called diamond

journals, which do not charge for open access or storage in open repositories (green route).

Our respondents pointed out that even in rich countries, the funds provided are not sufficient to cover the growing costs of publishing with large publishing houses.

> The general idea, at least for the [country's] research financing, is that all research should be done in open access whenever possible. So that's one of the things that really puts researchers in a different kind of bind, really, because a lot of the open research or open access research requires severe article processing charges, some of which are covered by your university. **But it's a different kind of extortion game, really, because it's not the individual researchers themselves who are financing, but it's rather the funding bodies that are being exploited here, because some of the processing charges, especially in terms of the added value that people do, are extraordinary and completely out of proportion.** So that's another issue that the European Union and other funding bodies will have to seriously look into – that they're not necessarily breaking the hold of Springer and Elsevier and such major dinosaurs of scholarly publishing **if they're just going to shift the financing issues to other players in the field, because in the end it will just cement their power.** (OP01)
>
> Open access is [...] actually a non-enforced directive of the [national policymaker] that we should try to publish as much open access as possible, but it's non-enforced because, especially for the STEM people – you know, the hard sciences – it's often very hard to find open access journals that, you know, have the same amount of impact as the questionable Nature and Science and that kind of stuff. (OP08)
>
> We have this directive now that we need to put everything that has been funded by public research, [into] open access. (OP17)
>
> I think the biggest influence is – it's those who give money. (OP18)

3.7.2.2 Funders' policies

Certainly, the policies of the funders who want to boost public investment benefits in research through an open approach have proven to be the most effective. The European Commission introduced open access as a pilot project under the FP7 funding framework, while under Horizon 2020 it stated that "each beneficiary must ensure open access to all peer-reviewed scientific publications relating to its results" (European Commission Directorate-General for Research and Innovation, 2017).

The European Commission's recommendations have been followed by other funders such as the Wellcome Trust, and the National Institutes of Health, among others. In 2018, European funders formed cOAlition S, which then adopted Plan S: a policy that mandates open access without embargo for all scholarly publications that result from research funded by public or private grants provided by national, regional, or international research councils and funding bodies. Though criticised and praised in nearly equal measure, Plan S certainly has had a major impact on the development of open access in Europe

Our respondents also recognised that funders and policymakers were the main stakeholders in the process of changes to the scholarly publishing system. Nonetheless, a general dissatisfaction with the open access initiatives encouraged by funders, was expressed by respondents, believing that the initiatives did not change anything in the system of scholarly publishing itself (OP03).

3.7.3 **Economy of open access publications**

One of the questions we asked the respondents was about the economic circumstances of publishing their works, and their ability to cover the potential costs of publication. The answers provided an interesting overview of the differing opinions and experiences related both to paywalled and to open access publications. The main issue raised of affordability was present in both settings.

A significant portion of the content (both books and journals) relevant to SSH researchers is still only available through subscription or purchase, which some universities might not be able to afford. Moreover, the older digitised content (the so-called backfiles that are not part of current subscriptions) and print content are even more likely to be unavailable to many interested users.

One of the most popular remedies to these affordability and accessibility problems (according to the responses in our interviews) are shadow libraries, but the interviewees also indicated that the "possibility of distribution of digital form[s] on the grey or black market" could be a serious threat, especially for small publishing houses (OP32).

3.7.3.1 APCs and BPCs

The public debates on open access in recent years have largely been dominated by considerations about the “author pays” model. A similar predominance was present in our interviews. The great majority of our respondents were familiar with Author Publication Charge (APC) and Book Publication Charge (BPC) models.

Many authors were willing to pay APCs if they had project grants or institutional funding available. Still, in many responses, it was clear that authors were aware of the “green route” to OA, and regarded paying for APCs, instead of self-archiving, as wasted money. Moreover, in several responses, the idea of paying APCs (in the humanities especially) was described as nonsensical or unrealistic.

> I just have no funding for that. So it's not even in the realms of possibility […] for humanities researchers, it's just so far removed from the idea that we ever will have any money to pay, that it just doesn't influence things. (OP10)

And while there was some approval for the idea of paying for publishing in journals, BPCs were considered unacceptable to the majority.

> And for most people, it's an impossible thing to override. […] open access to books in many publications are already gigantic amounts. (OP13)

Interviewees also indicated some other worrisome aspects of the APC/BPC model, which could threaten equity in publishing. BPCs are especially unaffordable for early career researchers who want to publish their theses as a book. They often have

no funds available or insufficient amounts (OP02, OP03). At institutions that have OA mandates, authors are obliged to

> ask their institution to pay for their papers to be published as open access – and they do pay. But what this means, since they are automatically paying the price asked, is that there seems to be little incentive for publishers to lower this price, and I think it's just incredibly […] you know […] exploitational […] the level of these prices. (OP31-138)

The recent concept of transformative agreements is closely related to the APC/BPC model. Again, aside from the many possible positive outcomes of such deals, our interviewees were very cautious about several aspects.

- The overall situation with financing research (and funders mandates) is very **complicated from the individual researcher's perspective**, so much so that publishers are trying to "educate authors into making inquiries about publication fees prior to making a submission" (OP09).
- There are **no transformative agreements for books**: "what's different with books is that we cannot really do this kind of transformative agreement because they usually only work with journals" (OP26).
- Even though equity is proclaimed as being the ideal of open access initiatives, there are **serious indications that the APC/BPC model threatens equity.** "That leaves all the global south in a terrible, terrible situation. Leaves all the independent researchers that do not have access to funds in a terrible situation […] if policies go in that direction […] What they're doing is probably getting a free card for their own work to be accessible for others, but they stop authors from other countries, from more difficult environments, publish[ing] at all" (OP15).

And while the following notion does not fit within the scope of the APC model, it has to be mentioned that there can be other hidden costs that are particularly important in publishing SSH texts, which are often covered by the authors personally: costs of proofreading, language editing, and translating.

3.7.3.2 Other OA business models

Some interviewees voiced ideas that were in favour of a different publishing ecosystem, without paywalls or author charges.

> I really feel like it's important for open access to support not just equity in the distribution of the text, so that everybody can read it, but also equity in terms of participation in the publication. (OP04)

> If it's not subscription revenue and it's not APC's, there's got to be another model. (OP04)

Some advantages of possible alternative business models were mentioned in several of our interviews.

- **A library subscription or membership model:**

> Open Library of Humanities has done a really good job of thinking about library subscriptions and how that kind of membership model can help support all of these publications. (OP04)

- **Diamond journals and public subsidies:**

> [These] don't have lots of journals which are completely open access and also free for research to publish in. But they do exist, and luckily they are some of the top journals in the field […] entirely funded indirectly or directly by public money. (OP17)

- **OA university presses with institutional support:**

> So you can publish a monograph in [press name], which is funded by the university, which again, it's a solution. It is especially valuable for our PhD students. (OP19)

According to our interviewees, it was definitely worth thinking in a non-market oriented way, and looking for a way that publications "can be funded publicly or […] collaboratively through a kind of a commons framework […] And then when you start to take the book out of the marketplace, or take scholarship out of the marketplace, then you realise that the audiences can be whoever you want them to be" (OP10).

3.7.4 **Ethical issues**

Publication ethics are vital for all disciplines. At a time when the number of published papers is skyrocketing, when scholars work in a "publish or perish" atmosphere, when commercial interests dominate the publishing environment, in which open access is more and more present, some existing ethical issues have become more relevant and some brand new ones have appeared.

The most common ethical issues in scientific publishing relevant to SSH are:

- authorship
- conflict of interest
- plagiarism
- research misconduct: data fabrication and falsification
- copyright infringement, intellectual theft
- redundant publications
- timeliness of publication
- "predatory" publishing.

Our respondents reported a problem related to **pre-prints,** and the possibility that they could cause publishers to consider the author's work had already been published. Moreover, preprint policies seem to be a grey area, without clearly defined rules.

> **I originally released [a pre-print] at the same time as submitting it to the journal. And that was [a] kind of anxiety inducing process, because you don't have a clue at that stage what people want.** I hadn't sent it to, literally,

> anyone. I just put it up. It turned out people kind of liked it and that was good. But it was slightly scary [...] because I didn't know what people would think of it. But it was also scary in the sense that I had tried to research the publisher's pre-print policies. And I wasn't entirely sure even after doing so, that it was kind of allowed. (OP10)

One of the interesting case studies described by a respondent was that blogging about intermediary findings was considered by the university faculty as **potential intellectual theft,** and the respondent "**had been asked to take it down, if possible"** (OP07).

"**Predatory journals**" are often highlighted as being the main ethical issue concerning open access. But it should be borne in mind that there is a wide range of editorial policies with regard to quality assurance and that the term "predatory journal" is less and less used. Some publications can be misleading and false by not providing a decent peer review process, and such publications are definitely an ethical concern. However, there are journals with lower quality editorial processes, but it does not immediately imply that they are predatory journals. We should especially keep in mind apparent "predation" when it comes to commercial, even prestigious publications, which charge seemingly unreasonably high prices for open access publishing, for which we can hardly find a justification. Some respondents reported that they were invited to become editors on the editorial boards of such a suspicious book or journal. Our respondents agreed that predatory publishing is spoiling the good image of open access. For one interviewee this was a source of uncertainty about open science for humanities scholars: "**a large part of the humanities research, in general, still has the idea that open access is directly tied to that, you know – that they only got to know open access through these predatory publishers. And I think that has done an awful lot of bad publicity for open access**" (OP08).

One respondent extrapolated the concept of predatory journals to academic research networks, claiming they tend to operate in a predatory manner:

> Academia.edu, it felt like spam. It was almost like a predatory platform, it was even difficult to leave it. I kept receiving emails, it was difficult to unsubscribe. It's private and [...] there were some issues on how they work with the data that you actually submit. Research Gate, in my understanding, is a bit more of a likeable platform, but still, I find it a bit confusing and a mess. (OP16)

It appears that the word "predatory" seems to refer to practices in scholarly communication that are aimed at maximising profit, while scholarly quality is of lesser importance: "**I think that open access nowadays, unfortunately, means for *me*, in most cases, [the] predatory structure of people, editors, and journals trying to milk money out of people who got funded by some projects, they don't really care what they publish**" (OP29).

3.7.5 **Summary**

- Scholars unreservedly support open access as a tool for improving visibility, showing their responsibility towards benefiting society, changing present prestige patterns, and changing the scholarly communication system.
- There is still a large gap between the apparent benefits of open access and the present criteria for academic career advancement, and scholars fear that publishing in open access can impair their chances of employment, diminish the value of their CV, or reduce their career prospects.
- Open access publications offer endless possibilities for improving traditional publication channels by connecting open content and taking full advantage of hypertext and available technologies.
- Policies requiring that all research should be conducted in open access whenever possible put scholars in a position of balancing between high open access fees in publications that are "prestigious enough" for their career advancement, and the possibility of their works being published in "less prestigious" publications, free of charge.
- Although scientists have no doubts about the advantages of open access, they believe the funders, as the primary stakeholders in the process of changing the scholarly publishing system, should redirect current funding from big publishing houses to other "players on the ground."
- The great majority of our respondents were familiar with the Author Publication Charge (APC) and Book Publication Charge (BPC) models, and they often found current APCs and BPCs "unaffordable," "gigantic," "exploitative," "impossible to override," and "leading to a less equitable world," especially for early career researchers. Open access must support "not only equity in distribution, but equity in terms of participation in publication."
- Many authors are aware of the “green route” to open access, and are opposed to paying APCs instead of self-archiving.
- Possible alternative business models have been proposed: a library subscription or membership model, diamond journals and public subsidies, or OA university presses with institutional support.
- Our respondents agreed that "predatory publishing" is spoiling the good image of open access, and called for a broader understanding of "predation" and a need to differentiate between misleading, deceptive, and false publications; publications with lower quality editorial policies; and predatory practices in scholarly publishing aimed at exclusively maximizing profit.

3.8 AUDIENCE

For a thorough understanding of the nature of scholarly communication, as well as for every attempt to model its future, it is crucial to understand the relationship between authors (and other agents in the communication chain, like publishers) and their audience, both existing and desire. Who are the readers and users of scholarly outputs? How to reach them most efficiently? How can authors benefit from expanding their audience? Can it help their careers? Will they gain prestige, or receive the necessary feedback? What are the best channels to communicate the diverse types of content or to reach diverse types of audiences, scholarly or amateur, national or international? Is there a way to meaningfully measure the impact of the scholarship on its audience? And finally, is every kind communication and interaction equally valuable?

The answers to these questions varied immensely among our interviewees, showing that scholars across SSH disciplines do not always share one common set of fundamental values.

To begin with, when asked about the nature of the scholarly text, many of the respondents opted for definitions containing a reference to the audience, but with a range of diverse and sometimes opposing attitudes. For some, the primary determinant of scholarly text is its direction toward the academic audience: “For me, scholarly text, I think, is really about the audience. […] it's for scholars, by scholars” (OP02), or even further, the absence of any orientation towards an audience:

> [A] scholarly work is a pure kind of expression that doesn't necessarily have to have an audience in mind. It's just – these are my thoughts, these are my data, whatever, I'm going to put them out there. (OP10)

Still, the very fact that a text is scholarly, is not unambiguously regarded as positive for all.

> My initial reaction is that it is something boring, not very useful. Really. I don’t have the impulse: “Oh I want to read it!” I have to force myself to read. (OP30)

On the other hand, in some areas of the arts and humanities, the line between academic and creative outputs and their audiences can be very porous, which is perceived as a welcome expansion of academic space. “[I]n the creative space […] they are trying to get their creative pieces recognised as a scholarly text” (OP06).

The ability to reach wider audiences is one of the fields of innovation, as discussed earlier in this report. Some interviewees look at the innovations that brought scientific knowledge closer to a general audience as an important breakthrough in recent history.

> [W]e had some breakthroughs in certain periods like the breakthroughs that came with the advent of the Web in the 90s […] this was [a] kind of innovation, not in the way papers were prepared, but how they were distributed. Nowadays

> I would connect innovation mostly with the distribution and access to scientific work. (OP32)

More importantly, the way the audience is defined has also broadened, it is not just readership, since reading is not the only possible way to interact with scholarly text anymore.

> I think that the text, the notion of the scholarly text, is starting to open up a bit. It's […] starting to become something that can both be sort of read and consumed in the ways that we have historically read, but it can also be processed, right? It can be data mined for interesting connections and for [the] understandings of the history of a field. (OP04)

3.8.1 **Different venues and audiences**

From the interviews, it is apparent that many respondents are very pragmatic in choosing the best possible venues for reaching different audiences, or even in reaching the same audience but in different contexts. For authors, it seems clear that each publication venue will have a specific readership, that there are different expectations, different discovery paths, and different ways of using the content.

There are many advantages to providing more than one communication channel so the audience will be able to find content via several paths (blogs, Twitter, Google search…).

> I also know that people who actually follow me on Twitter and read my stuff don't go on my blog, so they don't read my blog. […] And also, I think people, when they Google it, don't actually find content on Twitter, but usually would rather read a blog. (OP02)

And for many scholars, expectations are firmly divided between what they see as the potential for traditional publication, and the potential for innovative (or at least non-paper based) formats.

> And if you can produce a creative, interactive visual piece, I think it's really sexy to governments and any other public people […] it just delivers like something that people can immediately absorb and interact with. (OP06)

Furthermore, there are some expectations tied exactly to specific genres, like websites, social media (Facebook, Twitter), or blogs, as described here by a PhD candidate in information sciences:

> You've got lots of freedom with [blog posts] because the audience expects, well, it expects that it's not going to be an article. It expects a bit more personal tone, maybe some information that you wouldn't put in the paper because the wording is a bit crude or controversial or just hesitant or – anything. So, and it's also shorter usually. (OP17)

Many respondents have expectations that new formats and communication channels will more easily find their way to unusual or wider audience, like the general public, journalists, and politicians, but also in terms of geography.

> [Y]ou can then introduce new kinds of audiences, which may be from the global south countries, or […] civil society, or whatever. (OP10)

> These innovative forms can and probably do attract a more general [section of] society. (OP18)

The question is whether new channels will simply expand the existing audience, or if they will enable authors to reach a completely new ones. Some believe that the audience is basically the same, regardless of the medium.

> Someone might do something innovative with the purpose of sharing it with a wider non-academic audience, and someone else might do it with the same audience in mind as for a traditional publication. (OP21)

Others had the opinion that, with new media, there is also a completely new audience.

> [T]he next generation, my students, they would always prefer to have something online, [something] they use on their smartphones and they find it convenient. So I am sure that of these three thousand or so many readers, at least three thousand, would not have read this article if it were not online and open access. (OP19)

Some of these new communication channels are perceived as secondary, only as an addendum or a signpost to the main dissemination venues – something that can help spread the information about important scholarly ideas. Such can be the case with websites, blogs, lectures, etc.

> [My personal website] doesn't have a specific purpose, it offers a way for people to Google me and find me. (OP16)

> [T]hese new publication channels allowed me to gain cooperation, projects, grants, recognition, and so on. But they [function] only as an addition, that is, a medium for propagating something else. (OP13)

There are some characteristics of communication formats that make them suitable for certain types of uses or for specific audiences. For instance, audio recordings or podcasts can be “quite engaging […], I don't have to focus on the screen so I can do lots of other things and still keep up with [the] information” (OP17). And more importantly, they can be more “accessible in terms of disabilities” (OP28).

For many SSH scholars, a relationship with the wider, non-scholarly audience is very important. Many approve of the current developments, where new possibilities for engaging with the public are emerging that highlight the importance of connecting with audiences:

> Scholarship, in the traditional sense, I think, is a dying art. So I think one of the great goals of innovative scholarship is to grow an audience and to demonstrate the relevance of the work for the non-academic audience. There's always the dream of the academic to write for the broader public. (OP24)

However, it should be noted that scholars still consider other scholars to be the main audience for their works:

> I think it has potential to have a massive audience, but the people who tend to view it and analyse it and appreciate it are academics, because the amount of creative thought that's gone into it is usually driven by academic theory and practice. So normally the public won't get everything out of it that the academics have put into it. […] So, I think these works are mostly valued by academics, but they could be appreciated by others. (OP06)

3.8.2 **Scholarly community**

In the previous paragraph, we described the different possibilities for broadening the audience of scholarly works, while concluding that for most scholars, their primary readership remains within the community of their peers: "What matters for me is the scholarly community I am actually participating in" (OP01). The advancement of scholarship through dialogue is possible, primarily, among fellow scientists.

> [T]hese are the people with or against whom these new discoveries, say, or new interpretations of reality, are made. So in this respect it is most important, because these are the people with whom you work directly, to whom you talk directly, gather reviews, and so on. (OP12)

Within the scholarly audience, individuals can differ in the way their ideas are received: "someone who is a good populariser is not necessarily a good researcher" (OP25). There are great researchers who are also great communicators. But then "there are the authors who are hyper-cited and who play a secondary role in the academic world" (OP25).

3.8.3 **Professional community**

SSH disciplines vary with regard to the extent to which their research results are applicable in practice, but for some interviewees it is crucial to come out of "the academic bubble at least to an expert bubble" (OP12). The professionals they address could be teachers, educational managers, psychologists, school principals, librarians, documentalists, archivists, journalists, or policymakers. The results of ongoing research can be presented in public speeches, workshops, or lectures. Collaboration and communication with professionals can have practical advantages for scientists and their institutions, especially if they are also university teachers.

> [I]f you teach and you need some professionals to come in and do some part of a course for your students because you need some real expertise from the field, then you have some people to call. (OP17)

However, research institutions and universities will not always acknowledge this kind of engagement: “I think the institution has a difficult time recognizing a diversity of scholarly outputs” (OP24).

The innovations that result from scholarly research can have a beneficial impact on professional spheres. While they will be welcomed and embraced by most, it is possible that some professionals will see them as a threat and will be reluctant to engage with innovative practices. They will “enter [a] sort of competitive relation because they will want to be more innovative,” or will perceive them as disruptive (OP17).

3.8.4 **Popular publications**

When looking for the right publication venue for their outputs, scholars will not always be exclusive in their choices. Writing for publications specialising in popular opinion pieces, where academics can engage with a wider public and do public outreach, can be intrinsically rewarding, but can also be a source of prestige (see the section on prestige).

For some topics (for instance in literature studies or archaeology), researchers will look for journals that are not purely scientific but have the potential to address a wider professional and amateur audience. However, it is very often the case that societal impact does not correlate with the bibliometric one, therefore journals with high visibility in wider audiences probably will not have high impact factors.

3.8.5 **Style**

Among many of the interviewees, there was the general understanding that most scholarship is written in a dry and complicated language, incomprehensible to the wider audience. As this post-doc in biblical studies put it: “the way that scholars write – the general public doesn't want to read [it]” (OP24). Whether this is a positive or negative fact is not a matter of consensus. Some are frustrated about the illegibility of their texts: “I find it so frustrating when I mention the title of my dissertation and people instantly switch off like ‘this is too difficult for us to understand’” (OP07).

Others will point to the fact that a technical and dry style is inevitably connected to correctness, detail, and scientific rigour; or that even writing casually will not enable understanding without the reader being highly trained in a particular discipline.

> I think I write quite well […] that I'm quite eloquent and it reads as a story because I really try not to be, you know, like academic. […] But still, even with all the jokes, you actually really need to be highly trained in linguistics to get the point (OP02)

Articles written in collaboration are even more likely to be written in a very dry and technically scientific language, as was described earlier in the section on collaborative writing.

New communication forms and channels allow for a more casual style of conversation, and a more free engagement with the audience, but could require special effort and talent.

> You've got lots of freedom with [blog posts] because the audience expects, […] a bit more personal tone, maybe some information that you wouldn't put in the paper because the wording is a bit crude or controversial or just hesitant or – anything. (OP17)

> I look with interest at such forms as short podcasts, short video forms, which are terribly difficult for scientists. Because scientists generally don't know how to express themselves in such an engaging, relatively light way – that is a rare talent. (OP12)

3.8.6 **Reuse**

The free availability and open access of scholarly work for a wider audience enhances the possibility of reusing, translating, or creating derivative content; and some of our respondents perceived that as a great benefit: "I really like the fact that people have transcribed some of my books into Wiki sources, that you might get an unauthorised translation of them" (OP03). In that context, using open licenses can foster the reuse of scholarly content (especially for educational purposes).

> I think whenever we teach a course where we take excerpts of people's work and juxtapose it with other people's stuff, that we're creating a new kind of hybrid tapestry of the scholarly landscape. (OP03)

3.8.7 **Feedback**

Many interviewees expressed hopes for getting feedback from their audience, exchanging comments and engaging in discussions.

> I would like as many people as possible to read my papers and to comment, which is very rare. I would like to get emails from them and [have] discussions on my paper. (OP30)

Besides the absence of feedback, another possible negative outcome is adverse public reaction fuelled by misunderstanding.

> [No] wonder that some scientists […] are quite reluctant to talk to the general public […] there's quite a number of negative comments on the scientist, um […] during COVID and other things. Because, people don't understand science. (OP30)

Books and book presentation events generate a unique kind of response – they can initiate discussions and reviews, which can be profoundly rewarding for their authors.

> I think I managed to provoke such unofficial Facebook reviews somehow. Somewhere someone was wide open and he thought it was good. (OP14)

> I managed to make such an evening in the theatre for over a hundred people. And it was an event that gave me satisfaction. (OP14)

Non-traditional online platforms (for instance academic blogging platforms) can bring additional feedback and engagement options with new roles for audience members.

> In those web platforms, the readership, the community, has a huge role, as a kind of corrective, with the option to comment [on] works. (OP32)

3.8.8 **Value of popularisation activities**

The motivation for SSH authors to address audiences outside of their primary scholarly network is related to their individual set of values as well as to prevailing community perceptions of popularisation activities. For many, it is regarded as an important aspect of scholars' work, which can even become a source of prestige within academia.

> Writing for a non-academic audience is a real, real source of prestige and symbolic remuneration. (OP25)

> I don't know if it's academic prestige, but I think it's also part of our job to be a public figure as well. (OP27)

But there are dissenting voices, raising concerns that popularisation will not be regarded as serious scholarship.

> [H]e is seen as a lower quality researcher or an inferior scientist because he's not actually doing a lot of empirical work, but he is doing science communication more. (OP01)

The question remains: what is the value of popularisation and of widening the possible audience base for scholarly outputs? There were several answers offered:

> ***It is a contribution to the public good*:** the humanities are going through this anxiety, why do we exist? Why should research happen? So we need a tap on a shoulder that we need really to make sure that we contribute to society. (OP15)

> ***It provides feedback and enriches research results*:** what's going on in the humanities, […] especially if it's in concert and being done with local communities, with other kinds of organisations, [this] is really valuable. [It n]ot only feeds into the research outputs, it makes and enriches them, but actually feeds back into teaching and education. (OP09)

> ***It legitimises the public funding of science*:** since we do things that are from the budget that all citizens pay for, it would generally be good for those people who aren't necessarily involved in science to feel that their money goes to something that is somehow important. (OP12)

> ***It brings visibility to the institutions*:** my employer has been cooperating with a PR company for science institutions for a long time, precisely in order to increase the institution's visibility. (OP12)

> ***It creates opportunities to take part in public policy-making*:** participating in various discussions about the topic and going around the country and giving workshops and lectures, you are somehow more noticed by, for example, the Ministry of Education – you are involved in policy-making processes. (OP18)

> ***It builds trust in scientific knowledge among a wider audience*:** Especially today, in the context of coronavirus, we see how important it is to have reliable and valid insight from experts and how important it is to know how to communicate sophisticated research findings and complex knowledge to a broader public of non-experts. (OP32)

> ***It can be deeply and intrinsically rewarding on the individual level:*** it's sometimes hugely rewarding, writing something that you believe can be educational, can help or provoke discussion in a public audience. (OP31)

If popularisation activities and widening the audience base are perceived as being valuable in the research community, it would be expected that they are also perceived as such in various evaluation systems. So, how do they score in various assessment procedures? In many cases, the non-traditional and popular communication outputs are not registered, or are considered inferior, primarily because they are not easily quantified and measured.

> I really am worried that these kinds of nontraditional publications do not get registered in the scientific system with the same kind of merits as a traditional one right now […], and especially, hiring bodies do not understand the impact these make. And because they are not easily quantified, they are invisible to the system. (OP01)

> I'm not sure that […] we have good tools to, sort of, distinguish between really important contributions to public life and something that is basically irrelevant. (OP31)

However, having a public reputation can help in applying for a position or a research grant, even if it is not a requirement: "if a name is known, it might immediately catch the eye among the 80 applications or however many are received" (OP20). Moreover, public engagement can be perceived very favourably by funding bodies, especially in certain humanities disciplines.

> So anybody that's going after research funding in the humanities has to be very clued into that public engagement side of the work that they're doing. And it's not just, sort of, communicating out, it is how they're actually interacting with the public to do that work in the first instance. (OP09)

3.8.9 **Visibility and discoverability**

Considerations about the audience and widening the communication potential of scholarship lead us to issues of the visibility and discoverability of scholarly outputs.

Their importance was emphasised in many interviews, so much so that we can introduce the concept of "**economy of attention**" (OP14). Since visibility and discoverability will determine the size of the audience, we were interested in factors that can contribute to their enhancement. From the interviewees' responses, we discerned several possible factors:

> ***Online presence*:**
> things available online are then also strongly, or more strongly, received. (OP20)
>
> ***Coverage in citation and bibliographic databases such as Scopus or Web of Science:***
> not everything is indexed in every possible platform and, therefore, journals that might not be indexed in the major databases like Scopus or Web of Science – those will be problematic in terms of, like, how much of a reach it will have in the scholarly community. (OP01)
>
> ***Discoverability by Google or Google Scholar:***
>
> whatever pops up in Google Scholar, and then you have it, and those things will get cited no matter what. (OP01)
>
> ***Deposit in open access repositories (can also provide trust):***
> Institutional repositories are also interesting, as they give additional trust to the source" (OP32).
>
> ***Affordability, because not every academic institution will be able to subscribe to every needed resource:***
> you have to have a very good subscription indeed to find arguments, or know somebody at a library who has access to that stuff. (OP01)
>
> ***It also relates to the problem of the "deluge of material" that is published:***
> There is an awful, awful amount of text that is being put out and nobody in their right mind can read everything, even in very, very narrow fields. (OP01)
>
> ***Quality of metadata and using Persistent Identifiers:***
> I think that [PIDs] are major drivers of accessibility. (OP01)
>
> ***Increasing visibility with the help of social media:***
> discussions or recommendations on Twitter can do a lot – both to reach out to other scientists with your own ideas and to spread your own publications. (OP13)
>
> ***Using public online profiles:***
> The cloud version of a reference manager system like Mendeley and Zotero are also huge communities nowadays […], so it is wise to have a profile there. (OP32)

> ***Publishers assisting authors in outreach activities:*** there are a number of different tools that our marketing and communications teams have put together to help authors. (OP09)

3.8.10 Accessibility beyond paywalls

One of the factors that contributes to the visibility of content that we need to emphasise is open access:

> [OA books] have at least 30 to 40 per cent more downloads and more usage than books that are similar in their field that are not open access. (OP26)

More than just advocating open availability, our respondents suggested reconsidering the whole market-oriented context of the scholarly communication chain. We address those issues in the section on open access and the publishing economy.

3.8.11 Criteria for selection

One challenge that our respondents often faced was the need to make choices when discovering new scholarly content. They were often selective in choosing what to read, mainly due to time constraints or because of the abundance of materials being published.

> [I]t's much more important to actually target audiences correctly, because of the sheer amount of information, there is so many niche topics that people just don't have the time to read. (OP07)

The criteria on which scholars base their selections can vary. First, there are **geographical, but consequently, cultural criteria**, due to the different economic, political, and social contexts, some topics are not equally relevant to European and US audiences: “the academic debate on Twitter is so US based and so on, and so mono cultural, […] it makes no sense in the European market” (OP02). Second, a **refusal to accept some services:** “I never take anything from Academia (Academia.edu). Really, it's bullshit” (OP02). Finally, **institutional or regional affiliation:** “I follow institutions and people who are active in the region in the field of regional history, regional culture, historical regional studies, […] because this is central for me” (OP20).

Authors are sometimes also selective about the audiences they address and the venues they choose, and the reasons can again be time constraints.

> This is one of the reasons I stopped publishing. Because it just took too much time. There were days when I got 20 emails a day for the blog. (OP13)

But even political context can be important when someone is deciding on disseminating research results to a wider audience: “I'm very careful about publicly disseminating the findings from my studies because of, more or less, political issues” (OP31). This interviewee, for instance, refers to a personal experience, when after participating in the development and promotion of a new health education curriculum, he was exposed to a series of serious personal attacks in public.

Not just authors, but publishers also must be selective, and consider the potential audience as an important criterion when defining their selection policies if they want to keep a distinctive profile and serve their community.

3.8.12 Referencing

Citations and references can serve as useful indicators of how someone's work will be received by a scholarly audience; but this indicator is seldom unambiguous, because different social and non-academic motives can distort the referencing activities. Therefore, we must be careful when coming to conclusions about the reception of a scholarly output, based on the citations it receives (see the section on power structures for a broader discussion).

3.8.13 Societal vs bibliometric impact

For many academics, finding the balance between reaching out to a scholarly audience and to the wider public can be a challenge throughout their career.

> Some people are better at talking to the public than writing top-quality research. But again, we need all of those things and it's difficult to compare them. (OP16)

> Well, it depends at what point in your career you do that. […] If writing a popular science book or something like that delays your dissertation by three years and stretches your CV, then it's a problem. […] Now, this is such a difficult line to walk. (OP20)

As well as at the individual level, the balance between seeking societal impact and academic impact needs to be made by institutions too.

> [I]f you went to our provost and explained that, they would say that it's not scholarship. But if you're within our narrow world, it absolutely is scholarship. I think the institution has a difficult time recognizing a diversity of scholarly outputs. (OP24)

3.8.14 Interdisciplinary research

In the course of the interviews, several respondents brought the special problems concerning audiences for interdisciplinary research to our attention. Due to the interdisciplinary and collaborative nature of their research, it is not unusual that authors are not sure who their audience is.

> So that's one of the drawbacks of writing collaboratively in this way, and writing in a kind of cross-disciplinary way, so we don't really know where we fit. It would be hard to know where we would actually promote this directly to the audience that we have in mind, because, frankly, we don't have an audience in mind. (OP10)

And finding the right proxy (a publisher) who will be able to target the right audience can also be a challenge.

> But finding a publishing house for such a book is very difficult. […] books at the intersection of fields/disciplines are very difficult to find a place for. (OP13)

3.8.15 National vs. international

For those authors whose primary language is not English, one of the most important decisions that has to made when writing and choosing a publication venue concerns the language, reach, and scope of the publication, because these will determine whether the output has a national or international audience. This decision is not always entirely related to the research topic, because even local or regional issues can be presented to a wider international audience.

> I try to publish in English. […] Because even if it concerns Polish issues, I want it to have some sort of overtone and to reach beyond the [Polish] borders. (OP11)

Besides writing in English, reaching an international audience also requires choosing international journals and publishers: "I hope to reach a much larger audience with my research than if I published it in Polish" (OP11). Besides publishing originally in the English language, another option for reaching a wider audience is translating or opting for multilingual editions: "We decided to do it both in English and in the authors' language, to have a wider audience, and we would take charge of the translation (or of the translation costs)" (OP22).

However, in making a decision on the language and the publication venue, authors often not only have to consider the intended audience, but also the different evaluation systems. In many European countries, assessment criteria are changing, and authors are pushed more and more towards publishing in foreign languages and in journals indexed in Web of Science, Scopus, or ERIH (even though the very fact of being listed in such a database cannot serve as an assurance that the publication is really reputable).

3.8.16 Summary

- Authors are usually pragmatic when choosing the best possible venues for reaching different audiences. They are aware of the various discovery paths and the different ways of using content.
- Many respondents have expectations that new formats and communication channels will more easily find their ways to new or wider audiences, even though the primary addressee for scholarly works remains the community of their peers.
- Writing for popular publications can be a source of prestige, although societal impact will not always correlate with bibliometric impact.
- There is the general understanding that most scholarship is written in a dry and complicated language, incomprehensible to the wider audience, but that new communication forms and channels will allow a more casual style of conversation and a greater freedom of engagement with the audience.

- The free availability and open access of scholarly work for the wider audience enhances the possibility of reusing, translating, or creating derivative content, and of receiving feedback.
- Popularisation activities are regarded as an important aspect of scholars' work, which can even become a source of prestige within academia; but concerns have been raised that non-traditional and popular communication outputs would be considered inferior (primarily because they are not easily quantified and measured in the various assessment systems).
- The origins of the perceived value of popularisation, and a widening of the potential audience are multiple: they contribute to the public good, provide feedback and enrich research results, legitimise the public funding of science, bring visibility to the institutions, create opportunities to take part in public policy-making, build trust in scientific knowledge, and finally, they can be deeply intrinsically rewarding at the individual's level.
- Factors that can contribute to the visibility and discoverability of scholarly outputs by their potential audience are: online presence, open availability and affordability, coverage in citation and bibliographic databases or search engines, quality metadata, using PIDs, help of social media, and public online profiles.
- Various social and non-academic motives can distort referencing behaviour. Therefore, we must be careful when coming to conclusions about the reception of a scholarly output based on the its citations.
- In their decision on language and publication venue (national vs. international), authors often do not only have to consider their intended audience, but also different evaluation systems.

3.9 POWER STRUCTURES

Since the main ambition of this report is to analyse existing innovative practices and the potential for future innovation in scholarly communication, it seemed crucial to explore the existing power structures in this area. The interviewees were very much aware that there were power structures within the world of academia and within the landscape of scholarly communication. However, what differed was the perceived importance of the different groups or actors and the emphasis the respondents placed on various processes.

3.9.1 Power structures in scholarly communication and publishing

We asked the interviewees who they found to be the most powerful actors or groups within the SSH scholarly communication and publishing landscape. Our special interest was on who could facilitate or hinder innovation. We also wanted to know who they thought the main gatekeepers were, especially in the context of peer review (for a detailed analysis of this aspect see D6.6 *Report on Quality Assessment of Innovative Research in SSH*). Below are some insights about the perceived role of the various

actors: the research community, editors and reviewers, institutions and funders, and publishers.

3.9.1.1 The research community as the gatekeeper of tradition?

Unsurprisingly, numerous respondents saw the scholarly community, as a whole, as a powerful collective with real influence on the publishing processes. Moreover, many researchers perceive scholars as a conservative group who prevent newcomers from experimenting, whereas some others also saw the community as a place where new ideas could be discussed, resulting in innovation and changes in practice.

The guardians of traditional forms. It is clear that a lot of researchers see academic culture as quite conservative. Moreover, many think that these attitudes were passed on to the early career researchers (ECRs), who started following the same rules they had been taught.

> [...] if the subject culture of monographs in history is such that the printed monograph is most important, and not the monograph published online or published exclusively in digital form. Then this naturally continues in the younger generations, because the doctor, mothers, and fathers also advise accordingly, because the funding organizations perhaps prefer to support print publications rather than digital publications, and so on and so on. (OP20)

Some saw the potential for innovation in the hands of ECRs, but found it impossible to persuade senior academics that innovative writing forms should be properly valued in academia. However, they suggested things might change with time as generational change takes place.

> It's going to take that long for these types of outputs to be recognised. I think we still have a few generations of older scholars who would never recognise a blog post as something valuable. So, I think we sort of have to let them speak for themselves a little while longer. (OP28)

The innovators. There were also positive voices, presenting scholars as a group that could really change academic publishing. Digital humanities, as a discipline, was brought up as a positive example.

> [...] there is a space for bottom-up engagement in scholarly communities: I think it is the case in the DH community, where there has been a lot of reflection from the start about questions of open access, inequalities, and internal validation within the community. (OP21)

3.9.1.2 Editors choose reviewers

Editors were often described as "absolutely the strongest group" and "gatekeepers" (OP13) in SSH scholarly communication. A strong argument that was brought up by some respondents was that it was the members of the editorial boards of journals who "determine[d] who will peer-review the paper" (OP27), or "have the most influence on

how [the peer review process] will look in fact (OP18). Thus, assuming that the process of peer review is important in decision making, selecting the people who will perform it is even more crucial.

However, most respondents also saw reviewers as “an important part of the process” (OP15), since they were the ones assessing the newly submitted research and might greatly contribute to its quality (see D6.6 *Report on Quality Assessment of Innovative Research in SSH*).

3.9.1.3 Institutions and funders

The interviewees often referred to institutions, both larger European and national bodies such as funding agencies, and entities like universities: “all the people who are involved in the process of assessing people's research within the academic hierarchy; these are quite important stakeholders as well, and I would rank as the third [most] important stakeholder, of course, in terms of the larger grants, the funded research, etc. funders – especially the European Commission" (OP15). Occasionally, personal stories were shared where the organisation’s regulations or restrictions had been biased against innovative solutions.

> A couple of years ago, we applied – there is the Bulgarian [funding agency] program. They have a special program to fund scholarly journals. And we applied for funding and they refused us because the journal did not have a paper version. (OP19)

> I could mention this fact that when I was trying to come up with the format for my dissertation, I very much wanted to make it 70 per cent digital and only 30 per cent in academic writing form because it's a full blown edition. But then we had to negotiate with the university and the standards that are present in the literature department. So we had to stop at the reverse. So 70 per cent was in the traditional form of a dissertation and 30 per cent only as a test case edition. (OP07)

Sometimes the respondents mentioned the national funding bodies who support open access: “You're not penalised in any way if you don't publish in open access. But it's ever more stressed that it's important to do that and to, you know, publish open access, if possible, and non-OA only if really there's compelling reasons not to” (OP08). There was a consensus that funders could promote OA through their rules and regulations but not all solutions adopted by funding agencies were seen as equally valuable. For example, a Polish researcher criticised the guidelines endorsed by the National Science Centre in Poland: “I'm against what the NSC did. I think forcing people to pay for open access is a radical misunderstanding of how it works” (OP13).

3.9.1.4 Publishers

Although less frequently mentioned than the editorial boards, publishers also seemed to form a strong group that the interviewees referred to. This was particularly evident in the case of discussions about scientific monographs.

> If we are talking about books then, first of all, it is the publisher and editors that are connected with a certain publishing house. It could be a non-profit publisher like the institutional publisher of an academic institution then, it is governing boards and editors connected with that specific publisher. (OP32)

Some also saw publishers as potential innovators, the group who could truly bring change to the scholarly communication landscape.

> [...] scholars still have the most influence based on their decisions about where they're going to publish and what kinds of venues they're interested in. [...] not all scholars, obviously, but many scholars tend to be conservative in going where they see prestige and so that reliance on prestige is still of enormous influence. But it's not what's moving publishing forward. I think what's moving publishing forward is a series of initiatives that are coming from publishers, from libraries, and from scholars who are working with publishers and libraries to create new forms. (OP04)

3.9.2 The complex world of power structures

Sometimes it was difficult for the interviewees to specify who the main players were in SSH scholarly communication power structures, or else they listed many different actors. This can be explained in two ways.

Lack of transparency. One researcher specified that he was unclear about who held the most power in scholarly communication processes: “If I knew what was going on behind closed doors, I would have a better answer for that. Right now, I'm just sensing the power dynamics and how they impact [on] me, myself” (OP01). In many cases there seems to be no clear, transparent rules that all institutions, publishers, journals, or funders followed. Many processes seem to differ depending on the people involved and internal agreements (see D6.6 *Report on Quality Assessment of Innovative Research in SSH* where this issue is discussed in reference to peer review).

The interdependence of various players. What was also apparent in many of the interviews is that the researchers saw the publishing landscape as a complex space with many interconnections, where various stakeholders did not act independently from each other, and the decisions of one group affected all the others.

> [...] they all [influence ...] one another. What funders decide to fund is key in shaping what the future of the discipline looks like; but that is determined by [the] researchers who conduct the peer review. And they conduct that at different stages in their career, facing different pressures at different points.[...] in some disciplines, there are editors at particular book series, who have huge influence on what is published and shaped – the discipline in that space. In other disciplines, it's a journal peer review process that really doesn't have that single point of editorial acquisition or oversight. [...] I think it's a network that conditions how, what the future looks like and you've got lots of moving parts that are co-dependent on one another. (OP03)

This, however, does not need to be a criticism. One of the interviewees expressed hope for the future of the humanities, arguing that, instead of having a landscape of a few strong actors, the power might lie in the hands of successful alliances, bringing together different partners and representing more diverse needs.

> [...] I don't think it's the publishers that have the most strength. I think it might be more those communities of practice. It might be the fantastic organisations like DARIAH and others who are building that capacity amongst a range of players. And I think, for me, it feels like the strength in change will come from more of those partnerships, where there's an alliance, where there's more connectivity. There's a bit of a collective action that might be able to be initiated through bringing together various different stakeholders. I think it's that multistakeholder dimension that feels more productive in terms of strength to make a change rather than having one strongest player. (OP09)

3.9.3 Early career researchers: finding a place in academia, becoming a well-known scholar

In the context of innovation, it is common understanding that early career researchers (ECRs) are the most vulnerable, and perhaps take the most risk when engaging with novel research solutions while being expected to prove themselves in traditional formats: “if you are a young researcher, you have to also do traditional publications in addition to the innovative ones” (OP21). Because ECRs aim to gain recognition within the community, it is important for them to advance in their careers and to ensure that they score well in formal evaluations. The representatives of this group recognised their own fragile position, with one respondent claiming that she would not be interested in signing a peer review with her own name (thus practising open peer review) because she “need[ed] to feel safe as an early career researcher” (OP02).

A difficulty perceived by some of the interviewees lay in the fact that it might be challenging for people without connections or an established position in the academic community to advance in their careers. This is partly because of the publish-or-perish culture, meaning there is an excess of academic texts published in SSH and people need to be very selective about which texts they wish to engage with. Some research may not get the recognition it deserves. One respondent complained:

> There is an awful, awful amount of text that is being put out, and nobody in their right mind can read everything, even in very, very narrow fields [...], and if you're a middle tier scholar, I'm not even talking about the lowest of the low, but if you're a middling scholar, you are just not going to get cited that much because everybody – the publishing system – demands and rewards [the] citing [of the] big names in the field. And then you could have a couple of people who you might be expect to cite because I know they're your colleagues, former colleagues, or good acquaintances. And then genuinely good research that is being done is just getting lost and being left by the wayside. (OP01)

The same respondent referred to the fact that older, more established traditional research is more likely to be repeatedly referred to by scholars:

> That also means that traditional papers, which have been in currency for a much longer period of time, get more cited over time. And younger research, especially, is more overlooked because everybody wants you to cite the 20 experts in the field. (OP01)

Apart from tendencies to cite the "big names" and people one personally knows, the interviewees noted other social factors that can influence citation behaviour.

- Forming **citation networks** based on language, nationality, or geographical location, and excluding those who don't belong, can have a serious impact on scholars' reputations or on the development of the field: "And those American scientists who write about Poland quote American scientists – they are also gatekeepers in fact. So deliberate omission of some positions from scientific communication is a very powerful force" (OP13). As a result, for some authors, the reception of their works (demonstrated in citations they received) might come as a surprise, especially when it accrues over time and outside their known networks (linguistic, national, or disciplinary).
- The importance of the **citation choices of senior and renowned scholars**: " if they point out, quote, [or] praise – they will actually give life to someone or condemn them to oblivion" (OP13).
- The **reluctance** of senior scholars **to cite younger researchers, or publications that are perceived as less reputable:** "you wouldn't get a senior professor probably citing someone's thesis" (OP06).

The main reason mentioned as an explanation for the importance of citations was evaluation, especially in relation to the various criteria for formal assessment: "you will eventually be evaluated by the fact [of] how your papers are cited" (OP01), or "if you choose these journals, you have a chance to be registered in databases such as Scopus and Web of Science. If you are registered and someone quotes you, you will be able to see and show these ratios. And I need that because I care about getting grants. And to get grants, you need extra points for quotations" (OP11). The number of citations received can also have a positive, informal effect on someone's career: "my last piece has been […] cited and read quite a lot. […] I got invited to be a keynote [speaker] by a professor from a university I had never heard of" (OP02). Naturally, the drawbacks of such a system were also observed: "And then genuinely good research that is being done is just getting lost and being left by the wayside" (OP01). Citations can also serve as an indicator of quality, and help readers gain trust in the text they have discovered – and it is not just the number of citations that counts, but also the provenance of the citations received (who is the person who quotes?).

The issue also occurs within the area of contributing to the scholarly community through peer review, where certain groups may get marginalised due to the doctoral

school where they studied and the connections they built (see the discussion in D6.6 *Report on Quality Assessment of Innovative Research in SSH*).

It is within this culture that most ECRs learn about the demands of the academic job. Many are advised by their seniors on where to publish and how to approach their careers in order to excel. This seems to re-establish old structures and traditional expectations for younger generations of scholars. An ECR confessed that he had felt obliged to publish in a subscription-based journal despite believing in open science.

> But even I feel the pressure to release my research in conservative ways. [...] I published an article that was in a subscription journal. I mean, I'm an open access advocate and I wasn't hugely happy about doing that. But it was in a journal that I thought was sufficiently prestigious in order to say that I've done that this year. So it's just a kind of a box I had to tick. (OP10)

However, some respondents were more optimistic and thought that despite their vulnerability, ECRs might initiate innovation in scholarly communication, as "[...] many early career researchers who do take chances and publish in a new way, and insist on open access, and can actually have a real impact both on the field and on getting their work [...] on the level of impact that their work is able to have. Right?" (OP04). Still, in most cases, being an innovator is not enough for an ECR to succeed. They need to combine such activities with other ways of "scoring" in evaluative systems. An interviewee from Poland who had participated in a number of international projects, led large teams, and published in journals with a high impact factor, admitted: "I have not yet met with someone who would build his or her scientific position on innovative forms only and exclusively, if this is not supported by the baggage of scientific publications" (OP13).

3.9.4 **People, metrics and innovation**

It was in the context of the presumed vulnerability of ECR's that one of the respondents pointed out the hypocrisy of the academic world:

> [...] this is like the mafia saying: "Right. I'm really sorry you didn't choose to play by our rules. Now we have to break your legs." It's equivalent. You know, these people who are saying early career researchers are going to be disadvantaged if they do this, are the people who make it so the early career researchers are disadvantaged if they do that." (OP03)

While sometimes assessments or formal evaluations are discussed as if they are somehow independent of the academic community and forced onto it, the people responsible for metrics were occasionally pointed out as being important actors in the scholarly communication landscape. Their role is even more important when one realises that their approach can hinder innovation.

> So it is much more a question of, like, following the money and making sure that you are good with administrators in the eye of people who quantify things, people who are doing metrics. And it's an uphill battle and it really does

> discourage even more innovative forms of dissemination of research knowledge. (OP01)

However, well-known scholars in their field are seen as an opinion-changing force; they may affect the way processes are done. For example one of the interviewees saw their potential in raising the prestige of open access publications (see 3.10 Prestige): “big names in science, if they publish their papers there, it would help” (OP30).

Since innovation can be seen as “a disruption” (OP17), what happens to power in academia if and when researchers increasingly pay attention to novel forms of scholarly communication? Perhaps the broader community of readers becomes the new powerful group and replaces the traditional editors: “If you have a big blogosphere, it works more like a publishing house with a number of editors. In those web platforms the readership, the community, has a huge role, as a kind of corrective with the option to comment on works” (OP32). Here innovation implies not just a possible disruption but a shift in the traditional power structures.

3.9.5 **Summary**

- Researchers themselves, or the community more broadly, are recognised as important actors in the SSH scholarly communication landscape. Depending on their approach, they can play the role of guardians of the status quo or innovation facilitators.
- Our respondents often pointed to editors as being one of the most important groups in academic power structures, especially in the context of publishing, as they select the reviewers for scholarly texts and tend to make the final decisions. Other important groups included institutions, funders and publishers, and the “big names” (scholars recognised in their communities).
- The relative importance of the different groups is often hard to determine. The respondents referred to issues with the power structures’ transparency and complexities as explanations for this difficulty.
- Innovative forms of writing could challenge the traditional structures, giving more gatekeeping power to the wider community of readers.

3.10 PRESTIGE

Another topic that has been important for us to investigate in the context of scholarly writing is the issue of prestige in SSH scholarly communication. We did not ask our respondents directly how they defined prestige or what prestige meant to them, but rather explored such issues as whether certain publication types counted more than others in their academic careers, how they decided if a publication was trustworthy or worth reading, or about the prestige of innovative genres and open access in academia.

However, from the answers it seems clear that the notion of prestige was more general than strict bibliometrics and official evaluations, with one scholar claiming that "at a certain stage, what counts more is the place measured by non-parametric prestige" (OP13). Many of the comments suggested a certain reluctance to accept some elements of prestige. One of the interviewees questioned the choice of the word in reference to innovative forms of writing:

> I don't think that prestige, as a concept, has a lot to offer science [...] Why are we talking about prestige and not about truthfulness? Why are we talking about prestige and not about precision? Why are we talking about prestige and [...], you know, [not] clarity of language? [...] Why would prestige be an important concept in science? [...] I don't think that new forms of media should be ranked by "prestige." (OP29)

Another respondent voiced a similar concern, claiming that prestige was too narrow a concept as it is applied by academics today, and should be broadened if we want to talk about the prestige of innovative forms.

> [...] but it is another type of prestige, connected with the capabilities of using recent technologies or current technologies in a way that solves real problems. I absolutely connect prestige with creators of certain software packages, creators of different modules or plug-ins for a different software, especially if they are distributing that as an open-source solution. We need to take a broader perspective when considering prestige. (OP32)

Prestige has been linked to the idea of scarcity, which often contrasts with the official bibliometric evaluation: "scarcity and how difficult it is to get a publication in[to] a particular venue – so either with a press or at a particular journal. I don't think, often, this is formalised terribly well, so, you know, in my discipline, there's a kind of informal hierarchy of presses and journals" (OP03). Moreover, while prestige might seem a difficult concept to define, there seems to be a contrast between one's own internal views on what is important in academia and the external requirements on the scholarly landscape. Respondents appear to be balanced between the two. While at the start of their careers, young scholars need to look at bibliometrics and requirements, more established scholars may be able to focus a bit more on their own needs and interests. For a more thorough discussion on bibliometrics see D6.6 *Report on Quality Assessment of Innovative Research in SSH*.

3.10.1 Publication types and career advancement

When asked if they thought that some types of publications counted more in their career than others, an overwhelming majority of the respondents agreed that this was the case. There seemed to be a hierarchy of publications that was shared by many of the scholars. An assistant professor from the US summarised it in the following way:

> It's the monograph; the single author monograph is still always going to be king. In our institution, and in American institutions in general, it's less

> important how well received your monograph is than the fact that it was published, and that it was published by a certain number of presses. The second would be co-authored or co-edited volumes, and then third would be peer reviewed journal articles, and then way down on the list would be chapters in an edited volume. We get almost no credit for that. (OP24)

Indeed, unsurprisingly the monograph is very often mentioned as the most prestigious output type in SSH (see the section on SSH specificity), but its importance does not only come from its established position but also from the opportunities that the book's format provides. Sometimes the book is chosen as the appropriate output and the more wholesome approach to a given topic, a more consistent story (see the section on choosing the publication type). One of the interviewees admitted that a collaborative project resulted in a book co-authored by several persons (which is often considered to be a less prestigious form than a single-authored monograph), even though it was not the most "economic" decision in terms of bibliometric results.

> We started out writing it as journal article. And it very soon became clear that we could have written, you know, three or four journal articles. But actually, it was nicer to package it together within a single narrative space of a book. [...]. And I think it gives a nice story as a result. But I also think that it's probably going to hit us in terms of usage metrics. [...] So we might have sacrificed some citations for the sake of the kind of neat packaging of it within a book form, but I'm okay with that. (OP03)

Similarly, another respondent referred to an edited volume she published several years earlier:

> the rewards for publishing a chapter in that volume are just not that high in terms of, like, "what does this count for in your annual merit review" and so forth? It's just not seen as being that important. But that volume is of such importance and has been referred to and is used so much that it turns out to have a much more lasting impact than an average journal article might. (OP04)

The excerpts above show that the prestige and importance of particular publication forms are not always based on bibliometrics but rather on the form that fits the narrative and that reaches the intended audiences. However, generally speaking, most SSH scholars, especially those engaging in traditional research, perceive edited monographs as inferior in terms of prestige. Some think that the decision to publish a particular chapter is not based only on merit because "there are a lot more networking opportunities that precede the publication" (OP01), which could affect the final quality of an edited monograph. Similarly, conference proceedings are much less rewarded (than books or articles) in terms of formal assessment (see the section on specificity of scholarly communication in SSH). Still, the interviewees see the value of engaging with these formats. For example, the important role of conferences is seen, especially, at the start of the ideation or planning process, as it allows one to develop an idea that can later become a larger output that is published in a different form.

> We also do conferences and publish conference proceeding, [...] first, because it's quicker and easier to establish a presence. When you want to go into collaboration with other people, they will ask you: "What are you doing?" So you do need some written background about what you're doing and conference papers serve that purpose quite ok. [...] Another reason is, although it is costly, divided into two sub-reasons: a) You build relations with people whom you may have a potential partnership with for research in the future, and b) I send my younger colleagues to these conferences so they can get to it and simply gain some experience. (OP30)

3.10.2 Publication prestige: what makes an output trustworthy

Although not strictly a matter of prestige, though closely intertwined with it, is the matter of the credibility or reliability of a new publication one encounters in their research. We asked our interviewees to tell us how they made the decision to trust a scholarly output. Usually such decisions are made based on a whole list of factors, including a metric-centred approach.

> [...] the journal and citation counts are important to me; so if I do not know the research, then I look at what journal it is published in, I look at whether other people have cited it before me. That doesn't mean I wouldn't cite something that hasn't been cited before, it just means I give more scrutiny to less credible sources. If it's a prestigious journal, widely-cited paper, then I tend to trust the paper way more. I look at affiliations, so if it's prestigious institutions, I tend to trust those sources more, and [as] the last resort is just that I read it and make up my own mind. But for the preliminary search it's – I pay attention to these external factors. (OP16)

Often different criteria are applied depending on the type of output, and also how the respondent learns about the new text.

> I think I trust the things that I consult, for different reasons, depending on where they're coming from. Things that are informally published online, for instance, blogs and things like that, I tend to develop trust based on knowing who the author is and knowing their prior work. Things that are more formally published, some of the trust comes from the history of that publication venue, whether it's a university press or it's a journal; or just knowing how they operate and the ways that their processes work produces some trust. But a lot of it still comes back to, and this is a little bit apart from trust, it's not just, like, do I think this thing is authoritative and good and valuable, but, like, the kinds of decisions that I make, I mean, I've only got so much time to read things, right? And how I make the decision that this is a source that I want to consult, that I want to spend my time reading this thing, often comes from recommendations and connections of third parties, right? Whether it's a source that somebody else cites in a text that they have published or a recommendation from someone that I trust online saying you should read this thing or that. It's that kind of

> recommendation and reference that will often lead me to take the time to read something and think about it in the context of my work. (OP04)

Authorship is something that is often mentioned in the context of trustworthiness, as the choice of what sources to read is often influenced by the importance of certain researchers in a given discipline: “There is obviously a sense of you learn[ing], you start to learn the names in the field and you learn to know, roughly, what people stand for. And so that might mean you read something in a bit more detail, or it might mean you can use them as a sort of a placeholder for a certain perspective that, you know, that they stand for” (OP10).

Although it is to be expected that recognisable or well-known names in a discipline appear as more trustworthy to readers, this tendency could also bring about the risk of researchers becoming stuck in a disciplinary bubble, where the same group of scholars read and cite each other, and the less experienced researchers with less privileged backgrounds can find it difficult to see a breakthrough in their careers (see subchapter on power structures).

Our interviewees also often referred to the publisher or the journal as a source of trust: “[...]where it's published, like what journal or where it's published. And if I've never heard of the journal or if it's a journal that's kind of – because I'm so interdisciplinary – if it's really outside my discipline, then I try to see if it's the right fit” (OP06). In the case of papers, peer review is also often mentioned, with many scholars admitting they trusted peer reviewed texts more, particularly if these are outside their expertise. A doctoral student from France presented a very strong position: “if it's not peer reviewed, I can't cite it. It's just part of my personal ethos as a researcher, it has to have been reviewed” (OP17). Thus, whether or not the publication has been peer reviewed is important to researchers, although they rarely present as strict a view as the one mentioned. Peer review is discussed more extensively in this context in the D6.6 *Report on Quality Assessment of Innovative Research in SSH*.

3.10.3 The prestige of innovative forms

Some examples of nontraditional publications mentioned were blogs, tweets, podcasts, virtual exhibitions, multimedia, presentation slides, websites, Wikipedia articles, reports, pre-prints, software tools, videos, recordings of talks, comics, and radio programs. Many of the respondents were positive towards innovative forms, but they are often seen as “less prestigious” (OP21), especially in the metric-centric sense, as in not being recognised by official evaluation systems. These are not just theoretical arguments, they are also raised by “innovators” who use these novel forms: “I published a lot on my blog but it doesn't count – at all” (OP02).

An interesting insight came from a doctoral student who engages in innovation but recognises that some scholars may view this not as “progress,” but rather as a “disruption,” something they need to protect themselves against.

> [...] Innovation can be disruptive. All the tools that I'm using and promoting can be very challenging to use for some people who are not at all used to that

> system and who see them as a threat to the efficiency of their process, because they say, “what if I have to change all my workflow to [use] these tools, and I can't use them, and I can't use them well, and they don't bring anything to me,” and all that. (OP17)

The topic of trustworthiness and the prestige of innovative sources is sometimes linked to their citability. Interviewees found non-traditional SSH outputs tricky to quote in official publications:

> And the same for radio shows, it’s really difficult to know how you’re supposed to cite certain things that don’t fit into the traditional journal or book. [...] my students, when they quote websites, they have no idea how to quote them. So, I think we need to embrace the different forms and have a system for citing them. (OP28)

Some forms are perceived to be ephemeral: “tweets or other things are rarely worth quoting. They are rather an ephemeral form that serves to interest someone, and this communication happens somewhere else” (OP13), or the absence of peer review is the obstacle. For software tools or data sets, replicability and reproducibility are the main incentives for citing them.

Several other issues with citing nontraditional forms of scholarship were mentioned:

- the phenomenon of the doubling of effort, discussed earlier in the innovation section: “If you cite something innovative, a video, a recording of a talk, we still feel that we have to cite another, traditional, publication” (OP21);
- tools that are very widely used do not get cited anymore, but that could lead to them not being recognised as worthy of funding to keep them alive;
- citing content that isn’t permanent: “But the problem with citing blogs is, blogs go away. You can cite a blog, and in five years maybe it’s gone” (OP29).

Citation-related difficulties also pose problems for authors who apply novel scholarly forms and go beyond the traditional publishing landscape.

> In terms of how the work is evaluated, then I think the innovative ways have a problem because they don't get the same recognition as the traditional sources, so if I publish a data set, nobody is going to be interested in the numbers of downloads in the same [way] as they are interested in the number of citations of my journal article, or something like that. So if impact is understood in that way, then I think the newer forms are at a great disadvantage [...] One of the issues with innovation is that many of the new platforms that exist don't have the reputation and the prestige of the traditional journals and publishing houses, so in that way it's not recognised. (OP16)

However, it is not always the case that innovative outputs are overlooked in academia. Certain scholars now expect to see more digital solutions in their field and are suspicious if these do not appear, for example, for the sake of transparency: “[...]

even if there is a source and it's been hidden away in a pile, it already begs the question why" (OP07). Some are also keen to find ways to include novel forms in their bibliographies and make an effort to cite them. One of the respondents described the sources he used in his book:

> So I'd like to say that I evaluate everything independently, and I cite [a wide range of] sources. So in my most recent book [...], I cite tweets in that book, as well as citing blog posts, as well as citing formal journal articles and scholarly monographs. So the things I cite are not format independent and I'd say I try and bring the same scrutiny to whichever media form I'm citing. But I also cite, you know, different artefacts in different ways. (OP03)

As a result, a consistent way in which to evaluate these new forms should be applied (see T6.6 report). Yet scholars do not have a ready blueprint for assessing non-traditional research.

> Well, it could be, for example, how many people have watched or – I don't know – sometimes how many put their "like" to the thing, but I'm not sure how objective it is and what it gives [you] when you know that one or two hundred people have watched [it] and if there has been any impact or so. It's hard to tell. (OP18)

Comments made by other researchers under online content or re-shares on social media (such as Twitter) may increase the perceived quality of prestige of innovative forms, and, in the eyes of some researchers, may even be seen as a form of peer review (see D6.6 *Report on Quality Assessment of Innovative Research in SSH*). While social media (especially Twitter) are seen as a good way of disseminating information (see subchapter on audiences), their impact on one's career or position in the academic community is ambiguous.

> I don't think it would impact my career at all; or also, if I disappeared from Twitter. I also don't think it would mean anything [...] That I'm not sure actually, because Twitter has made me very visible, so I'm not sure, but yeah, if I was less active or, I dunno, I don't think it would make a real difference. (OP02)

Furthermore, there are mixed opinions on whether some innovative forms are more prestigious than others (some researchers think otherwise) and anecdotal evidence rather than consistent solutions are usually offered as supporting evidence.

> It depends on, for example, things like collaborating with multiple institutions, working with libraries, getting external funding. Those are the kinds of things that make a project more viable or more legitimate, in my mind, than some others, than a blog post for instance. (OP24)

> I would say that in the environment in which I operate, posting on this London School of Economics blog is often much more important than publishing it in a magazine with an impact factor. (OP13)

Some interviewees also understand open access to be a form of innovation. To explore various views about OA please see the next section.

3.10.4 Open access and prestige

The prestige of open access (OA) publications has been a topic that has generated a wide range of answers. Three distinctive views may be distilled:

1) a positive link between a publication being published in OA and its prestige;
2) a negative relationship between the two;
3) no links between OA and prestige.

3.10.4.1 OA is prestigious

When seeing the links between OA and prestige, one of the arguments is that top-down guidelines and directives sometimes force the opening up of research so the institutions can communicate about their OA achievements in a positive tone.

> We have this directive now that we need to put everything that has been funded by public research [into] open access. [...] there is some prestige in doing that, but more at the laboratory level. So, for example, the lab will say, "this is all our output for the year 2019. It's all open access, only open archives. Look at it. What a good job we've done." (OP17)

Another topic that is brought up is the fact that OA is so easily accessible to wider audiences and, thus, texts published in OA may be cited more often.

> [...] there was a certain prejudice against open access, and in favour of those heavy scholarly journals by commercial publishers, something published by Sage, or Routledge, or Cambridge University Press, or whatever. But for the past [few] years, we can see that open access journals, because they're open access, they really reach a very wide audience. So this is the way to be seen, to be heard. And I think they are becoming, again, they are becoming more prestigious than just a couple of years ago. (OP19)

3.10.4.2 OA is not prestigious

Researchers who refer to OA as unprestigious often do not themselves question the value of OA but rather think that their broader community does not appreciate it. A theology researcher commented:

> In my field, no, it's the opposite. It's not a legitimate scholarship, which is terrible. I think in certain fields, yes, in certain fields you're evil if you're not putting your scholarship out there. But we're not there yet. (OP24)

The issue that is often raised in this context is one of presumed poor quality. One of the respondents commented on the composition of OA journals' editorial boards.

> I find that just purely by design and layout of the publication and the small number of people on the review board, it makes it look like it's not reputable. So it's not about the open access aspect of it. I think it's the amount of professionalism that's put into it. So it could be [that] maybe there's three or

> five people on the board or the editors or that review. But you're just like, well, if three or five people, instead of a huge network of reviewers [...]. So it's just, I think, it's a matter of image, almost, with the current open access that I've seen. (OP06)

There is also a presumption that OA publications are not peer-reviewed and that affects their credibility (see D6.6 *Report on Quality Assessment of Innovative Research in SSH*). Predatory journals, discussed in greater detail in the open-access section of this report, also affect the image of other OA publications and put them in a negative light.

3.10.4.3 There is no relationship between OA and prestige

An interesting viewpoint presented by some respondents was that there are few or no links between OA and prestige. One of the arguments was that **OA is something that ought to be taken as a given, independently of the quality of the publication**.

> The fact that something is open access doesn't necessarily mean it is connected with prestige. It should be standard nowadays. It is a commodity; we, as researchers expect, to have access to research publication, so I wouldn't connect it with prestige – it's a prerequisite. (OP32)

Another argument refers to the format of OA publications (contrasting it with traditional printed texts). Here the assumption is that the **digital OA format has not really altered the essence of SSH publications and, thus, the change has not been particularly revolutionary**.

> The form is not so much influencing the content so far, I have to say. If you have a printed book with a linen cover and gold [lettering], or if you have an OA e-book, it's still the same text. This is again the question: is it just the form, or is it also the content that is changing due to new ways, digital ways, that we have to explore our source material? But I don't see a new dynamic, it is still very old-fashioned. I'm sure there is something that will change. (OP23)

Our respondents recognised that there were different extents to which researchers cared about open access. It is also clear that even advocates of OA are sometimes willing to compromise if an exciting non-OA opportunity comes along. An interviewee who used to pay a lot of attention to publishing in OA during her doctoral studies admits that the situation shifted once she obtained her PhD and started collaborating with other scholars.

> And it's a bit different because when you get someone who asks you to write, I don't know, I usually don't raise the open access issue to be honest, because I'm already flattered and I'm still quite a, you know, at the beginning of my postdoc career. (OP02)

A senior researcher also referred to certain opportunities in academic life that seem exciting but unfortunately do not result in OA publications:

> But sometimes we cannot choose. For example, there's a good author and a good team, and a very interesting concept and I've heard of publishing houses that are respected in academia, like Routledge itself and so on, and that's – I can't decide what to do. (OP18)

3.10.4.4 Potential to increase OA's prestige

How can open access be promoted in SSH? As in the case of most of the solutions suggested by our respondents, there are no universally accepted answers. For example, a senior researcher shared a plan he developed at an individual level and described how he balances publishing in high impact journals (which, in his discipline, are often not OA) with keeping his research as open as possible for most of the time:

> And so you have to, you know, have a little strategy, I guess, that you try to publish an article once every while in a high, highly visible journal like Digital Scholarship, and put most of the other stuff in nice open access journals. (OP08)

Other respondents saw hope in the research community (if more authors publish in OA it will become more prestigious) and in the funders who may adjust their requirements to meet the standards of open scholarship.

> I think the only way to change this is by the numbers. When the majority of places where people are publishing at are accepted as a place to publish open access then it will become a normal thing and perception will change. (OP30)

> The only thing I'd say is that the more that funders insist upon open access, and pay for it, the more it becomes normalised – that high quality, project funded work is open access. So if you see work in the medical humanities that's been funded by the Wellcome Trust, for instance, it will be open access, and people start to say, "Oh, yeah, I was funded by Wellcome," so it's probably good work. And it's open access because it's funded by Wellcome. So they sort of associate the open status with project funding. (OP03)

3.10.5 Writing for non-scholarly audiences as a source of prestige

The majority of respondents saw value in reaching out to non-disciplinary audiences. This issue was addressed in the audience section, here we are just putting it into the context of prestige. Moreover, many researchers identify writing for non-scholars as a source of prestige. The skill to communicate with larger audiences can be understood as an ability to explain things more simply: "[...] the majority of us, the common scientist, should be able to explain what they do in plain language to others, and those who can do this clearly are significant contributing the science and should get prestige" (OP30).

One of the interviewees drew an interesting comparison from the perspective of his discipline:

> If you publish an article in Le Monde, as a historian, it is a huge source of prestige. There was a professor saying: “to publish for the general public media, for a historian, is the equivalent to publishing in Science or Nature for a scientist.” (OP22)

However, while most do indeed see reaching broader audiences as something possible, not all find it extremely important. It may be that the strictly academic work addressed to experts comes first.

> If writing a popular science book or something like that delays your dissertation by three years and stretches your CV, then it's a problem. If you can do that on the side, i.e., write a dissertation for 3–4 years and then write the bestseller on the easy subject, okay. (OP20)

Moreover, in certain cases popular scientists are contrasted with experts who are recognised within their own disciplinary field and who mainly produce publications addressed to the scholarly community.

> [...] there are people who are relatively well known due to [their] media positions, or forms of intervention, or types of books, even though they are quite marginal in the scientific space and are not considered to be leading researchers. (OP25)

3.10.6 Summary

- Many scholars admit that there is a strict hierarchy of publications, with the monograph often being mentioned as the most prestigious output in SSH.
- Factors that influence the trustworthiness of a publication in the eyes of fellow scholars include: relevance (including recommendations or the fact that the publication has been cited elsewhere), bibliometrics, authorship, the publisher or the journal, and peer review.
- Innovative forms of writing do not have an established position in academia yet. Some respondents already expected novel solutions from their colleagues, and have referred to digital outputs (such as blogs or tweets) in their own work, whereas others saw them as undervalued and difficult to cite.
- There is no consensus on the relationship between open access and prestige.
- Although most respondents agreed that writing for non-scholarly audiences was a positive phenomenon, they attached different levels of importance to it.
- Prestige in academia is something more than strict bibliometrics, for example, relying on the idea of scarcity. Still, prestige may be difficult

to define and some perceive it as too narrow a concept, one that should be broadened to incorporate innovative forms of writing.

4 Case studies: the state of innovation

4.1 Methodology

In this section we explore the current landscape of scholarly writing, focusing on the actual directions of the innovations in scholarly communication. We were interested in how the providers of tools and services define and respond to the needs of researchers and which innovations are considered useful.

Altogether 56 cases were identified for analysis through snowball sampling for which we used various sources, such as literature review, and suggestions from interviews or presentations during events on innovations (e.g. Open Book Fest). We aimed for sample diversity rather than representativeness, hence we did not include some important projects if we felt that their specificity was satisfactorily addressed in another case. The study, therefore, was not aimed at creating an exhaustive typology of innovative cases but rather to provide a rich analysis of the various features observed in those innovations.

On the basis of a literature review, a detailed study protocol was prepared to serve as a guide for researchers. It was based on the DiMPO protocol for meta-research in digital humanities proposed by Maryl et al (2020b). It was attuned to the particular issues identified in the earlier stages of the project and consisted of the following categories:

- **Basic data** such as project title, type, authors, links, status, date of creation, languages, and licensing.
- **Abstract** – a brief description of the project.
- **Users and their needs** – reasons for the project's development, response to certain needs, user research, projected role of users, etc.
- **Data and technology** – description of content, data formats, programming languages, features, links to databases, etc.
- **Affiliation, authorship, workflow** – how is authorship handled, what is the team's workflow.
- **Availability and accessibility** – entry requirements to use the service, conditions of access, compatibility with other services or different browsers.
- **Sustainability** – business model, sources of funding, security measures, and persistent identifiers.
- **Evaluation, trust, and authority** – basis of source credibility, peer-review and other kinds of evaluation or metrics.
- **Impact** – examples of use, number of users, societal and educational impact, communication strategies.
- **Bibliography** – creating a portfolio of documents about the given case.

Each section came with detailed questions to help define innovative areas in each case. All notes were gathered together collaboratively in a Google spreadsheet. Case studies were listed in a separate file and simultaneously saved to Zotero.

Case studies were analysed by a team consisting of two main researchers plus collaborators (including trainees), who had a chance to contribute to the study while gaining analytical skills in the study of innovation.

All citations from home pages are marked with a case code in brackets, e.g. (C1). Other citations are referenced in the footnotes. The list of codes with links is available in Annex 3.

4.2 Overview of innovative cases

We looked at the future of scholarly writing through all the stages of the research cycle, from the first idea to all post-publication activities like promotional strategies or the reuse of data for subsequent projects.

According to DiMPO typology (Dallas et. al 2017), scholarly activities can be categorized into five non-linear phases:

- To discover, collect, and create research assets.
- To organize, structure, and manage research as sets.
- To annotate, enrich, and curate research assets.
- To process, analyse, and visualise one's own research assets.
- To publish, disseminate, and communicate about research.

In fact, the discovery stage is ongoing throughout all the phases. Yet, the cases mostly connected to this stage are the digital cultural heritage projects like Polona, "one of the most modern digital libraries in the world and, at the same time, the largest library of its kind in Poland" (C13) filled not only with modern monographs and scholarly articles, but also with "illuminated manuscripts, the oldest Polish printed books, engravings, drawings and popular publications – postcards, old primers, children's books, cookbooks, and old handbooks" (C13). SSH Open Marketplace is also completely dedicated to this phase: "a discovery portal which pools and contextualises resources for social sciences and humanities research communities: tools, services, training materials, datasets, and workflows. The Marketplace highlights and showcases solutions and research practices for every step of the SSH research data life cycle."[6] Also high in discoverability are projects that group pre-texts as data, such as Octopus, the primary research record, publishing "all kinds of scientific work, whether it is a hypothesis, a method, data, an analysis, or a peer review;" (C32) or Protocols.io, "a secure platform for developing and sharing reproducible methods." (C29)

As for the stages concerning organizing, annotating, and processing, we found cases that wanted to deliver all these features in a single tool, for example, Rebus Ink, Authorea, Jupyter, and Recogito. On the other hand, there is also a tendency towards narrow specialisation, such as with Pundit for annotating websites, or ContentMine, which opens source text mining services. We looked at different text processors (for example, FairCopy Editor, Overleaf, Fidus Writer, dokieli), content management systems (like Muruca, Omeka, Mukurtu), and other types of software (docloop, Editoria, Hypotheses, Publons).

[6] https://marketplace.sshopencloud.eu/about

When it comes to the publishing phase, we investigated examples of publishing houses (Language Science Press, OLH) and platforms (Manifold, Janeway, NextBook, Electronic Book Works, New Panorama of Polish Literature). We also analysed specific outputs (in the form of digital collections, digital editions, and databases etc.), like *Metagaming: Playing, Competing, Spectating, Cheating, Trading, Making and Breaking Videogames*, and *The Chinese Deathscape: Grave Reform in Modern China*. There is also a representation of cases focused on post-publication activities, for example, Peerage of Science, “[a] free service for scientific peer review and publishing"(C36) or PubPeer, an online platform for post-publication peer review.

We observed the progress of the inclusiveness trend in designing with regard to scholars from the humanities and social sciences. Mukuru is a platform dedicated to creating digital heritage collections, whereas Muruca focuses on digital scholarly editions. On the one hand, there are projects of a more general scope providing for diverse data types, methods, and disciplines (SSH Open Marketplace), while on the other hand, they may address relatively narrow areas like research activities around manuscripts (FromThePage), or a specific discipline like Polish literature (New Panorama of Polish Literature).

The community aspect is very important at all writing stages. For instance, Manifold allows chapter drafts to be made visible to the community. To some extent participating readers become co-authors or collaborators, providing new content (as designed in SSH Open Marketplace), or reviewing source code (Jupyter) and incoming functionalities (Rebus Ink, FairCopy Editor).

Openness is greatly appreciated, and is applied not only to final results, but also to the whole writing process including at the technological level, which manifests itself in Github-type repositories. This might evolve into a more radical form, as Janeway’s creators state: “[We] will never accept commits of, or ourselves write, paywall features into Janeway.” (C45) When it comes to for-profit cases and pricing, individual researchers still often have the opportunity to use products free of charge, although with basic functionalities only.

Impact and prestige measures are not homogeneous. Apart from the well-established forms like citations or metrics (based on statistics), the number of workshops and amount of training given by creators is a sign of the popularity of a case. It also seems that social media engagement isn't used to its full potential, perhaps due to the need for regular payments. Moreover, developing and following a sophisticated social media strategy may not be an element of the typical scholarly skill set.

In the next sections we will analyse such tendencies more thoroughly.

4.3 Users and their needs

In this section we discuss how tool and service developers identify and define the gaps in scholarly communication they wish to address with their products. We focus on users from two directions. First of all, we look at the reasons behind the development of a particular project and its current aims, i.e., what gap is this project trying to bridge.

Second, we attempt to reconstruct projected user roles and their communication with the creators.

4.3.1 Needs

4.3.1.1 Reasons and gaps

The reasons for developing a particular project are most often related to the identification of a specific problem in the sphere of scholarly communication. These reasons are quite diverse and variable depending on the circumstances, for example, at the time of its creation in 2014 **Authorea**, a web service for collaborative writing, did not have any typical scholarly competition. Github, MS Word, Google Docs, and Latex word processors had similar functionalities (Lomas 2014), but were not necessarily focused on the academic community. Except for some Latex editors, at the time, Authorea was an innovation. As the authors explained their motivation for starting this project: "We were frustrated that other writing tools didn't fully understand the needs of researchers – especially researchers in a web-first world – and we wondered why the Internet age hadn't yet delivered a modern toolset for scientific collaboration."[7] This need to design a writing tool or service dedicated to researchers deepened and was addressed by the development of projects like Fidus Writer, FairCopy Editor, Rebus Ink, etc.

Rebus Ink, a web-based research workflow application, currently in the beta version, has also addressed the need to overcome the obstacle of incompatibility between different writing tools applied during the typical writing process, often resulting in makeshift, ill-fitting workflows (Hyde 2020). Their aim is to design a smooth workflow for writing that will be clear and intuitive for users and that will mitigate the incompatibility issues between the various applications by keeping the whole process in one place. Hence, it will provide a more "streamlined" path through the whole writing process: "Users should always understand where they are, not just within the app, but where they are within their process of workflow. We must provide clear navigation, structured architecture and clear indications of things like recently touched sources, recently edited notes etc." (Hyde 2020). **Janeway**, a journal management software, has been designed to streamline the journal publishing process by bringing the entire workflow "from submission and review, to editing, production, and final publication"[8] together in one place by ensuring that communication between researchers and editors is maintained throughout. Communication with users seems to be key to the process of identifying particular gaps.

Often projects are created to overcome a apparent obstacle in scholarly communication that, according to the project's creators, has been misidentified. This is borne out by the example of **Peerage of Science**, a service for scientific peer review and publishing, which counters the belief that publishing early drafts of academic works or sharing the ideas freely online increases the risk of intellectual theft. On the contrary,

[7] https://www.authorea.com/aboutus

[8] https://janeway.systems/

as the creators justify, if editors and scholars from your field see your preliminary research, it will later be recognised as yours. **Peerage of Science** is also aware of the problem that reviewing work in the scientific community lacks recognition: "[The] traditional peer review system does not give Reviewers citable academic recognition or other compensation for their reviewing work."[9] Dedicated to STEM sciences, the microPublication.org service likewise publishes single research observations and findings, and positive and negative results, each of which is peer-reviewed and assigned a DOI. Some projects build on the desire to embrace the social media affordances of digital technologies (discussed later in the subsection on promotional strategies).

For **Periodicals**, a lightweight virtual journal that has oe of its users as editor-in-chief, it is important to acknowledge that scholars are not satisfied with the current state of scholarly journals: "The traditional journal has changed remarkably little in centuries and many people feel that scientific publishing is stuck in a rut, subject to a corporatist drift, and is not serving science optimally," (C38) thus, researchers need freedom in setting editorial policy to select the most interesting and useful manuscripts. This approach can also be linked to the need for a decentralised model of communication, already advocated by **dokieli** on their mainpage: "a clientside editor for decentralised article publishing, annotations, and social interactions," (C43) which also asserts on their main site that "no central servers to monitor or control your content or the interactions of your reader[s]." (C43) For **PubPeer**, a platform allowing post-publication peer review, dissatisfaction with the current publishing system is also due to the many mistakes that appear in texts despite the reviewing and editing process. A solution, they claim, would be to allow for open social commentary that could enable a process of constant improvement: "Often – due to the sheer numbers of people involved – this is the only way that flaws and inconsistencies that have gone unnoticed in the original peer review process can be revealed" (Stoye 2015).

At the heart of many projects is the need to propose an open, sustainable alternative to the commercial academic publishing business models. A recurring problem is the prohibitive prices for access to monographs and scientific journals, which was the catalyst for **Language Science Press** to establish their own publishing system: "Some publishers raised book prices by more than 100% in the past ten years while consumer prices only raised by 22%."[10] Also, big publishers' marketing capacities, which allow outputs to be promoted through various channels, is often a substantial factor in a publication's success, which may be of great disadvantage for authors willing to cooperate with smaller presses. Language Science Press addresses this situation: "The solution to the problem is the publication on a central storage and archiving server in combination with print on demand services. The copyright is granted by the Creative Commons CC BY, which allows the work to be printed and figures to be reused, provided the original work is cited."[11]

[9] https://www.peerageofscience.org/solutions/

[10] https://langsci-press.org/motivation

[11] https://langsci-press.org/motivation

The **Open Library of Humanities** (OLH) recognises the limitations of the APC (Article Processing Charges) model for journals in the humanities. Hence, they have proposed their own business model based on sponsorship of their publishing platform by institutions and libraries around the world: "We are funded instead, through a model of Library Partnership Subsidies to collectively fund the venue and its array of journals. A large number of libraries and institutions worldwide already support us, which makes for a sustainable, safe platform."[12]

Omeka, an open-source web publishing platform for sharing digital collections and creating media-rich online exhibits, can be seen as a fine solution for GLAM (galleries, libraries, archives, and museums) and scientific institutions interested in publishing their content online with basic standards. Omeka matches the basic needs of small projects, even without any funding, thanks to a collection of plugins and themes. Furthermore, Omeka can be independently developed within large digitisation projects or can become a basic web infrastructure for an institution's digital resources.

"Efficiently and inexpensively" (C33) is also **Editoria's** strategy for publishing. It is a solution for researchers who perceive traditional ways of publishing as costly and out-dated. Editoria's team notes that "a lack of funds for scholarly communication has meant a curtailing of many forms of publishing" (Rühling 2018).

It seems that the reported needs are nevertheless linked to a fairly specific need, or, more often, problems, and the impetus for their creation, often lies in a "lack" or "deficiency" of something. This is often connected with the urge to explore and experiment with new solutions and opportunities beyond the gap analysis. To give a counterexample, in 2006 the digital version of the book ***Gamer Theory*** was created to "explore the possibility of a new textual form in social web media: a middle space, somewhere between the sprawling public discourse arena of the blogosphere and the collaborative knowledge factory of Wikipedia."[13] Sometimes a special issue of a particular journal or series might be devoted to seeking new means of expression, as in ***The Disrupted Journal of Media Practice***, whose reason for experimenting with the form of discussing media practice was the need to embrace a plurality of media in light of the fact that "scholarly forms of production and communication remain predominantly text-based."[14]

Cases that have not responded to this tendency thoroughly, have been those connected with responding to and overcoming the COVID-19 pandemic and enforced remote working, international travel bans, and similar. The proliferation of webinars, online seminars, and digital workshops is huge, but a particularly interesting case was a Twitter Conference under the title "DH in the Time of Virus," organised by the "Athena" Research and Innovation Center.[15] The "transcript" of the presentations and discussions at this conference, and one of the first innovative examples for dealing with lockdown situations in SSH, can be explored using the hashtag #DHgoesViral on

[12] https://www.openlibhums.org/site/about/the-olh-model/

[13] futureofthebook.org/gamertheory2.0/index.html@page_id=2.html

[14] http://journal.disruptivemedia.org.uk/intro/

[15] https://apollonis-infrastructure.gr/2020/03/12/dh-in-the-time-of-virus-twitter-conference-02-04-2020/

Twitter. The main advantage of this form is its concise and serialised way of presenting, which is imposed by the character limit of a single tweet. Moreover, the presentation appeals to broader audiences in the social media environment, who can freely engage with selected arguments.

4.3.1.2 Lowering the barriers to uptake

Language Science Press has identified an interesting paradox: "While you have full control over your product, the disadvantage is that you have full control of your product."[16] Users should be aware that by carry out all publishing functions without the help of a traditional publishing house, they then have numerous activities to oversee and perform: "quality control; content; proofreading; typesetting; marketing" (C23). Thus, with great control comes great responsibility, and a need for training.

When a user is interested in new, innovative tools or services there is often a need for training. Tutorials and training opportunities might be considered a form of marketing strategy, especially in the case of *for profit* services and startups. For instance, **Mukurtu,** a digital cultural heritage sharing platform, prepared a dedicated support site[17] for new users. It is also possible to organise training for teams through web conference, phone call, or email. **Pundit** has included an exemplary video tutorial on their website,[18] which showcases this tool's most important features for beginners. **E-Editiones** provides a YouTube workshop[19] for beginners and even a Slack channel for solving problems. Similarly, **FairCopyEditor,** apart from video tutorials,[20] runs a support forum[21] for the community. Despite the plethora of avenues available to engage with the product, users don't seem to be very active, for example, on support forums. It seems that tailored training should stem from user research carried out before and during the development of the tool or service in order to target the problems encountered by actual users.

4.3.2 **Users**

These services engage with their users on many levels in order to ensure that services address the actual gap and that the training provided fits scholarly needs. In this section we look at how the platforms define their users and provide space for interaction.

4.3.2.1 Testers

Identifying needs and knowing your target users can be achieved through surveys.[22] However, a method that is gaining popularity is to invite testers (people outside the circle and network of developers) to research and open-test a tool or service, which is usually done at the development stage.

---

[16] https://langsci-press.org/motivation

[17] https://mukurtu.org/support/

[18] https://thepund.it/videos-pundit-web-annotation/

[19] https://www.youtube.com/channel/UCAPhSZdBwFRCEFWNNYOC4Ww

[20] https://vimeo.com/showcase/faircopy

[21] https://support.faircopyeditor.com/

[22] As such, Rebus Ink's creators were asking users about their accessibility needs in the survey.

Rebus Ink states on their website that they are looking for beta testers for their service. The testers' experiences are posted on a special forum.[23] For more advanced users, **Fidus Writer** provided an option to test new or in-progress features for everyone by logging into a temporary server[24].

The SSH Open Marketplace, a platform developed for the European Open Science Cloud, makes an interesting case concerning elaborated testing procedures: four external testers had a week to test the alpha version across three dimensions: 1) UX design, navigation, and search experience; 2) content and curation of content; and 3) curation, trust, and governance. After a week, testers shared their results during a dedicated workshop with potential users of the service.[25]

The procedures for becoming a tester are clear and fast in most cases; sometimes only email signup is required, as with the early access program in **FairCopy Editor**.[26]

Interestingly, **Janeway** declares that "testing will be applied to security modules and whenever a post-launch bug fix is committed."[27] The nature of this process is "selective" in contrast to the above examples.

The testing phase serves as pre-evaluation by the users, and the results should be shared openly with the community.

4.3.2.2 Types of users

The services analysed here are focused on individual and institutional users. For instance, **Kudos** delivers to scholars who "[...] want assistance with increasing usage of, and citation[s] to, their publications. Kudos is also for institutions and funders looking to increase the impact of the research that they fund, and for publishers wanting to develop closer relationships with their author communities and increase publication performance."[28]

Science Open users can be divided into three main groups: publishers, institutions, and researchers. **ScholarLed** targets, firstly, small publishers, who are looking "for collaboration rather than competition." (C21) They chose that kind of user in order to help small-scale scholarly OA presses. In some cases, for example, **ContentMine**, it seems that the users are only institutions: "ContentMine provides text and data mining services on a consultancy basis to academic institutions, organisations, companies, and not for profit institutions." (C55)

Yet, it is still the individual user who receives more attention. When it comes to individuals, projects identify two broad categories of users: the tech-savvy digital humanist that is often experienced in coding, and the digitally inexperienced researcher who wants to use a particular tool or software without obtaining programming skills for the long run. The distinction between the two is rather fluid in

[23] https://support.rebus.ink/ and https://support.rebus.ink/c/beta-info/

[24] https://staging.fiduswriter.org/ (established in 2018 to test track changes features)

[25] *Agile Development of the SSH Open Marketplace: Alignment with User Requirements* workshop: https://www.youtube.com/watch?v=AZTrzHtil3Q&ab_channel=Science2.0Alliance.

[26] https://www.faircopyeditor.com/en/

[27] https://janeway.systems/about

[28] https://www.growkudos.com/about/user_guide

practice. Authors and developers are aware of the different needs corresponding to these two categories, and often try to meet the expectations of both.

The authors of ***Black Quotidian: Everyday History in African-American Newspapers*** (an instance of the Scalar 2 platform run by Stanford University Press) mention three types of visitors: "(1) skimmers, who browse a handful of posts or sources; (2) swimmers, who delve more deeply into particular essays and paths; and (3) divers, who spend significant time exploring the site and engaging with the project's methodological questions."[29] **Manifold**, a platform for scholarly publishing, has a clear statement on its website: "Different publishers have different needs. If you've got the resources and technical skills, you can install and maintain Manifold yourself. If you need assistance, Manifold Publishing Services can help you install and maintain your web publishing workflow." (C2) Similarly, the digital library **Polona** provides an API for more advanced users to "facilitate access to the resources shared in the digital library" (Rosa 2019, 27). For **New Panorama of Polish Literature**, a platform for digital scholarly collections about Polish literature, users are found to be diverse, depending on the specific collection. For example, the *Prus Plus* collection has a popularizing character and is intended for students and enthusiasts of the novel *Lalka* (*The Doll*) by Bolesław Prus. In turn, *Sienkiewicz Postmodern* is a collection of scientific articles with non-linear and video game-inspired control and navigation. Thus, it may interest both literature experts and non-specialists.

FairCopy Editor's authors declare that their tool is designed for non-tech-savvy researchers (mostly from the humanities): "Transcribing isn't just for XML experts anymore." (C54) Also **Muruca**, a CMS for digital scholarly editions, states: "To work with Muruca, no technical skills are needed." (C31) The Muruca case is specific, because their target audience is quite narrow in comparison to other cases. Their users are digital scholarly editors (therefore SSH researchers), yet cultural heritage institutions might also be interested in this tool.

Likewise **FromThePage** is a tool dedicated specifically to editing manuscripts. It is, therefore, designed for scholars working with manuscripts, and primarily meets their needs with features like the ability to compare manuscripts with transcriptions, even different versions of each. The **Next Book** project ("An open platform for publishing and reading on the web. Defining the standard for book reading in Q2 of the 21st-century and beyond" (Kocurek et al. 2016)) states that it is designed for academics, researchers, students and seekers of knowledge, and the self-educated.[30] This last type seems most interesting as it implies a public science edge. Similarly, the *Ports, Past and Present: Cultural Crossings between Ireland and Wales* project was created with and for users outside academic circles (local communities are encouraged to share their stories about their surroundings; the Wexford County Council is one of the partners involved in the initiative). Their target audience may be tourists – and the tourist sector is named as a big stakeholder – as well as arts organisations, writers, and artists. For the post-publication peer review platform **PubPeer**, the audience is

[29] https://blackquotidian.supdigital.org/bq/themes

[30] https://next-book.eu/en/book/01-about.html#idea2

researchers in a specific phase of their scholarly communication process: those who want to share and comment on papers already published.

Tools might also be designed for scholars wishing to follow a very specific workflow. For instance, **Docloop** provides an ideal user's case, namely, for "the author working on an open educational resource (OER) using LaTeX and publishing the source code in an online repository." (C56)

Finally, when we talk about users' needs we need to think in plural terms. Across most of the analysed cases, the future of scholarly writing seems to apply to a collaborative approach. Even if a project was originally designed for an individual, it still provides a connection to other users. A good example is **Publons,** a tool for tracking publications, citation metrics, peer reviews, and journal editing work for a single researcher, which offers an option to invite other researchers to your private dashboard.[31] Collaboration features were also declared by other projects, for example, **Scalar** and **Authorea,** and will be addressed more thoroughly in the Features section.

4.3.2.3 Co-creation, community

The future of collaborative writing is not limited to working on one document, but encompasses communicating with each other on different levels and platforms. The ultimate aim for a significant number of the analysed cases is to turn users into the co-creators of a particular tool, service, or project. The team behind **PubPub**, a tool enabling group work on publications, believes that a community-driven approach has a positive impact on research in terms of inclusion and higher quality: "When researchers collaborate the result is often more impactful research. And, as the preprint model has demonstrated, publishing work early can lead to more opportunities for feedback and higher publication and citation rates. But there are larger benefits to choosing a community publishing model as well. PubPub users [...] have shown that a community-driven approach can also invite people likely to be affected by research into the knowledge creation process. As a result, research becomes more transparent, more inclusive, and ultimately, more trusted and impactful."[32] Researchers are vital for PubPub's team, not only as users of the tool, but also as "co-producers" who are very keen to comment and provide feedback.

One of the most common ways to participate in a digital project is through Github-type repositories, yet this is an option for rather tech-savvy users. Such repositories are open and available to many of the analysed cases.[33] For instance, **Gitenberg** puts a lot of emphasis on collaboration. It is "a prototype that explores how Project Gutenberg might work if all the Gutenberg texts were on Github, so that tools like version control, continuous integration, and pull-request workflow could be employed" (Hellman 2018). Basically, the very choice to share code on Github is an invitation to collaborate, to correct, and to polish Gitenberg; so it is a very scholarly way of thinking

31 https://publons.freshdesk.com/support/solutions/articles/12000075331-granting-dashboard-access-to-others

32 https://www.pubpub.org/about

33 Manifold, docloop, Authorea, TOPOI, Faustedition, SchoalrLed,PubPub, docloop, ContentMine, SSH Open Marketplace, Protocols.io Octopus, Editoria, Recogito, Janeway, GITenberg..

about a project, especially because it is an open source method, and is developed and internalised by thousands of people, so it practises a radical openness (Himmelstein 2020). On the other hand, there are still a lot of people who aren't used to working on Github, and they don't want to wait until a document is merged before they can work on it. In addition, the publishing industry is unwilling to deal with versions and forks. Getting things out of Github into academic circulation poses a significant difficulty (Himmelstein 2020).

Another crucial aspect in the area of co-creation is crowdsourcing. The above mentioned **FromThePage**, a manuscript transcription tool, proposed the "Your First Crowdsourcing Project" workshop on its blog:[34] "The session covers selecting material, finding volunteers, developing transcription conventions, keeping volunteers engaged, and what to do with your transcriptions once you're done." The barriers to crowdsourcing uptake might be connected to the issue of voluntary labour and the quality control of works transcribed by non-professionals. Ben Brumfield, when referring to the idea that "scholarly editors or professional staff at libraries and archives can be replaced by a crowd of volunteers who will do the same work for free," (C46) observed: "Decision-makers seem to understand that crowdsourced tasks are different in nature from most professional work and that crowdsourcing projects cannot succeed without guidance, support, and intervention by staff" (Brumfield 2020). Thus, he highlights that these kinds of works still have professional staff involved – they are trainers and supervisors for volunteers, so it is more about delegating certain activities than replacing one group (volunteers) with another (scholars).

Language Science Press also sees the potential of crowdsourcing, and utilises various types of activities that might be performed by contributors: "The scientific community is directly involved in the publishing process. Next to authoring or reviewing books, community members can also take over roles in proofreading, typesetting, illustrating, market-ing via crowdsourcing" (Nordhoff 2018).

Other kinds of activities that might be delegated to the community through crowdsourcing are: open peer-review (**Publons**), sharing information to increase publication impact (**Kudos**), translations (**Fidus Writer**), and UX design and research, as well as coding and code review (**Jupyter**).

The community's contribution is perceived as the ultimate goal and symbol of project success. For **Editoria**, the users' attention and work is all that constitutes this project: "Here the user is, by design, as important as everyone else, perhaps even more important. Users are the people who know best what changes can improve Editoria" (Rühling 2018).

However, with such openness also comes issues. Firstly, how should this work be attributed and evaluated, and who is/are the author(s) after all? Secondly, how are less technically advanced users enlisted?

4.3.2.4 Communication between users (and authors)

It is worth distinguishing the options available for collaboration and direct communication between the users and/or authors of a particular tool or service.

[34] Blog address: Content.fromthepage.com. Workshop's form: https://bit.ly/2RIUgiN

Communication channels are vital for developing communities; thus, it is remarkable that many cases still have not include this option for their users – which is even more surprising in the times of the pandemic, where there is a vast range of options dedicated to this.

Jupyter has the most diverse list of options available to their users for communicating: a forum, two mailing lists (general and education), a Google Group, a chatroom on Gitter, the Jupyter site on Stack Overflow, and community guides on the Jupyter site about events and a second one on how to contribute.

In other projects, there is a preference for one main communication channel. Apart from a contact mail address, there are options like: Mattermost (**Editoria**), Slack (**E-editiones**), and Discord (**Janeway**). The forum format seems to be scoring a small comeback in such projects as Fidus Writer and Faircopy editor.

4.4 Data and technology

In this section we are interested in the content delivered through the analysed services. First, we look at the research data formats used in these projects, and second, we analyse the features and functionalities that are provided.

4.4.1 Data

Scholarly text in the electronic environment is often supplemented with, and well connected to, underlying data. Many services try to facilitate the connection between text and data.

4.4.1.1 Research data

Data is not only a final product of the research life cycle, but emerges at all stages of the research lifecycle, thus, many projects focus on outputs before the final paper. **Authorea**, a collaborative platform for publication, sees the range of data inclusiveness in the academic writing process as follows: "We are building a platform where researchers can collaborate and write their findings including not only text and figures, but also all the important 'products' that are currently lost upon publication: notebooks, data, analysis and code" (Cantiello 2016).

Research data is often organised in datasets for future reuse. Interestingly, these datasets might contain additional data that are not accessible in another form, as in the digital book ***The Chinese Deathscape. Grave Reform in Modern China***: "the original datasets contain surplus data, in the sense of data points that, while rich and potentially insightful, are not exploited by the online platform."[35] On the upcoming SSH Open Marketplace platform, *Datasets* are promoted to a main category, alongside *Tools and Services*, *Publications*, *Training Materials*, and *Workflows*. Yet, at the moment, there is a great disproportion between these categories: *Tools and Services* has 1606 entities, *Publications* 2986, *Training Materials* 140, *Datasets* 2, and *Workflows* 29.[36]

[35] https://chinesedeathscape.supdigital.org/data

[36] As of 21st of January 2021, https://marketplace.sshopencloud.eu. It is necessary to add, that this platform is still in the beta version.

Datasets are also an output from the evaluation campaign, **PolEval**. These annotated datasets can be used for the teaching and evaluation of systems for Natural Language Processing. Additionally, an open repository of resources and methods is planned (Kobyliński 2020).

Data is also seen as providing richer contexts for scholarly texts: visuals, graphs, audio etc. For instance, **New Panorama of Polish Literature** contains digital collections about 19th-century and contemporary Polish literature, whereby the text is supplemented with maps and images. The "Disrupted Journal of Media Practice" identifies articles, and the "series of curated conversations" around them, as their data.[37] Similarly, digital scholarly editions are also enhanced with supplemental materials. For instance, ***Faust. Historisch-kritische Edition***, was supplemented with "[...] relevant prints published during Goethe's lifetime and over 1500 testimonies to the creation of the work" (Sonntag 2018).

Octopus is a platform for scholars willing to "[p]ublish work that you cannot publish elsewhere: hypotheses, small data sets, methods, peer reviews." (C32) Octopus also has a unique standpoint – to share scholarly failures too: "Good science isn't necessarily a good story. Good science can be the careful collection of a small amount of data, or careful analysis of data collected by someone else, or a good hypothesis (regardless of whether data later supports it or not)."[38] New data might also appear after publication; thus, **PubPeer** collects post-review publications for further discussion. According to PubPeer's main website, the number of recently commented publications is 104,148.[39]

4.4.1.2 Formats and forms

When it comes to data publishing, a move beyond the mere PDF is recommended by **Authorea**[40] and **Protocols.io,**[41] and has been put into practise in the analysed cases. The choice of formats for data publication correspond with the type of project. Here is an outline of forms and formats used in the cases analysed in this study:

- **TEI**: Melville Electronic Library. A critical archive (MEL), Faust. Historisch-kritische Edition (Faustedition), TEI NPLP, FromThe Page;
- **XML/RDF**: Polona, Black Quotidian: Everyday History in African-American Newspapers;
- **LaTeX**: Language Science Press, Refereed;
- **PDF**: Polona, GITenberg, Refereed;
- **JPG**: Polona, Mukurtu;
- **JSON**: Polona, TEI NPLP, Black Quotidian: Everyday History in African-American Newspapers, Jupyter;

[37] http://journal.disruptivemedia.org.uk/intro/

[38] https://demo.science-octopus.org/about

[39] As for 25.01.2021 from https://pubpeer.com/

[40] "Not just PDFs. You can publish d3.js and Plot.ly graphs, data, code, Jupyter notebooks" (https://www.authorea.com/)

[41] "Unlike static PDFs, they (protocols - ASZ) are "runnable" on the web and mobile devices (both iOS and Android)" (https://www.protocols.io.)]

- **EPUB**: GITenberg;
- **TIFF**: Polona;
- **CSV**: Publons, The Chinese Deathscape. Grave Reform in Modern China, From The Page;
- **RIS:** Publons;
- **BibTeX**: Publons;
- **HTML**: Black Quotidian: Everyday History in African-American Newspapers, From The Page;
- Kindle Ebooks: GITenberg;
- **plain text:** GITenberg, TEI NPLP;
- **MP3**: Mukurtu;
- **d3.js**: Authorea.

The crucial issue here is to choose formats that will be reasonably easy to reuse and interoperate without the necessity of downloading any additional software. A radical example of not meeting interoperability requirements is the case of **after.video**, an edited collection of assembled and annotated video books. By combining the PDF format with the Raspberry Pi operating system (necessary to "read" this format), the video book has a file size of almost 30 GB – similar to a video game – which makes it quite awkward to use. Also, installing Raspberry Pi for only one publication seems an excessive effort.

4.4.2 Technology

4.4.2.1 Languages and formats

Given the increasing popularity of browser-based projects, web languages such as HTML, CSS, JavaScript would be the natural choice. For instance **Manifold** uses "JavaScript (48.8%), Ruby (38.3%), SCSS (7.6%), TeX (3.4%), CSS (0.9%), HTML (0.8%), Other (0.2%)."[42] It appears that information about languages and formats is divided with regard to core or backend languages, and frontend languages. At its core, Omeka's technology is PHP, Zend Framework on a LAMP server (Linux / Apache / Maria DB - MySQL / PHP), and uses standard CSS/SCSS and JavaScript (as jQuery) for the interface. Also, themes can be built into any front-end standard. API in the REST standard can be used to build a "headless project," providing content using any modern JavaScript framework without the need to link it to a particular output. **Muruca's** creators present architecture and technology grouped into: "1. Backend: Wordpress, elastic.co + database MySQL, 2. Frontend: Muruca, Publisher 3. A connection-framework between back and front through a Rest API: serverless. The frontend is flexible and highly configurable, and completely disconnected from the backend" (Andreini 2020).

The choice of technology and programming language is a challenge for many SSH researchers wishing to present their research in digital form. Rarely, however, is the rationale for the choice stated clearly. The exceptions here are the developers of **Janeway** journal management software. They decided to use a platform written in

[42] https://github.com/ManifoldScholar/Manifold

Python instead of PHP, the language of Open Journal Systems (OJS), because of the more stable and secure characteristics of the former: “We knew Open Journal Systems (OJS) well at this time, but were not big fans of PHP, the language in which it is written. We also were aware of the work being done by Coko (the Collaborative Knowledge Foundation) in Node.js. What we really craved, though, was a scholarly communications platform written in Python/Django. Python was the most popular programming language in 2017, which would mean that a platform in this language would be comprehensible to a wide range of programmers. So, we chose to write in Python using the Django framework, since this is a well-known, stable, and secure framework for the development of web applications” (Eve and Byers 2018). Similar reasons, concerning stability and sustainability led the team at **New Panorama of Polish Literature**, in 2015, to choose Wordpress as their CMS for digital monographs instead of **Omeka**. At the time, **Omeka** did not provide the functionalities needed for the project; thus, opting for a more stable yet commercial CMS that was well-known and actively developed by its community was a more justifiable approach.

4.4.2.2 Features and functionalities

Of the numerous functionalities, those listed below seem to recur and be the most important for the analysed projects. We discuss each feature using one selected example, while providing others for reference:

Annotation – in **Manifold,** every part of the book can be highlighted, cited, and annotated. Users can choose whether they want to see private, public, or group annotations. Moreover, every highlight can be shared via Twitter or Facebook, or referenced. The citation can be generated in three styles: APA, MLA, and Chicago. Users also have the possibility to use the reading tool as well as additional options for ease of reading, such as changing the width of margins, changing the font and its size, and selecting night mode.[43]

Other cases with a strong focus on annotation are **Pundit, Scalar, The Disrupted Journal of Media Practice, FairCopy Editor, Recogito, docloop, PubPub, and dokieli.**

Collating – **Pulter Project**[44] (digital edition) contains 102 poems. There are two ways of displaying and comparing poems from different types of editions: the “Elemental Edition” (non-academic users) and the “Amplified Edition” (professionals). The latter option displays critical commentary and high-quality scans of the manuscript. Moreover, the platform provides the transcribed version of a particular poem, for both “elemental” and “amplified” editions. All of these might be turned off or on. There is also a search engine but it works only in a basic mode.[45] **Faust. Historisch-kritische Edition** is also an excellent case here. In this digital scholarly edition every line (!) has its own metadata and provenance information, and each version of the text can be

[43] https://manifold.umn.edu/read/metagaming/section/7ec09519-aa44-4f90-aaf6-eb7d50c4c0e7
[44] http://pulterproject.northwestern.edu/about-the-project.html
[45] http://pulterproject.northwestern.edu/poems/vm/the-eclipse/

collated with other versions from different years. Other projects focused on this functionality are **Octopus, FairCopy, and Editor**.

Real-time collaboration – **Overleaf** is very rich in collaboration features: "Every document you create [...] is private by default, with two easy ways to share your work with collaborators: by private invitation or by link sharing. Link sharing allows you to share your projects via secret links. Just turn on link sharing, send the link to your co-authors, and they can review, comment and edit. Or simply turn the link off to make your project private again." (C17) It also provides an automatic real-time preview: "Overleaf compiles your project in the background, so you can see the output PDF right away."[46] **Overleaf** provides many features for communicating with collaborators: "With real-time commenting and integrated chat, you can discuss your work without having to switch to email, printed versions, or any other tool. You can leave comments, give quick feedback and resolve issues, all within Overleaf."[47] Other projects focused on this functionality are **PubPub, Editoria, Peerage of Science, and From The Page.**

Collecting data and organizing notes – **Rebus Ink** is an open-source tool with an open API that has three components: Sources, Notes, and Notebooks. It serves as a source library, reading and note-taking interface, and note library. It has researcher-designed features for tagging materials and annotating videos (Hyde 2020). **Recogito** "provides a personal workspace where you can upload, collect and organise your source materials – texts, images and tabular data – and collaborate in their annotation and interpretation. Recogito helps you make your work more visible on the Web more easily, and to exhibit the results of your research as Open Data."[48] It allows researchers to generate semantic data without the need to use formal languages directly, "while at the same time allowing the user to export the produced annotations in different formats such as TEI-XML, RDF, and GeoJson" (Castro 2019). Exemplary cases focused on similar functionalities are **Jupyter** and **Omeka**.

1.1.1.1. Compatibility and Integrity

Compatibility and integrity issues are separated from the other functions, as it seems that this option should be considered at several levels:

Internal integration – pooling resources together within a project. In **Octopus**, publications must be interlinked to form ordered chains: "Octopus accepts 8 types of publication – all must be linked to another publication somewhere in Octopus. The top of any "chain" is a publication that [defines] a scientific Problem. Below that you can publish a **Hypothesis** (theoretical rationale); below that, a Method/Protocol; below that, Data/Results; below that, Analysis; below that, Discussion; and below that Applications or translations in the real world. Reviews can be published attached to any of those 7 other types of publication"[49]

Infrastructure integration – integrating resources that come from different infrastructures. For instance, "**Polona** digital library exhibits collections from over 40

[46] https://www.overleaf.com/for/authors
[47] https://www.overleaf.com/for/authors
[48] https://recogito.pelagios.org/help/tutorial
[49] https://demo.science-octopus.org/about

institutions." OpenEdition, a comprehensive digital publishing infrastructure for the humanities and social sciences, provides services that pool together various publications: research blogs (**Hypotheses**), journals (OpenEdition Journals), and books (OpenEdition Books).[50]

Compatibility – compatible with external tools and projects, especially with those designed for scholars like Zotero (Rebus Ink, Overleaf, Publons, Hypothesis) or ORCiD. For instance, **Publons' integration with external services allows for automatic import of** "your bibliographic reference manager (e.g. EndNote, Zotero or Mendeley)," while providing "citation metrics based on the editorially curated Web of Science Core Collection" and "[y]our verified peer review and journal editing history, powered by partnerships with thousands of scholarly journals."[51] Other analysed services that also provide compatibility with non-specialist platforms are Dropbox (**Omeka**, **Jupyter**) and YouTube (**Scalar**).

4.5 Teams and their workflows

This section focuses on how affiliation and authorship are handled within the project. It also discusses the workflows behind the projects (team structure, leadership, responsibilities, and roles).

Teams working on digital projects need to consist of members who perform various roles. For instance, **Editoria** is developed by a "team of publishers, production editors, engineers, developers, and UX designers."[52] Descriptions of the roles of individual team members are becoming increasingly accurate and meticulous. In the digital edition of Faust, the team structure is presented with scrutiny and clarity: The project involved three "managers/main authors," seven collaborators, two proofreaders, eleven students and eight interns.[53]

The authorship of outputs like digital collections may differ across particular projects depending on the type of work involved. For example, each collection of the **New Panorama of Polish Literature** is authored by a scholar responsible for the content, two programmers, a graphic designer, and digital humanists. The team creates the publication's structure, functionality, and visual material (Szleszyński 2019). Most importantly, there is always a team behind any particular collection rather than an individual author, as is the case for the **Mukurtu** platform.

In the case of **Manifold**, apart from author(s) and editor(s), there is a third player – the community that forms around a particular book. Every project (a book) has its own community around it. Once a new part of the project is published, the community comments on it. The next step is to review the feedback and apply changes. There is strict collaboration between the authors, the editors, and the community in preparing the final version of the book.

[50] https://hypotheses.org/about-hypotheses

[51] https://publons.freshdesk.com/support/solutions/articles/12000009179-how-does-publons-work

[52] https://editoria.pub/about-us/

[53] http://www.faustedition.net/credits

Having a large network of collaborators might be profitable for the team. **Muruca** has a team of eight people including SSH scholars and developers (and is a part of Net7, IT company). The **Muruca** team collaborates with research teams, cultural heritage institutions, and other stakeholders (Andreini 2020). **Electronic Book Works** has a similar approach: "Our team of publishing experts – from design, editorial and software backgrounds – work with our clients to invent, plan and deliver complex publishing projects, empower their people, and improve sustainability."[54]

Yet, despite the dominance of the team structure, it is possible to find instances of contact with a single author, such as in **Octopus** or **docloop**.

Organising a workflow for digital publishers involves coordinating the work of diverse actors. **Electronic Book Works** focuses on simultaneous work between multiple editors and designers on their multi-format books: "This lets us have multiple editors and designers working on the book simultaneously, editing version-controlled, single-master content files, and seeing their changes instantly in both print and screen versions." (C44) The workflow for the Scalar-based digital book ***The Chinese Deathscape. Grave Reform in Modern China***, seems traditional, but one thing stands out: choosing reviewers with similar expertise who have experience with digital projects.

When describing team workflows, **Editoria** practices booksprint as a new workflow method: "many authors work at the same time on various chapters, each having the same capability to write and revise during the same creative session" (Rühling 2018). Collaborative workflow and certain standards are especially important when it comes to digital scholarly editions. Critical editions are obliged to follow certain editorial principles, defined for each case, which need to be integrated into the team's workflow to govern the process of data gathering, annotating, and publishing (e.g. **Pulter Project)**.

When the workflow seems too complicated for authors, it is worth considering presenting it more graphically, as is the case in Janeway, a journal management software developed by the Birkbeck Centre for Technology and Publishing for the Open Library of Humanities (OLH) (Eve and Byers 2018).

4.6 Availability and Accessibility

In this section we focus on the entry requirements for using the analysed services. We look at whether users need to sign-up, provide affiliation, or pay. We also analyse whether tools are accessible via an existing platform or whether they need to be installed and operated by the user's organisation. We also checked their compatibility with various browsers.

4.6.1 Entry requirements

No login is required to access the content of most of the analysed projects. Even if readers can create an account, the content might still be open, however, registration will unlock additional features. For instance, creating a **Manifold** account allows for greater interaction with text through highlighting and commenting. **Polona's** collections

[54] https://electricbookworks.com/about/

may be viewed and downloaded without registration, but users who want to customise them for their own purposes need to register. Very rarely, a special registration key is required to start using a tool or service, as for ***Black Quotidian: Everyday History in African-American Newspapers***, which is built on Scalar. Conveniently, one can use an ORCID account to log in to some services, for example, **Octopus**.[55] Academic affiliation is required in some cases, like for **POL-on**, the broadest repository of data in Polish science and higher education. For some services, some form of external validation is required. For instance, in order to become a "Peer" at **Peerage of Science**, one needs to "have published a peer reviewed scientific article in an established international journal" (C36) and after the registration the potential peer status will be verified.

When a project is still in development, users are sometimes given the opportunity to apply for early access. **FairCopy Editor**, **Rebus Ink**, and **SSH Open Marketplace** have such options.

The web-based projects analysed in this study usually work on most browsers without any problems. Web access tends to be treated as the main access point. For instance, **NextBook** "assumes only the use of a modern web browser and tries to consider the book experience first, technology second, and everything else only after that."[56] It should be noted that some projects provide dedicated software that has to be installed on the readers' hard drives, as in the case of **Mukurtu.**

In most cases, no dedicated mobile application is provided. Still, thanks to a responsive web design approach, analysed projects were available on mobile devices, although sometimes, due to their complexity (like digital editions using versioning) this is less convenient than on desktop devices. Interestingly, **Mukurtu** has created a mobile version of their platform, allowing for online or offline mobile content creation, storage, and access.[57]

4.6.2 **Free of charge, or payment options**

The majority of cases declare a free of charge approach to their content, for instance: **Scalar**, **Publons**, **Refereed**, **Mukurtu**, **Editoria**, **Peerage of Science**, **Janeway**, **Gitenberg**, and **New Panorama of the Polish Literature**. Janeway radically dismisses any paid-access options for their users: "We will never accept commits of, or ourselves write, paywall features into Janeway."[58] Free plans are often limited to individual users (**Kudos**, **Peerage of Sciences**), or to certain basic features, while enabling premium content in the paid plans (**Authorea**). Sometimes creators reward contributions made to the tool's creation and to the community (discussed in earlier sections) by unlocking paid features for contributing users (**Editoria**). See more about payment options in the Sustainability section.

[55] https://demo.science-octopus.org/about
[56] https://www.next-book.info/concept/
[57] https://mukurtu.org/support/what-can-mukurtu-mobile-do-for-me/
[58] https://janeway.systems/about

4.6.3 Approach to open access

In the majority of cases open source and open access approaches are declared: **PubPub**, **Recogito**, **NextBook**, **Gitenberg**, **Editoria**, **Fidus Writer**, **Jupyter**, **ScholarLed**, and **ScienceOpen.** For instance, **Language Science Press** advocates a lean, open science publishing model, whereby all unnecessary costs are eliminated in order to make the publications available worldwide without a fee. Some of the projects also embrace open software workflow like **Electronic Book Works, and Rebus Ink**. **Jupyter** is 100% open-source software, free for all to use, and released under the liberal terms of a modified BSD license. **FromThePage** software is open-source under an AGPL 3.0 license and SSH OpenMarketplace licenses its software under Apache License 2.0.

ContentMine specialises in "providing open-source text mining solutions for both Higher Education and Knowledge-based organisations." (C55) The **PolEval** campaign's outcomes, like new taggers and training corpora, will be mostly available in open access.

Data FAIRness, which is connected to the topic of openness, seems to have a slower uptake. For instance, **Muruca's** future development goals are focused on FAIR principles, open access, and long-term sustainability.

Reproducibility connected with openness is perhaps of greater importance for STEM disciplines. Some platforms provide data on demand for reproducibility purposes. **Authorea**, a collaborative publication platform targeted mostly at health sciences, introduces "data-rich articles," which are meant not only to provide underlying data upon request, but also to contribute to a new reward system "built not only on citation […] but also on data reuse" (Lomas 2014).

4.7 Sustainability

In this section we look at the sustainability and business models of projects in scholarly communication. The sustainability of tools and projects related to the future of academic writing is essential not only for their creators but also for users. For a potential user, information that a tool or service has a stable funding source or a successful business model is a signal that they may not have to look for a replacement in a year or two, or, in the worst case scenario, lose their data. A plethora of tempting initiatives and innovations awaits academic writers, but they may not know if the solution being offered is sustainable. Sustainability is primarily connected to financial cost (whether it is acquired from grants or from an effective business model), but it also refers to issues of, for example, persistent identifiers or project security. A separate issue is the lack of information about sustainability models, even if they exist. However, in a few cases, we can find transparent funding statements or precise budget descriptions (e.g. **Language Science Press**).

4.7.1 Funding sources

If information about the funding source(s) are present (and this is often a requirement of the grant provider), it is often more specific than the kind of general statement given by **Refereed**: "At the moment, Refereed is funded by founders' capital

and governmental business grants." (C28) Usually it provides at least the name of the funder, sometimes also an identifier, period, amount of funding, etc.

The most common source of funding in the analysed cases was the Andrew W. Mellon Foundation, which is dedicated mostly to projects in their initial phase. This institution granted funding to tools (**Rebus Ink**, **Manifold**, **Scalar**, **Recogito**), publishers (**Stanford University Press**, **Open Library of Humanities**), and other types of projects (**Editoria**, **after.video**, which is part of Archive.org). Other funders of the analysed cases include: the European Union's Horizon 2020 programme (HIRMEOS, SSH Open Marketplace), National Endowment for the Humanities (USA) (Scalar, Mukurtu), Gordon and Betty Moore Foundation (Protocols.io), the Chan Zuckerberg Initiative (Protocols.io), the Washington State University Foundation (Mukurtu), the Institute for Museum and Library Services (Mukurtu), the Fetzer Institute and the World Intellectual Property Organization (Mukurtu), Mozilla Science Lab (Octopus), Crosscloud (dokieli.eu), Grant for the Web (Electronic Book Works), INTERREG V-A United Kingdom – Ireland (Ireland – Wales) (Ports, Past and Present), the Shuttleworth Foundation (ContentMine), Open Knowledge Foundation, Deutschland (docloop), Bundesministerium fur Bildung und Forschung (docloop), National Program for the Development of Humanities (New Panorama of Polish Literature), Technology Agency of the Czech Republic (NextBook), Knight Foundation (GITenberg), CORFO (Spanish: Corporación de Fomento de la Producción de Chile) (Fidus Writer), and the German government-funded German Research Foundation, or DFG (German: Deutsche Forschungsgesellschaft) (Fidus Writer). In some cases funding comes from private donors (**Janeway**).

It is also common for a tool to be funded from several sources over a period of time; examples of this pattern can be seen in **FromThePage**, which is funded by four sources; and **Jupyter** and **The Pulter Project**, with ten funding sources each. In some cases a single funder provides several grants to support consecutive phases of development of a project. For instance **Melville's work** was initiated with the National Endowment for the Humanities Digital Humanities Start-Up grant (2008) and then supported by three Scholarly Editions grants.[59]

The resource's funding situation is also strongly connected to the national context. The project **Ports, Past and Present: Cultural Crossings between Ireland and Wales** is funded, in part, by the European Regional Development Fund through the Ireland Wales Cooperation programme. **New Panorama of Polish Literature** bases its sustainability on research grants that provide both new content and functionalities, given the difficulties in securing funding for sustaining existing services. Some public services, such as **POL-on**, a governmental platform for collecting data on research, have the stable backing of their respective ministries.

Crowdfunding is not a common option for professional academic texts. Among the analysed cases only **PubPeer** invites donations and has a Patreon account.[60] **Fidus Writer** accepts donations to cover some of its operational costs. **Language**

[59] https://melville.electroniclibrary.org/acknowledgments.html

[60] https://pubpeer.com/static/about

Science Press chose a distributed funding model in cooperation with Knowledge Unlatched, where research institutions worldwide contribute towards the financing of their publications.[61] **Free Ebook Foundation**, developers of the **GITenberg** project, opened a sponsors account on Github in 2019 asking for donations from 2 to 200 dollars monthly.[62] On the other hand, donations and grants might be considered insufficient to secure the future of a particular project, in which case a fee-for-service model is chosen, as in the case of **Refereed.** It's worth mentioning that some of the projects analysed are owned by large for profit corporations, for example, **Publons** belongs to Clarivate, and **Github** is owned by Microsoft. We discuss this issue in more detail in the sub-chapter on platformisation.

4.7.2 **Business and sustainability models**

The big question for almost every tool or service reliant on external funding is – what happens after the funding period ends? Some projects openly admit their uncertain future, such as **NextBook**: "Our goal is to become more diverse as soon as possible and sustainable by ~~April~~ December 2020." The crossing out of the earlier date when it was hoped that sustainability could be achieved, as well as the fact that this information is still displayed on the website in April 2021, shows the kinds of difficulties the teams behind such tools have in planning.[63] **Peerage of Science**, a for-profit company, stated that their revenue comes from "other organisations that want to purchase the peer review service to use in their decision-making, such as publishers, funding organisations, and universities."[64]

When it comes to business models, an interesting option is to apply diverse payment plans to different user categories or product versions. **Overleaf** offers their basic version for free, but it also provides premium functionalities (such as real time track changes, full document history, or priority support) in three pricing plans ranging from 9 to 28 euro per month. Under another tiered model, researchers as individuals have free access to particular tools or services, whereas institutions must pay a fixed or negotiated price, often to get access to advanced monitoring capabilities. Such arrangements are used by, for example, **ScienceOpen**[65] and **Kudos**: "The basic Kudos service is free for researchers to use; publishers, funders and institutions pay a fee for access to support tools, information on publication performance and author sharing effectiveness within Kudos, and also to supplement the data set available to help authors evaluate the impact of their use of the Kudos tools."[66] In some cases, there is no radical distinction between payment plans, as for **PubPub**, which differs mostly in the number of communities allowed to create free DOI numbers or access dedicated support.[67] Other projects use different plans: **Protocols.io** has Basic (free)

[61] https://langsci-press.org/knowledgeunlatched
[62] https://github.com/sponsors/EbookFoundation
[63] https://www.next-book.info/about-us/
[64] https://www.peerageofscience.org/faq/#faq1
[65] https://about.scienceopen.com/for-institutions/
[66] https://www.growkudos.com/about/user_guide
[67] https://www.pubpub.org/pricing

and Carbon (payable) plans; **Authorea** has Free, Premium (for Individual User), and Community (for institutions also?) plans; and **FromThePage** provides different plans for researchers, and small and large organisations. **FairCopy**, a TEI-editor in development, claims that it will be affordable to both individuals and institutions, and that one fee will cover unlimited use of the product and automatic updates.

It seems, in general, that the main differences between free and premium accounts/payment plans lie in, a) support priority, and b) unlimited access to functionalities (e.g. number of documents, projects etc.).

It should also be mentioned that one of the general business models supporting digital publishing is to sell the printed books (e.g. **Manifold**). This solution was applied as an experiment by McKenzie Wark, the author of Gamer Theory, a digital monograph in a form of a website: "Why give the book away free if you want to sell it later, it's an experiment to test the theory that if you make a gift out of something people feel better disposed towards it."[68]

Another sustainability strategy is to establish consortia for mutual support and shared services. For instance, **ScholarLed** is a consortium of five open access publishers formed to strengthen the position of small publishers: "Members of the consortium each retain their distinct identity as publishers, with different audiences, processes, business models and stances towards Open Access. What they share, however, is a commitment to opening up scholarly research to diverse readerships, to resisting the marketisation of academic knowledge production, and to working collaboratively rather than in competition." (C21)

4.7.3 **Persistent identifiers and standardisation**

Sustainability is achieved not only by financial means, but also through data standardisation and protection. **Muruca's** developers emphasise that their strategy in this area involves accessibility and using persistent identifiers: "Muruca has developed a module for publishing data in Zenodo, the CERN Data-Center backed research data repository:

- **Citable** – every upload is assigned with DOI, to make them citable and trackable;
- **Persistent** – a JSON with resources metadata is stored on Zenodo for as long as CERN exists;
- **OAI-PMH support** – and custom integration for database aggregators" (Andreini 2020).

Melville Electronic Library also adopted a standardisation of their graphic resources using IIIF standards "to ensure flexibility and integrity in uploading images for editions and displays, and to ensure sustainability and the lasting reliability of MEL as an online centre for Melville studies."[69] **Polona** refers to these standards [Rosa 2019: 27], while **Pulter Project** and the **Faust Critical Edition** apply the TEI-XML standard for digital editions.

[68] http://www.futureofthebook.org/gamertheory2.0/index.html@page_id=228.html

[69] https://melville.electroniclibrary.org/tool-kit

A DOI (Digital Object Identifier) remains the main persistent identifier, used or recommended by such projects as **Kudos**, **Peeriodicals**, **Refereed**, **Muruca**, **HIRMEOS**, and **Stanford University Press**. Other types of persistent identifiers used in the analysed cases were: Web of Science ResearcherID (**Publons**), ISSN (**Refereed**), ORCID (**HIRMEOS**), project specific unique identifiers (**dokieli**, **Octopus**), ISBN (**Stanford University Press**), and OCLC (**Stanford University Press**).

4.8 Impact and prestige-generating mechanisms

This section discusses how the analysed projects are used. We look at the examples of use and the impact these services may have on society and education. We are also interested in prestige-generating mechanisms and outreach strategies.

4.8.1 Metrics

The various types of statistics reported by platforms serve as a source of knowledge about their size and impact. **FromThePage**, a software for transcribing documents and collaborating on transcriptions with others, provides statistics for transcribed documents (1,136,555 as of 16 April 2021) and a "Recent Activity" feature. **Language Science Press** includes statistics in yearly achievement posts. For example, in 2019 they published 30 books and their “pdfs were downloaded 362,983 times (+222,550 compared to 140,433 in 2018), for a grand total of 680,057)” (Nordhoff 2019). **Protocols.io**, a methodology-sharing platform, displays a number of created protocols.[70] In some cases Google Analytics is used to collect data on user activity (**New Panorama of Polish Literature** and **Black Quotidian**).

HIRMEOS, a project that prototypes innovative services for monographs in support of open science infrastructure, proposes metrics collection and aggregation from third-party platforms through the Statistics Collection Agent. They collect information about metrics using Google Analytics, Access Logs, Google Books, Open Edition, OAPEN, Wikimedia, Unglue.it, The Classics Library, OpenAIRE, IRUS-UK, JSTOR, Matomo (Piwik), and World Reader and Identifier Translation Service to unify the different identifiers across many platforms (Arias 2018).

4.8.2 Citations

Citations are still considered to be the basic impact measurement for scholarly outputs. Although citations themselves aren’t anything innovative, we can identify new ways of “boosting” them. **Kudos** is a web-based service that helps researchers increase the impact – and citations – of their publications. After registration, Kudos will lead scholars through many steps “that prompt them to explain their publications; add context and enrich them with links to resources such as images and data; and share information about their publications via social networks and email.”[71] Authors can track the reach of their outputs, measured in citations, via the Kudos platform.

[70] What is more, plenty of organisations ENCOURAGE the use of protocols.io (like PLOS, eLife, Gordon and Betty Moore Foundations, NIH, University of California,Carnegie Mellon University and more), and among these are over 500 journals. (Open Publishing Festival: New Tools in Publishing, 2020. https://www.youtube.com/watch?v=66HvylmIwjwandab_channel=LiberateScience)

[71] https://www.growkudos.com/about/user_guide

Publons, a tool with a range of peer review solutions, provides a visually interesting feature: a citation map that serves as "a geographic representation of your citations across the globe."[72] This feature is for users interested in reaching international audiences, therefore, a more-locally oriented discipline might not be equally represented.

Analysed projects often list their outputs in order to show their impact. **Muruca**, a CMS for digital editions, is currently used by several digital scholarly projects, so developers created a portfolio of their collaborations.[73] Listing renowned institutions using a particular tool may be a source of prestige. For instance, **Jupyter** is currently used by Google, Microsoft, IBM, Bloomberg, Soundcloud, Michigan State University, NYU, and Berkeley University of California among others. Similarly, **Electronic Book Works** published a list of their clients, including Encyclopedia Britannica, Foundation for Human Rights, and Oxford University Press.[74]

The amount of training and number of workshops carried out, as well as the prestige of the trainees, can also be a factor that indicates the impact and prestige of a service or tool. **ContentMine**, a project with a broad range of text mining services, delivers workshops for institutions like "the Wellcome Trust, Virginia Tech, MSI Global, Jisc and FutureTDM; and at the Mozilla Festival, SciDataCon, and other events."[75]

Press coverage in influential (but not necessarily scientific) media outlets also generates and demonstrates impact. **Authorea** showcases media coverage about their project, including in Scientific American, Times Higher Education, The Hundert, Tech Crunch, Nature, and Huffington Post.[76] Various awards are also a natural prestige-generating mechanism. For instance, **Electronic Book Works** displays a list of awards they have received.[77]

4.8.3 **Social and educational impact**

To achieve social and educational impact it is always worth asking which features of a given project are especially important for the broader audience. Some projects provide popular releases targeted at non-scholarly users. For instance **Electronic Books Works** produced The Economy ebook, which can be used in higher education. Each chapter has colourful themes (History, Global economy etc.), links to figures and to external websites, quiz-questions after each chapter, and options for bookmarking sections of the book (headline, footnote, tips, questions, etc.). The book also has five different language versions: English, French, Italian, Spanish, and Finnish.

Some projects organise targeted workshops for teachers to demonstrate how their resources can be used in the classroom (e.g., the symposium organised by digital

[72] https://publons.freshdesk.com/support/solutions/articles/12000071842-what-is-my-citation-map-

[73] https://www.muruca.org/portfolio/

[74] https://electricbookworks.com/about/

[75] http://contentmine.org/training-and-workshops/

[76] https://www.authorea.com/aboutus

[77] https://electricbookworks.com/about/

edition's creators "Teaching with The **Pulter Project**").[78] Digital scholarly editions collected in **Melville's Electronic Library** are listed as a resource for teaching Melville in schools.[79] Even greater societal impact is available when a tool or service is recommended on governmental websites: like the digital collection The Roots of Janusz Korczak published by New Panorama of Polish Literature, which is listed on the portal Lektury.gov.pl as recommended reading for students.[80]

4.8.3.1 Promotional strategies

Social media remain the main channels for promoting projects. The most common strategy is to use at least one of the following channels: a blog, Twitter, or Facebook. It is quite surprising that channels like Instagram and Pinterest, which rely mostly on graphics and photographs, aren't frequently used, especially for such visual projects as the **Polona** digital library.

Twitter is probably the most popular choice among the analysed cases. What is interesting is that the developer (personal) accounts were often more popular than the product accounts. For example, Rebus Ink has 215 followers on Twitter while Zoe Wake Hyde, the project leader, has 812 followers.

The creators might also promote their project widely by participating in numerous conferences, seminars, or webinars (**Manifold**), or by creating their own and inviting others to them (**FromThePage**). Interviews and podcast invitations might also be a chance to gain an audience, like the **Pulter Project's** participation in a podcast about early English book culture.[81]

Considering the tendency for projects to be community-driven projects, it might also mean that promotional actions should also be left to them. **Peeriodicals'** creators assume it is up to the editor(s) of a particular "peeriodic" (created on their platform) to build audience for their journal and promote it. They are encouraged to do this by using social media, which Peeriodicals supports through their Twitter account.

4.9 Bibliography

Andreini, Giulio. 2020. 'Muruca Framework Presentation'. Presented at the DSE Project, November 6.

Arias, Javier. 2018. 'Collecting Inclusive Usage Metrics for Open Access Publications: The HIRMEOS Project'. In *ELPUB 2018*, edited by Leslie Chan and Pierre Mounier. Toronto, Canada. https://doi.org/10.4000/proceedings.elpub.2018.11.

[78] https://networks.h-net.org/node/73374/announcements/4726831/teaching-pulter-project

[79] https://teachingmelville.org/resources/

[80] https://lektury.gov.pl/lektura/krol-macius-pierwszy

[81] https://brocku.ca/brock-news/2020/07/new-podcast-episode-explores-early-english-book-culture/

Brumfield, Ben. 2020. 'The Decade in Crowdsourcing Transcription'. *Content.Fromthepage.Com* (blog). 2020. https://content.fromthepage.com/decade-in-crowdsourcing/

Cantiello, Matteo. 2016. 'How to Bring Science Publishing into the 21st Century'. Scientific American Blog Network (blog). 10 August 2016. https://blogs.scientificamerican.com/guest-blog/how-to-bring-science-publishing-into-the-21st-century/.

Castro, Antonio Rojas. 2019. 'Recogito-TEI Working Group: From Semantic Annotation to Minimal Digital Editions'. Medium. 21 June 2019. https://medium.com/pelagios/recogito-tei-working-group-from-semantic-annotation-to-minimal-digital-editions-4f578334968e.

Delmont, Matthew F. 2019. Black *Quotidian: Everyday History in African-American Newspapers*. Stanford University Press. http://www.sup.org/books/title/?id=29420.

Eve, Martin Paul, and Andy Byers. 2018. 'Janeway: A Scholarly Communications Platform'. Insights 31 (0): 15. https://doi.org/10.1629/uksg.396.

Hellman, Eric. 2018. 'A Milestone for GITenberg'. Go To Hellman (blog). 30 October 2018. https://go-to-hellman.blogspot.com/2018/10/a-milestone-for-gitenberg.html.

Himmelstein, Daniel. 2020. Writing Manuscripts on GitHub: Can Automated Software Replace Publishers? https://youtu.be/UTepYeOpFrk.

Hyde, Zoe Wake. 2020. 'Introducing Rebus Ink: A Research Workflow Tool to Help You Do More with Your Digital Texts'. June 24. https://youtu.be/lDX22HZjGnk.

Kobyliński, Łukasz. 2020. *PolEval 2020: Wprowadzenie.* YouTube. https://youtu.be/X3RTISOw8Pc.

Kocurek, Jozef, Matěj Málek, Jan Martinek, and Lukáš Porsche. 2016. 'Next Book'. Platform for publishing. Next-Book.Info. 2020 2016. https://www.next-book.info/.

Maryl, Maciej, Costis Dallas, Jennifer Edmond, Jessie Labov, Ingrida Kelpšienė, Michelle Doran, Marta Kołodziejska, and Klaudia Grabowska. n.d. 'A Case Study Protocol for Meta-Research into Digital Practices in the Humanities'. Digital Humanities Quarterly.

Nordhoff, Sebastian. 2018. Language Science Press Business Model. Berlin: Language Science Press. https://doi.org/10.5281/zenodo.1286972.

Nordhoff, Sebastian. 2019. 'Achievements 2019 | Language Science Press Blog'. 2019. https://userblogs.fu-berlin.de/langsci-press/2020/01/29/achievements-2019/.

Lomas, Natasha. 2014. 'Authorea Nabs $610k For Its Bid To Become A "Google Docs For Scientists"'. TechCrunch (blog). 22 September 2014. https://social.techcrunch.com/2014/09/22/authorea-seed/.

Rosa, Igor. 2019. 'Digital Library Polona: Digitization, Technology, Cooperation'. *Slavic and East European Information Resources* 20 (1–2): 23–30. https://doi.org/10.1080/15228886.2019.1628495.

Rühling, Barbara, ed. 2018. Editoria Community-Owned Book Publishing Platform. Shuttleworth Funded. https://coko.foundation/wp-content/uploads/2018/07/Editoria_Community-ownedBookProductionPlatform_screen.pdf.

Sonntag, Susanne. 2018. 'Weltexpresso - BUCHPUBLIKATION / DIGITALE EDITION 1.0'. 2018. https://www.weltexpresso.de/index.php/buecher/14188-buchpublikation-digitale-edition-1-0.

Stoye, Emma. 2015. 'Post Publication Peer Review Comes of Age'. Chemistry World (blog). 2015. https://www.chemistryworld.com/news/post-publication-peer-review-comes-of-age/8138.article.

Szelszyński, Bartłomiej. 2019. 'Digital Humanities: Project Funding versus Continuity of Research'. In *Proceedings of the Digital Humanities Congress 2018*. Studies in the Digital Humanities. Sheffield: The Digital Humanities Institute. https://www.dhi.ac.uk/openbook/chapter/dhc2018-szleszynski.

5 Prototypes: directions for the future

During the research process we identified two main innovations that seemed particularly important from the vantage point of OPERAS. These were the "living book" and the "digital scholarly edition." The former was considered to be an important form with which to update scholarly writing in connection with OPERAS Special Interest Groups. The latter was identified as an important (but underserved) form allowing for an innovative connection between the text and data.

The work on the living book was conducted in cooperation with WP3 of OPERAS-P, and the actual prototype is live on the OPERAS website.[82] The section on digital scholarly editions was consulted about with various OPERAS partners and resulted in a conceptual blueprint for further work. Both projects should be considered conceptual prototypes for future OPERAS services.

[82] https://www.operas-eu.org/special-interest-group-living-book/

5.1 The Living book

The following text was submitted as OPERAS-P deliverable 3.4, the living book. Work on the living book was carried out within the WP3 of the OPERAS-P project, which was focused on providing support for OPERAS' infrastructure, and was informed by the analyses and explorations of Task 6.5 (*Future of Scholarly Writing in SSH*). Focussing on the analysis of the innovation landscape as well as on the case studies presented in this report, technical specifications were defined by Mateusz Franczak (IBL PAN), and implemented by Yoann Moranville (OPERAS' technical coordinator) and Judith Schulte (OPERAS' Communication Officer), with contributions and support from Maciej Maryl (IBL PAN), Agnieszka Szulińska (IBL PAN), Pierre Mounier (OpenEdition), Chloé Lebon (OpenEdition), and Marta Błaszczyńska (IBL PAN).

5.1.1 The living book: embedding fluid and collaborative scholarly communication

5.1.1.1 Introduction

In August 2018, OPERAS launched a series of white papers prepared by its Special Interest Groups (SIGs), which covers the landscape of scholarly communication, addressing such issues as advocacy for open science, common standards among research infrastructures and service providers, platforms and services, the multilingualism of publications, open access business models, and tools.

Since the publication date, all documents have been viewed, altogether, almost 8,500 times and downloaded over 4,500 times, which indicates their relevance for the community. Although these papers provide a comprehensive, reliable overview of the state of the art, the three years that have passed since the time of their publication represents a significant period in terms of the ensuing development.

White papers are core to OPERAS' operations as they codify the state of the art as well as the paths the infrastructure can take. Moreover, they are products of sustained team effort by SIG members collaborating on the subject and contributing their unique perspectives. Hence, when discussing the present White Paper update, we dedicated substantial time not only to their content but also to choosing the right format that would accommodate the need for future updates as well as foster community discussions. This is why we started to investigate the concept of the living book.

This section provides a summary of our work on the living book, starting with the definition of this genre and an overview of notable examples. This is followed by the needs analysis and technical specification of the OPERAS living book, providing the rationale behind our technical decisions.

5.1.1.2 What is a living book

The concept of the living book was born out of a dissatisfaction with the limitations of print communication imposed on the new digital genres. Due to technological constraints, the printed book (or its digital equivalent in the form of an e-book in pdf or epub format) codifies its message in a finished, rarely updated volume. Through centuries of writing and reading, books have been positioned as products of certain

(usually long) processes, distributed to the audience in a slow and limited feedback circle: readers receive the work and can respond through established channels like reviews of references in other articles and books. However, as the authors of the reports stemming from the *Academic Book of the Future* project observe, the current technological advancement allows for a disconnecting of thought from the form in which it is disseminated, leading us to consider different, perhaps more suitable vehicles for our arguments. We are empowered "to produce new kinds of books, with extended texts, narratives, ideas, and arguments produced in new ways, with dynamic and interactive images, graphics and sounds; links within the text and to external sources; and facilities for updating and annotation" (Jubb 2017: 14). This means increased flexibility on many levels, as "writing may become more collaborative, and much more influenced by embedding or linking to a wide variety multimedia and other content, and the use of interactive features in the presentation of the book" (ibid.:33). This openness also entails linking the text with external resources and data through hyperlinks and references. The texts themselves become elements of larger ecosystems and user networks, where "social reference management allows individual users to share personal libraries and exchange reviews, notes, and recommendations in order to find the most valuable references through the collective choices of their peers" (Ren 2013: 745). Thus, new technologies allow us to reconsider the shape of the argument and the means to convey it.

As Xiang Ren (ibid.) pointed out, "a growing number of digital publishing initiatives are approaching scholarly communication in new ways and incorporating dynamics of openness, networking, and collaboration into their most basic functions." However, what seems to constitute a living book is its temporal dimension, i.e., a certain liquidity of content, which is impossible to achieve in print forms. It allows for alterations, additions and comments to be added along the way. That is why it is hard to draw a clear line between a living book and such genres as the open science notebook or wiki. It seems that being a "living" document is simply one of the features shared by all these formats. In order to broaden our understanding of this genre we conducted several case studies looking at the available tools for creating and presenting digital publications. Below we discuss the most relevant examples that will allow us to distil the key features of living books.

An interesting case to start with is *Living Books About Life*, published by Open Humanities Press (OHP), with funding from the Joint Information Systems Committee (JISC). This is an open access publishing series consisting of selections of texts about life, from both the science and humanities' disciplines. Each book in the series was developed and compiled from already existing open access publications, very often linked to external repositories. While the project was intended to be a "living" series, its innovative potential has been limited to remaining available to users who wish to edit, update, remix, and add comments to a given set of texts, suggesting the inclusion of other publications, attaching hyperlinks, etc. However, the form of the book itself has remained traditional, and the individual publications included in the book do not possess any unusual functionalities. Consequently, we can speak of a limited "livingness" in this case.

Another example is *Culture Machine Liquid Books*, a series of experimental digital publications that allows open editing and the addition of content. Users can rewrite, tag, remix, and reuse all the books in the series. Technologically, the series uses a wiki engine that enables logged-in users to create their own versions of articles, in a similar way to Wikipedia. Despite the use of various open solutions, most of the publications included in this collection consist of webpages with sets of links to external documents (some of the links have expired and have not been updated since). The individual pages look like blog posts and one can comment only on the entire page, not on particular paragraphs. The most recent user comments come from 2017. Liquid Books continues to exist and more publications are planned. However, the editors of individual publications together with Open Humanities Press have stated that selected publications will be frozen and published as finished texts (while remaining open access).

Apart from regular publishing initiatives, there are tools and plugins dedicated to providing "living" solutions for publications hosted on private and institutional websites. One such plugin is CommentPress, whose open source version was released in 2007. It is a WordPress plugin created by the Institute for the Future of the Book, whose goal is to apply the modern technological solutions used in blogs to more complex, slow-developing works that require advanced text organisation. One of the important innovations this tool brings to living books is being able to position comments next to the text rather than below it, which helps to achieve a visual representation of the dialogue and show the book as a work in progress, developed through the conversations of commenters. An additional innovation is the possibility of commenting on selected fragments of the text (paragraphs, sentences, words), which was not possible in earlier, typically blog-based solutions. The limitations of this tool include the rigid structure of the template, which makes it inconvenient for adding multimedia materials in a free and open manner. CommentPress is definitely a tool that gives primacy to the text, so despite innovative technological solutions, it treats the book in a traditional way.

Some tailored platforms offer more services to support and facilitate digital publishing. Scalar is a project developed by the Alliance for Networking Visual Culture (ANVC) in collaboration with Vectors and IML, and supported by the Andrew W. Mellon Foundation and the National Endowment for the Humanities. It is a free open source platform designed to make it easier for authors to write long-form, born-digital scientific texts online. Scalar enables users to attach different types of media and juxtapose them with text without requiring advanced technical knowledge. It is a semantic web development tool that provides a balance between standardisation and structural flexibility for different types of media. In addition, the platform supports author collaboration and reader commentary. Adapting the platform to more advanced needs, however, requires cooperation with a developer.

To sum up, we may identify two key features that are central to the concept of living books: fluidity and collaboration. The general concept of a fluid or unstable text precedes electronic forms of communication and is well established in literary studies as it signifies the modifications undergone by texts while functioning in culture, be it

editing, copying, reissuing, etc. This quality of texts is well recognised by creators of scholarly editions who need to deal with many sources while reconstructing a critical version of a given work (Cf. Bryant 2005). While the digital environment opens new possibilities for scholarly editions – like dynamic collation, variant analysis (McGann 2001) – it also allows for the fluidity of other texts, including scholarly outputs.

In general, this means that a given output does not have to be "frozen" in a particular moment, as "erasing the artificial distinction between process and product" (Priem 2013:495) is enabled by the Web. This may have two implications. On the one hand, the web allows the final output to be versioned and updated. It is a phenomenon Juhás et al. dubbed a paradigm of "continual improvement in scholarly publishing" (2018:245), a metaphor drawn from software development that treats the published version as a snapshot of an ongoing creation process subject to comments and feedback loops. A scientific paper, they write, is understood as "a dynamic document evolving in time, which can have different versions and releases, published online, enabling incremental and continual improvement in analogy to software as a service, with [the] software['s] new version releases and software support enabling continual improvement of [the] software" (ibid.) New versions of a text can be provided (as they are in many repositories like Zenodo or wiki-based publications. cf. Mietchen et. al 2011) and are open for feedback and comments (as in the case, for instance, of the reviewing process at F1000Research). On the other hand, the fluidity of web textuality enables genres and formats that don't have to be the final products of a given research study and could be published and constantly updated at various stages of a research project as scholarly blogs or open science notebooks.

The second living book feature we wish to highlight – collaboration – is interconnected with fluidity as they are both allowed by the affordances of web technologies. By "collaboration" we mean opening the text to other authors and contributors, who can then share early findings and receive quick feedback from interested communities. For instance, this is a key feature of scholarly blogs, which are dubbed "creative catalysts" (Kjellberg 2010) because they serve as platforms where new ideas emerge from such communication. Similarly, open publishing initiatives that allow the community to conduct open conversations on published works are considered more transparent and democratic than traditional, closed peer-review processes (Ren 2013:745). This immediate feedback is also an essential component of the continual improvement paradigm mentioned above, whereby reviews sketch recommendations for further improvement. With this overview in mind we began the process of designing the OPERAS living book.

5.1.2 **Technical specifications**

5.1.2.1 Requirements

The specific technological requirements and functionalities used in the living book are based on OPERAS' needs, which were defined earlier. We describe them here in greater detail:

a. **Location and sustainability.** The first important issue is to choose where the living book will be published. It was important to us that the publication be closely linked

to OPERAS and that the server chosen provide sustainability, technical support, and editing ability without involving external developers or service providers.

b. **Automatically updated Zotero bibliography.** Another requirement that had to be met was the implementation of a Zotero bibliography in the living book. Each report has its own collection in Zotero, which can be updated and edited at any time. It is important that all changes made in Zotero are immediately visible in the publication.

c. **Versioning.** As the activities described in the white papers are constantly being executed and developed, reports need to be updated regularly to reflect the state of the art. The answer to this is to implement versioning, which will allow for both updates and the comparison of new versions with earlier ones.

d. **Referencing.** Being able to conveniently cite white papers increases their visibility in research circles. We therefore felt it was important to find a convenient solution for displaying citations.

e. **Annotating and commenting.** The ability to comment and annotate is one of the most important features of a living book. By adding comments, hyperlinks, and building semantic relationships with other pages and articles; inter-research dialogue, collaborative writing, and research development become easier.

f. **Quality control (peer-review)** By allowing comments, publications are subjected to critiques and reviews, so they can be quality-controlled by a much larger number of researchers than traditional publishing models. This functionality also enables open peer review, which is becoming an increasingly popular review model, and is supported by the European Commission, policy makers, and funders, as well as many international research institutions.

5.1.2.2 Solutions

Preparation of the OPERAS living book was preceded by an overview of the existing tools for creating and presenting digital publications. In addition to meeting technical requirements, tools and services were checked for stability, technical support, regular updates, and sustainability.

While technical solutions for the living book are ready, the white papers will be updated and reviewed in the future.

a. Wordpress – the living book running on OPERAS' server

After considering various options, a decision was made to use WordPress with plugins ensuring the implementation of the required functionalities. We surveyed the available options but didn't want to use paid services for sustainability reasons, nor could we implement tools like Scalar, which did not support the default functionalities we defined as being crucial. The ability to freely transform, edit, use plugins, and overall sustainability were all in favour of this choice. By using WordPress, it was possible to publish the living book within the OPERAS website without any difficulties. Moreover, it thus became an integral part of the OPERAS environment, closely linked to information about current activities. WordPress is one of the most popular content management systems available. It is freely configurable, customisable to personal requirements, has good technical support, and is a well-known and widely used system (it has been in operation since 2002).

b. Zotero bibliography via ZotPress

By using WordPress it was easy to implement the bibliography collected in Zotero – it was enough just to install the appropriate plugin: ZotPress. This allows the Zotero collection to be linked with the WordPress site. In this way, the bibliographies of each white paper were automatically connected to the text. In addition, every time an item is added or removed from the collection, ZotPress automatically updates the bibliography on the website, so the reader can be sure that the references in the report are up-to-date.

c. Versioning solution through separate PDF files

Another issue to consider was versioning. For reports that describe the current state of knowledge, it is important to perform ongoing updates because reports can quickly become outdated. This is crucial due to the development of research and the production of new information. We decided to present new versions of the white papers within the living book, so users could navigate between versions with a single click on the report's subpage. In this way, we avoided linking to external sites and were able to present the updated reports in a consistent way while maintaining the same layout. The relevant collections in Zotero have also been updated for the new versions, making it easy to compare bibliographies and check new references.

d. Citations added manually

Initially, we planned to use the Cite plugin for citations, but after checking its technical capabilities, we found out that it does not allow multiple authors for one publication. Therefore, we decided to add citations manually, following the visual solution used in Cite.


OPERAS Advocacy White Paper, July 2018

Cite this article as: Elisabeth Heinemann, Andrea Bertino, Francesca Di Donato, Aysa Ekanger, Elena Giglia, Barbara Jędraszko, Michael Kaiser, Lisa Matthias, Alessia Smaniotto, Stavroula Sant-Geronikolou. (2018, July 30). "OPERAS Advocacy White Paper", July 30, 2018, https://doi.org/10.5281/zenodo.1324036.

DOI: 10.5281/zenodo.1324035

e. Annotations and comments using Pundit

The last functionality that needed to be implemented was an annotation tool. We decided to use Pundit Annotator, designed and developed by Net7, which already cooperates with OPERAS on various projects. With Pundit Annotator, you can comment on and highlight selected parts of the text, add hyperlinks, hold discussions and suggest modifications. The Pundit Annotator is free to use and it only requires the user to log into their account. Web annotations serve to enrich content, enable evaluation and addition of sources, and facilitate and support collaborative research. They can also be used for open peer review. The possibility of browsing comments using the user's notebook is very convenient.

f. "Reviewathons": quality control using Pundit

We decided to use Pundit not only for annotations, but also as a tool for applying open peer review to the living book as Pundit allows its registered users to add their comments to both individual fragments of the text and entire white papers. In connection with this, we plan to conduct "reviewathons," i.e., two-week reviewing marathons during which invited reviewers and members of the audience will comment on and discuss the paper. By enabling asynchronous communication and focusing on a limited time-period, we plan to stimulate a genuine conversation around the white papers, which may lead to the preparation of new versions. OPERAS Special Interest Groups will be responsible for coordinating community feedback for their white papers.

5.1.2.3 Conclusion

The living book prepared in the course of the OPERAS-P project is a prototype aimed at responding to the needs of the OPERAS community. Based on the needs assessment, the analysis of similar projects, and available technology we prepared tailored solutions aimed at addressing the key requirements of this community. We have discussed this process in detail above.

Now, it is the OPERAS community's turn to take advantage of this living book and use it according to the needs expressed earlier. We will work closely with SIG leaders during "reviewathons," observing how the discussion on the white papers unfolds and how actual users interact with this format. The lessons learned will inform our understanding of scholarly communication and perhaps modify the living book in the future according to the emerging needs of the community.

5.2 Bibliography

Bryant, John (2005) The fluid text: a theory of revision and editing for book and screen. Ann Arbor: University of Michigan Press.

Jubb (2017), Academic Books and their Futures: A Report to the AHRC and the British Library, London, https://academicbookfuture.files.wordpress.com/2017/06/academic-books-and-their-futures_jubb1.pdf

Juhás G, Molnar L, Juhásová A, et al. (2018): Continual Improvement Process in Scientific Publishing. In ICETA – IEEE International Conference on Emerging eLearning Technologies and Applications, Proceedings. edited by František Jakab, Institute of Electrical and Electronics Engineers Inc. 2018.

Kjellberg S (201): I Am a Blogging Researcher: Motivations for Blogging in a Scholarly Context. First Monday. 2010; 15(8).

McGann, Jerome J. (2001) Radiant textuality: literature after the World Wide Web. New York: Palgrave.

Mietchen, Daniel and Hagedorn, Gregor and Förstner, Konrad and Kubke, M Fabiana and Koltzenburg, Claudia and Hahnel, Mark and Penev, Lyubomir. (2011). Wikis in scholarly publishing. Nature Precedings. 6. 10.1038/npre.2011.5891.1.

Priem J (2013) Beyond the Paper. Nature. 2013; 495 (7442):, p. 437.

Ren X (2013) Beyond Open Access: Open Publishing and the Future of Digital Scholarship. In ASCILITE – Australian Society for Computers in Learning in Tertiary

Education Annual Conference. edited by Gosper M., Hedberg J., and Carter H., 2013. http://www.ascilite.org/conferences/sydney13/program/papers/Ren.pdf

5.3 Digital Scholarly Editions

5.3.1 **Scholarly aspects of editing**

Digital scholarly editions (DSE) are a development within a specific field of literary research, namely scholarly editing. Before we move on to consider the digital aspect and all the issues related to it, it is worth emphasizing the *scholarly* editing aspect, because it immediately points to some specific features of editing activities that are related to it.

In the context of the paper medium in particular, scholarly editing is the range of activities performed by professional scholars (mostly from the humanities disciplines[83]) in two main areas: 1) identifying the relevant variants of a work and comparing them, with the aim of proposing a definitive edition, often one that is intended to be "the nearest approximation in every respect of the author's final intentions" (Tanselle 1976); and 2) reconstructing all possible existing versions with a focus on their socio-historical contexts (McGann 1992) and to provide critical apparatus and scholarly commentary.

While the textual basis of a work, established as a result of scholarly editing, could (and even should) later serve as the basis for popular editions for a wider public, the scholarly edition itself is mostly intended for professional audiences (including university students). This focus on professional audiences affects the selection of texts for scholarly editions (for the most part, texts that are important to researchers are not necessarily attractive to a wider readership), the way they are presented, and the way the critical apparatus and scholarly commentary are drawn up.

The same holds true for digital editions, whereby editing activities are conducted by professional researchers (editors) for professional researchers (scholars) in a digital environment. Regardless of the involvement of IT developers and graphic designers, the final output is still quality-assured by the work and skills of scholarly editors. When compared to printed editions, digital scholarly editing projects produce a more widely accessible and reusable content. However not all digital affordances are equally attractive to all user groups. It is, then, vital to precisely identify the audience and users' needs. In the case of a digital scholarly edition, as we said, the main target should be SSH scholars, especially editors. It is advisable to distinguish the various kinds of editions for different groups of users.

For example, digital scholarly editions of contemporary Polish dramas will be published on the TEI.NPLP.PL platform and is intended for professional users. However, since the project also includes an aspect of popularisation in the preparation of a printed popular edition, its digital version will be available on the New Panorama

[83] According to "Digital Editions of Text: Surveying User Requirements in the Digital Humanities" (Franzini, Terras, and Mahony 2019): "With respect to the specific disciplines, of the 218 responding participants 44% are involved in literary studies and 22% conduct historical studies [...] Almost 82% of participants identified themselves as belonging to the Humanities, 10% as working across different disciplines and 7.85% as belonging to the Applied Sciences."

of Polish Literature platform (NPLP.PL) – a platform designed to publish digital collections having a more educational approach, and not loaded with critical apparatus. There are some attractive digital editions (such as the Canterbury Tales Project), where most of the functions are designed to enrich its popularizing aspect. As such, this edition has a read-aloud function for each verse, but lacks many functionalities that are tuned to the needs of scholarly editors, such as the ability to juxtapose different layers of commentary, or transcription with transliteration.

5.3.2 **Digital versus paper editions**

Digital editions cannot simply mimic paper editions. However, the development of specific digital methods of scholarly editing is in progress, although more standardisation is needed. These methods and standards must take into account hybrid models, where both print and digital forms of editions are being produced in parallel.

The case of the digital edition of Faust (a project lead by Prof. Anne Bohnenkamp-Renken, Faustedition.net) implements such a model. The book publication[84] contains a critical ("konstituierten") text of Faust and selected facsimiles with transcriptions. The digital edition incorporates all possible versions, and presents the source ("genesis") and metadata for every single verse of this drama. Additionally, "a comprehensive collection of manuscripts with over 2,000 pages of descriptions'' is included in the digital edition and "there are all the relevant prints published during Goethe's lifetime and over 1,500 testimonies to the creation of the work."[85]

An edition of letters by the Polish poets of the Skamander literary group (Kazimierz Wierzyński, Jan Lechoń, Mieczysław Grydzewski) also serve as a good example of a parallel edition. This digital edition on TEI.NPLP.PL, encoded in TEI (English: *Text Encoding Initiative*) standard, doesn't merely follow the annotation model of the parallel print edition, but also offers a set of tools and solutions that were *born digital:* a TEI editor designed for encoding letters, and an index of entities corresponding to the basic categories of semantic data found in that corpus (such as people, places, journals, works, and organisations[86]). Such hybrid models are often forced upon scholars by evaluation mechanisms, which disregard or undervalue novel, digital publications. Hence, authors decide to prepare a double publication, a strategy described earlier in this report.

It is hard to overstate the importance of the sustainability of digital scholarly editions given the amount of work and financial resources needed to create them. Analogue and digital editing paradigms, although not identical, can act as forms of mutual back-up, and substantially increase the overall accessibility of an edition to its readership. In the case of a book edition, it should be noted that it may reach a

[84] Goethe, Johann Wolfgang: Faust. Eine Tragödie. Konstituierter Text. Bearbeitet von Gerrit Brüning und Dietmar Pravida, Göttingen 2018 (Johann Wolfgang Goethe: Faust. Historisch-kritische Edition. Hrsg. von Anne Bohnenkamp, Silke Henke und Fotis Jannidis unter Mitarbeit von Gerrit Brüning, Katrin Henzel, Christoph Leijser, Gregor Middell, Dietmar Pravida, Thorsten Vitt und Moritz Wissenbach).

[85] https://www.weltexpresso.de/index.php/buecher/14188-buchpublikation-digitale-edition-1-0

[86] It is evident that in following editions, this typology should be expanded.

completely different audience with more traditional habits or with insufficient technological resources (not least a stable internet connection) to use a digital edition. The print publication also has a well-established scholarly distribution model through publishing houses and bookstores.

5.3.3 **Examples of digital solutions to editors' problems**

In this section we look at how the features identified above could be used in practice to resolve complicated issues. The projects discussed below are in different stages of implementation, some are still in the planning phase.

5.3.3.1 Collating

One text that has several very different versions is Janusz Krasiński's contemporary drama *Czapa*. The versions differ not only in length, but also in their intended use for particular mediums. First a radio version was created, then it was expanded and reworked for theatre purposes, and finally rewritten for television. In the print edition, the editor was forced to choose one of the versions due to the page limit and financial constraints of the project. In the digital edition, currently developed at TEI.NPLP.PL, all three versions will be published to present textual transformations between the three media and emphasise the performative character of the work.

Another interesting case is the 19th-century Polish novel *Dzieci* (*Children*) by Bolesław Prus. The novel originally circulated in the press as a serial novel in two different versions both entitled *Świt* (*Dawn*). It was later expanded and republished, again as a serial novel in a journal, under the title *Dzieci* (*Children*). Finally, a book version was released, and although it was severely censored, it served as a basis for later popular editions, being the last version released during the author's lifetime. Another interesting feature of this work is that a fragment of the original typescript of *Dzieci* (*Children*) has been preserved, together with the author's handwritten corrections and the author's creative notes documenting the reasons for the changes and their motivations. The digital edition will make it possible to show the author's creative process by collating the changes in subsequent versions, corrections made to the typescript, and creative notes. It will also allow the pre-censorship version to be presented.

The situation is even more complicated in the case of *Samuel Zborowski, a* drama by the eminent Polish Romantic poet Juliusz Słowacki. The main challenge for the editor is that the work's manuscript is available on separate, unnumbered pages – which, thus, does not indicate an obvious linear order to the reading of the text or the sequence of scenes. An additional factor is that there are 19th-century drafts of the work that are richer in content than the final manuscript, as well as a tradition of important theatrical performances of the text that imposed a particular order. All of this makes the digital environment ideal for presenting the work as in the primary, scattered form. Editors may add suggestions regarding the reading order, however the essence of the work is preserved without the necessity of imposing it.

5.3.3.2 Multilayered critical apparatus

The flexible layers of critical apparatus are also used in the aforementioned edition of the correspondence between Polish poets of the Skamander literary group. The user

can choose to read the plain text version of each letter, or use the annotated version in which different types of entities (people, places, published texts, organisations, periodicals, lexicon) can be highlighted in the text. This also means that the entities are interlinked with the text, which allows users to browse the corpus either in chronological order or by choosing particular points of interest, like particular people, places, or works mentioned in the letters.

5.3.3.3 Maps

Another feature made widely available in digital editions (already present in paper editions but scarcely used due to their many limitations) is the possibility of presenting maps, both in static and dynamic forms. These can be applied not only to those texts that relate to traditionally understood geographical, geopolitical, and historical matters, but also in all those cases where the visualisation of spatial relations and interconnectedness (whether on a small/local or global scale) can enrich the understanding of a given subject.

A relevant example here is the Atlas of Holocaust Literature, hosted on the NPLP.PL platform, which makes use of maps to present the history of the Jewish ghetto in World War II Warsaw, based on the written testimonies/diaries of its inhabitants, both in its human and urban context.

Moreover, maps – apart from being a means of spatial representation – can also serve as an additional way of interconnecting and linking the various entities, elements, threads of published text/corpora of data, and access points within the edition.

5.3.3.4 Multilingual editions

Digital media also opens editions to international and multilingual dimensions**.** This seems to be of special interest in those cases where versions in different languages were created by the author, but obviously, it is also applicable to any text existing in more than one linguistic version. Instructive realisations of such multilingual editions are discussed in the examples below:

The Beckett Digital Manuscript Project. This is a comprehensive digital project, delivered via an inter-institutional collaboration, which gathers and enhances the manuscripts of Samuel Beckett's works. This edition brings together the digital facsimiles of texts and documents, that have been dispersed across various libraries. It makes genetic research – reconstructing the writer's creative process – possible thanks to transcriptions of Beckett's manuscripts, as well as providing tools for bilingual version comparison. It also delivers an analysis of the textual genesis of the writer's works. This is an important example due to its comprehensiveness and scope, the importance of the author for contemporary literature, and the exemplary bilingual output. However, it needs to be mentioned that the edition is not available free of charge.

The Corpus of Ioannes Dantiscus Texts and Correspondence is a TEI-based platform bringing together the diverse body of works by the Polish 16th-century poet, diplomat, and printer, Jan Dantyszek (Ioannes Dantiscus), which includes his poems, speeches, and memorials written in Latin and German, as well as letters composed in more than a dozen languages. The platform serves as a comprehensive archive of

texts whose manuscript originals and printed editions have been dispersed, and systemises knowledge about the corpus. It incorporates the features of a database and enables data to be migrated to and from other information systems, providing a critical apparatus, including commentary and indexes.

Bi- or multi-lingual digital editions, whose affordances greatly exceed those of the potential printed formats, can be of great use in relation to texts from different periods, especially in those cases when a writer is also the author of a translation of his own text into other languages, or writes two parallel versions of the same text in different languag (even if it can't be regarded as a translation *sensu stricto*), allowing for cross linguistic collation.

5.3.4 **Advantages of digital editions**

We can identify the following features of digital editions, which leverage the affordances of the medium in terms of availability, accessibility, and overcoming the limitations of book publication (linearity, limited content). These are:

- the publication of all possible, or selected, versions of the text, allowing for dynamic collation. Editors no longer have to choose the "definitive" version of the text whenever this is not their explicit intention;
- the publication of contextual materials that enrich the edition: maps, manuscripts, audio and graphic files, etc. Genetic research may also use notes, drafts, and other "pre-texts" evidencing the creative process;
- creating a network of hyperlinks interconnecting corresponding fragments in different versions with contextual material;
- flexible layering of critical apparatus (the option to display one or multiple overlapping layers of critical apparatus and commentary);
- support for multilingualism.

5.3.5 **Serial vs. individual solutions**

5.3.5.1 Two models of design in scholarly editing

We can broadly distinguish two models that are used in the design of services dedicated to digital editions by referring to the categories of "Haute Couture" and "Prêt-à-Porter" as defined in an insightful article by Elena Pierazzo about the correspondences between DSE and fashion design:

> The fashion industry clearly distinguishes between two lines of products: the Haute Couture, and the Prêt-à-Porter. The former is characterised by the fact that each piece is unique and is often created for one person only to wear for a special red-carpet occasion. Haute Couture can and indeed usually must be innovative and creative and has more to do with art and innovation than with the production of wearable items [...]. Prêt-à-Porter ("ready to wear") is the term used to refer to the class of items people can actually buy in shops and wear in their normal day-to-day lives. Prêt-à-Porter clothing comes in different

> sizes and colours, and it is normally worn by its owners more than once. These items may be inspired by Haute Couture, but they simplify it, making it accessible and wearable. [...] If we adapt this metaphor to digital scholarly editions, we notice that at the moment the editions we produce bear more affinities with Haute Couture than they do with Prêt-à-Porter: digital editions are typically unique. Each is provided with a set of dedicated tools, and each is innovative, creative, expensive, and specific to the text for which it was created, and it is not normally available to others to use. The challenge here is to imagine what a Prêt-à-Porter edition might look like, that is, to model the digital editions of the future and their editors, or, better, the skills they need to acquire.[87]

From this diagnosis, Pierazzo draws, among others, the following conclusions:

- the same infrastructure should be portable and reusable for many editions;
- the interface and layout of the published texts should be familiar and recognisable to users in order to avoid disorientation and to help further assessments of the scholarly value of the edition;
- it should create the option to "plug in" essential, easy to set-up, and reusable tools;
- the edition should be easy to create for the scholar, i.e., not requiring a knowledge of programming languages or setting up a web server;
- it should be open to certain customisation, providing a choice of some basic models.

Pierazzo also stresses the fact that in such cases, the digital should be perceived not so much as a field of research, but rather as a backbone for research.

Infrastructure designed in accordance with these suggestions, Pierazzo observes, would enable the spread of digital editions, help to consolidate its subsequent developments, and provide a more sustainable and durable environment for them. In order to achieve such a goal and assure the stability of future infrastructure for digital editions, she sees the necessity of putting effort into creating preliminary models (which, nonetheless, can be based on already existing infrastructure, and take into account already emerging trends in its design and use). Also, there is a need to reach an agreement on features that are essential from the point of view of the

[87] Pierazzo, Elena. 'What Future for Digital Scholarly Editions? From Haute Couture to Prêt-à-Porter'. *International Journal of Digital Humanities* 1, no. 2 (1 July 2019): 209–20. https://doi.org/10.1007/s42803-019-00019-3.

scholarly community (in relation to particular types of text or scholarly problems), which would help shape essential models.[88]

Despite the imperative of the infrastructural integration advocated by Pierazzo, her considerations make clear that it is impossible to arrive at a single, ultimate way of organising DSE. While there are many advantages to a single service for different editions (sustainability, updates, maintenance costs, lower entry threshold), it is also impossible to deny the vital importance of editions dedicated to particular works or corpora (https://melville.electroniclibrary.org/, http://www.faustedition.net/ – analysed as earlier in this report). It is important that editors designing their digital edition can clearly choose which model meets their requirements.

5.3.5.2 Standardisation: an urgent issue

The Haute Couture and Prêt-à-Porter models of DSE should be perceived as complementary. The importance of editions designed for particular cases, with specific functionalities catering to distinctive aims, does not diminish the general need for the wide adoption of common standards in the field of DSE, and digital humanities in general, which becomes ever clearer.

The existing state of fragmentation sets limits on the retrieval, accessibility, and reusability of scientific resources and information, as well as constraining the possibilities for communication and cooperation between scholarly communities. The current diversity of formats, publication versions, content types, workflows, and operational models give rise to an increased necessity for the implementation of global standards that would set a common framework for scholarly communities operating in a digital environment.

Standardization – perhaps the most significant, current challenge for digital humanities – requires a thoughtful, gradual course of action. It must be a process that includes a wide range of tasks, such as the identification of needs (that includes discipline-specific standards, already developed practices, and predefined formats); deliberation on, and adoption of, shared principles; and the implementation and promotion of newly introduced common practices and standards. Only such a pre-agreed general framework of rules, models, and practices can guarantee the reinforcement of openness and availability, and interoperability and interconnectivity of scholarly output and its further processability (content reuse, meta-search possibilities, long-term preservation).

5.3.5.3 Advantages of standardisation in DSE

There are already a number of foreseeable, concrete advantages that should arise from establishing unified editorial and operational models.

Once we decide to use a certain standard, for example TEI (Text Encoding Initiative), we will be able to juxtapose and collate different texts, and their structure will be readable by anyone who is already familiar with the TEI format.

[88] Pierazzo, Elena. 'What Future for Digital Scholarly Editions? From Haute Couture to Prêt-à-Porter'. *International Journal of Digital Humanities* 1, no. 2 (1 July 2019): 209–20. https://doi.org/10.1007/s42803-019-00019-3.

The possibility of re-using data is not purely a theoretical one, as can be demonstrated by the example of the DraCor platform, which has collected together more than ten extensive corpora of dramas written in different epochs and languages. Each corpus was created in a different project, but their common feature – use of TEI Drama standard – made it possible to gather them into one collection.

The set of annotations and software available on the site open further innovative research facilities, such as collating and juxtaposing manifold statistics, charts, and graphs for each of the more than one thousand dramas processed beforehand.

Another important advantage of a platform that collects different editions is the possibility for creating a structured base of commentaries and descriptions, which can be applied to many (also forthcoming) editions and contribute to their mutual enrichment. They can be interconnected, thus also enabling the integration of resources. Standardisation (e.g. TEI) also provides the ability for data sharing and making it FAIR.

5.3.6 Needs concerning the scope and functionalities of a future DSE service

On the basis of a literature review and consultations with editors, we prepared the following list of requirements that should be met by a future DSE service.

5.3.6.1 Scope:

- **Wide variety of texts**. It should serve to edit a wide variety of types of texts, e.g., different forms and genres;
- **Diverse commentary models.** Enable and make use of diverse models for annotations and editorial commentary;
- **Different historical periods.** Be open to include texts from different historical periods and provide facilities that answer the specific challenges and problems of a given text and its cultural and historical context;
- **Possibilities for customisation**. Provide possibilities for customisation that would allow issues specific to a particular work to be handled, such as, for example, the mark up of censored texts.

5.3.6.2 Essential functions need to enable the following operations:

- **Collation/comparing versions of text**. Features permitting the collating and juxtaposing of different text versions allow for comparative reading with a high level of detail. As previously mentioned, this function is of particular importance in cases where texts do not have an agreed upon, definite version.
- **Tracing and presenting the genesis of a text.** Drafts, notes, diaries, manuscripts, typescripts, and other materials should be

interconnected with the work to allow for analysis of the creative process.

- **Adding visual/multimedia context**. Visual additions such as pictures, films, scans, charts, and graphs should allow archival materials such as manuscripts and related documents to be presented, which can then create a context for the edited work. Hyperlinked materials not only refine and widen traditional models, but also open a wide range of possibilities for creating entirely new narrative models, methods of text presentation, and modes of lecture, unattainable in print editions. Visual elements enriching the user's experience have additional significance for societal impact. Another area highly dependent on visuals is the quantification of textual elements, which benefits from such tools as graphs, diagrams, charts, figures, etc.
- **Maps and geospatial tagging**. Digital editing opens new ground in all the areas of the humanities where geospatial relations are particularly important. It can also support applications in research areas influenced by cultural geography and the so-called “spatial turn,” for example in geopoetics. Such tools as static or dynamic, multilayered maps and geo-tagging make it possible to visualise a network of interrelated entities (place/time/person/event), highlighting the social factors of the edited work.
- **Text explanation features** must ensure the possibility of creating interlinked comments and annotations that will provide readers with all necessary explanation and information so that they have a deeper understanding of the text and its background. This task requires implementing commentary at different levels, such as:
    - **Context explanation**, i.e., the historical, cultural, and social background of the text; the environment and conditions in which it came into existence and its first publication; and social impact. Emphasis should be given to elements of the original context, which may be unclear for contemporary readers;
    - **Lexical explanation**, i.e., clarification of set phrases and idiomatic expressions, figures of speech, rhetorical figures, and other broader semantic units, which, due to their historical or local cultural and linguistic context, may be incomprehensible today;
    - **Named entities tagging**, identifying and interconnecting people, places, institutions, events etc., that appear in the text;
    - **Intertextuality**: establishing the interrelationships between the presented text and other cultural texts that may broaden the readers’ understanding.

5.3.7 **Problems and challenges**

There are a number of problems and challenges that have to be taken into account from the outset:

Problem	Possible solution	Comment
High financial threshold for entry	Creating free tools for DSE	Creating a DSE usually requires considerable financial resources and digital humanities expertise (see Pierazzo 2019). We should seek to minimise this problem by offering free access to infrastructure/tools.
Lack of general understanding of infrastructure, workflows, and business models in undertaking DSE projects.	A combination of DSE infrastructure and staff with relevant expertise.	Many scholars who are willing to get into digital editing may find it difficult to gain a general orientation in the field, not knowing its possibilities, advantages and limitations, organisational and practical intricacies, or proceeding modalities. These obstacles cannot be overcome by means of infrastructure alone. The human factor is equally important: access to expertise on how to choose a digital editing model and how to use the tools available.
Cooperation with “traditional” editors who have low digital skills.		There is, obviously, an imbalance in the level of digital skills among academics. Digital editing should also work towards levelling “digital literacy,” but cannot abandon the knowledge and theoretical competence of “traditional” editors.
A variety of display/analysis tools.	Creating a detailed catalogue of the tools that are	

The variety of digital editing tools currently in place	available at any given time on various platforms. Developing complementary tools across different platforms.	Such a dispersion of tools and mark-up standards – though perfectly natural in a relatively new and dynamically evolving digital field of research – poses problems for both newcomers to digital editing and users, which also limits the possibilities of cumulative learning. Developing common standards, then, should be perceived as part of good practice and a forward looking strategy.
A variety of mark-up standards.	Precise agreement on common standards, and adapting tools accordingly.	
Cost of maintenance and problems with sustainability.	The integration of tools into a single research infrastructure, coordinated by all partners and with higher possibilities of receiving governmental support and funding.	As for the case of SSH Open Marketplace, "a discovery portal which pools and contextualises resources for Social Sciences and Humanities research communities: tools, services, training materials, datasets and workflows,"[89] it seems that the integration of such resources would be beneficial for the research cycle in SSH. Similarly, there is the possibility to enhance this with more DSE-specialised entities.

5.3.8 **Operating models**

[89] https://marketplace.sshopencloud.eu/about

Another issue of great importance is the organisational and business model, especially in the case of scholarly enterprises with broader time horizons and a scope exceeding the edition of a single text corpus.

Here are some examples of teams and institutions, each with differently projected *modus operandi*, but with a sustained commitment to DSE:

- **TEI.NPLP.PL**: the IBL PAN platform, with an NPLP team of (mostly) literary researchers, who also perform an intermediary function between the team of editors who are interested in digital publishing (having various levels of digital skills, though mostly newcomers), software programmers, graphic designers, and possible subcontractors. All editions are hosted on the same platform and use the same tools on both the backend and the frontend – although the way the tools are used is specific to each edition (new functionalities are sometimes created for specific editions, but must be compatible with the overall editing system and its performance on the platform).
- **Muruca**: developed by NET7. A team of developers, engineers, production editors, designers, and publishers from various institutions, working with a host of technological partners. Muruca is aimed at selling software to scholarly institutions, offering a broad range of SSH-focused services dedicated, in the first place, to digital scholarly editors, but also cultural heritage institutions. Each edition (Wittgensteinsource, Nietzschesource, Burckhardtsource, Galassia Ariosto, and many others) is hosted on a different site, although it uses the same software, customised for specific editions in collaboration with a research partner.
- **OpenEdition**: an internet publisher, including for the DSE sector, providing a digital infrastructure for academic communication in the humanities and social sciences, offering four complementary platforms designed for journals (OpenEdition Journals), book series (OpenEdition Books), research blogs (Hypotheses), and academic events (Calenda). Its activity is directed towards the development of open access digital publishing, dissemination of digital publishing related skills (open laboratory programme), and the development of new approaches to text and data mining. All content is on the same platform and uses the standards adopted for it.
- **ILIESI** (The Institute for the European Intellectual Lexicon and the History of Ideas): A very large number of digital resources of very high content quality, but created over the years in different standards and with the use of different tools – as a result, there are very different tools used to present different resources on the one site/platform.

It should be clear that, as with the previous criteria, it is impossible to identify a single appropriate organisational model for DSE undertakings. While a single platform

for different editions may be the most cost-effective solution, and platforms managed by academic institutions may offer advice from the researchers they employ, many users would like to have their edition hosted autonomously, or have specific needs regarding functionalities. Moreover, many institutions funding digital editions are reluctant to see their project hosted on a platform managed by other institutions.

5.3.9 **Conclusions**

It seems that a future infrastructure for digital scholarly editions should go far beyond the tool aspect by providing advisory services for choosing the best solutions for particular users in order to find the best fit for the financial possibilities of the project, but also with the assumed aims of the edition as well as other factors (team size etc.).

A consistent tagging, data, metadata system should be established and implemented across all platforms, and a catalogue of tools that all partners have available should be compiled.

It also seems that it would be a mistake to abandon previously developed editorial infrastructures and tools – the planned infrastructure should strive to integrate these into a system of tools/infrastructures, additionally, extending those functionalities that the existing infrastructures do not have.

Nothing prevents the creation of new tools that are not currently available to any of the partners, or proposing approaches that have never been applied before (e.g. DSE as data).

Bibliography

- Franzini, Greta, Melissa Terras, and Simon Mahony. 'Digital Editions of Text: Surveying User Requirements in the Digital Humanities'. Journal on Computing and Cultural Heritage 12, no. 1 (13 February 2019): 1:1–1:23. https://doi.org/10.1145/3230671.
- Irakleitos Souyioultzoglou, Marina Angelaki, Eelco Ferwerda, Paola Galimberti, Arnaud Gingold, Pierre Mounier, Ronald Snijder. (2018, July 30). "OPERAS Common Standards White Paper", July 30, 2018, https://doi.org/10.5281/zenodo.1324066.
- McGann, Jerome J. 1992. A Critique of Modern Textual Criticism, Foreword by David C Greetham. Revised ed. edition. Charlottesville, Va: University of Virginia Press.
- Sharing knowledge digitally; the Muruca case study (Chiara Aiola, Giulio Andreini, Francesca Di Donato, Tiziana Lombardo)
- Pierazzo, Elena. 'What Future for Digital Scholarly Editions? From Haute Couture to Prêt-à-Porter'. International Journal of Digital Humanities 1, no. 2 (1 July 2019): 209–20. https://doi.org/10.1007/s42803-019-00019-3.
- Tanselle, G. Thomas. 'The Editorial Problem of Final Authorial Intention'. Studies in Bibliography 29 (1976): 167–211.

Annexes

Annex 1: Codes and demographic information concerning the interviewees

Code	Gender	Career stage	Discipline / area of research:	Country
OP01	Male	Post-doc/ECR	Cultural studies, gaming studies	Hungary
OP02	Female	Post-doc/ECR	Linguistics	France/ Germany
OP03	Male	Post-doc/ECR	Literature, scholarly communication, DH	UK
OP04	Female	Senior researcher	English studies	USA
OP05	Male	Senior researcher	Digital humanities	Hungary
OP06	Female	PhD student	Digital humanities	Ireland
OP07	Female	PhD student	Digital humanities	Ireland
OP08	Male	Senior Researcher	Digital humanities	The Netherlands
OP09	Female	Publisher	Arts, humanities, media	UK
OP10	Male	Post-doc/ECR	Information studies	UK
OP11	Female	Post-doc/ECR	Communication, information science	Poland
OP12	Female	Post-doc/ECR	Sociology	Poland
OP13	Male	Post-doc/ECR	Philosophy	Poland

OP14	Male	Senior Researcher	Cultural studies, literary anthropology, cognitive semantics	Poland
OP15	Male	Senior Researcher	Information; digital humanities; digital heritage	Canada
OP16	Male	Post-doc/ECR	Science studies	Czech Republic
OP17	Male	PhD Student	Information and communication science	France
OP18	Female	Senior Researcher	Narrative studies, digital humanities	Latvia
OP19	Female	Senior Researcher	Cultural memory studies	Bulgaria
OP20	Male	Senior Researcher	Early modern history, regional history	Germany
OP21	Female	Post-doc/ECR	Biblical studies, digital humanities	Switzerland
OP22	Male	PhD Student	History	Switzerland
OP23	2 x Female	2 x Other	Religious studies / global history	Germany
OP24	Male	Post-doc/ECR	Biblical studies and digital humanities	USA
OP25	Male	Senior Researcher	History	France
OP26	Female	Publisher	SSH	The Netherlands
OP27	Male	Senior Researcher	History	Luxembourg
OP28	Female	Post-doc/ECR	History	Luxembourg
OP29	Male	Post-doc/ECR	Psychology, statistics	Croatia

OP30	Male	Senior Researcher	Electronic Systems and information processing, education	Croatia
OP31	Male	Senior Researcher	Sociology	Croatia
OP32	Male	Senior Researcher	Information and communication science	Croatia

Annex 2 : Interview questionnaire

1. Episodic knowledge

1.1. Writing

1.1.1. Which digital technologies do you use when writing academic texts? Please, describe the process from the idea formation to the final draft and provide some examples of tools.

1.1.2. Do you use any digital tools enabling workflow planning and monitoring? Give examples.

1.1.3. For how long have you been using these tools?

1.1.3.1. How did you learn to use them?

1.2. Publishing

1.2.1. What types of scholarly outputs have you published in the course of the past two years? (It could be in the form of a journal article, book or book chapter, paper in an edited volume or conference proceedings, SSH blogs/platforms, data source, software, multimedia)

1.2.2. Which ones did you write individually and which ones were co-written in collaboration?

1.2.3. Now I would like you to choose one output which you'd find most interesting for our discussion in terms of form (If applicable, it could be an example of an innovative genre of scholarly communication, like blog post, project website, multimedia scholarly edition, social-media post, etc.). We will talk about this output in more detail.

The interviewee chooses one output for further discussion.

1.2.3.1. Why did you choose this particular form for this output?

1.2.3.2. What were the main challenges in finding the appropriate publisher or publication channel?

1.2.3.3. When choosing a publishing venue are you attentive to bibliometrics (h-index, open/new metrics)?

1.2.3.4. What are the advantages and disadvantages of using that form?

1.2.3.5. How long did the publication process take from the moment of finishing the draft?

1.2.3.6. Which parts of the publication process were the most time-consuming?

1.2.3.7. What could have been done more efficiently? How?

If this subject didn't surface in the responses so far, we ask about the cooperation with the publisher/editors and the reviewing process in case of the chosen output (1.2.4, 1.2.5).

1.2.4. Please, tell us more about the cooperation with your publisher/editors?

1.2.4.1. What could be done more efficiently? How?

1.2.5. What did the reviewing process look like?

1.2.5.1. How long the peer review period took?

1.2.5.2. Who made decisions about approving your output for publications?

1.2.5.3. Were you involved in finding/referring reviewers to your submission? Did you have any difficulties?

1.2.5.4. Did you submit this output for other kinds of evaluation i.e. formal or informal feedback, comments or review? How did that process look? Why did you choose this form of evaluation?

1.2.5.5. Did your work benefit from this process and/or peer-review?

1.2.5.6. What could have been done more efficiently? How?

1.2.6. Do you perceive any difference between papers which resulted from collaboration as opposed to individual pieces?

Additional questions for respondents who indicated experience with collaborative writing.

1.2.7. Have you used any digital tools enabling collaborative writing?

1.2.7.1. For how long have you been using them?

1.2.7.2. How did you learn about them?

1.2.8. Who was in charge of the writing process? Did you have a "leading" author or was it an equal collaboration?

1.3. Evaluating

1.3.1. When discovering new scholarship or doing literature review, how do you make decisions about trustability and quality?

1.3.1.1. Do you trust work more if it has been peer reviewed? Why?

1.3.2. When you receive an invitation to peer review a scholarly object, what are the circumstances that help you to decide whether to accept or decline?

1.3.2.1. Have you ever been rewarded for the reviewing activity?

1.3.2.2. What would increase your motivation to peer-review?

1.4. Communicating

We are still discussing the chosen output.

1.4.1. Is the output you have chosen for this discussion available openly online?

1.4.1.1. If yes, where and why did you choose this dissemination venue? [*Prompt: possible options may include: OA journal, repository, website (institutional, private, publisher's), scholarly social networks (Academia.edu, Researchgate)*]

1.4.1.2. Are you a part of any online group or network for researchers?

1.4.2. What other methods of communication about this research did you use? Here we ask about communicating your outputs through diverse forms in various phases of the research process.

1.4.2.1. Did anyone help you with that?

1.4.3. Which communication channels are useful and appropriate for communicating with the audience in SSH?

1.4.4. Do you see a need for changes in the field of communication about scientific papers in SSH? What could be improved from your perspective?

2. Semantic knowledge

2.1. Traditional and innovative forms and genres

2.1.1. When you hear “scholarly text” what comes to your mind?

2.1.1.1. What role do other materials (data, images, software, etc.) play in the process of writing and publishing?

2.1.1.2. What is your opinion about publishing the entire material from a given study in SSH (E.g. whole interviews, annotated texts, annotation schemas, corpora, research protocols, data collected in the research process etc.)?

2.1.1.3. Do existing metadata schemes cover the needs of scholarly writing and integrating various metadata: publication metadata, research data metadata, non-textual content metadata?

2.1.2. What is innovation in scholarly communication?

2.1.2.1. Which innovative genres and formats of scholarly communication are you familiar with? (e.g. website, software, blog, social media posts, etc)

2.1.2.2. What are the advantages and disadvantages of using innovative genres and formats?

2.1.2.3. Which innovative forms and genres of scholarly communication are the most useful for SSH researchers?

2.1.2.4. What is the audience of innovative forms and genres? Does it overlap with the audience of more traditional forms of scholarly communication?

2.1.2.5. How to assess the impact of new genres in comparison to bibliometric impact factor of traditional communication? E.g. alt-metrics (number of downloads, mentions in social media).

2.1.3. What are the difficulties of evaluating innovative genres in existing peer review processes?

2.1.3.1. Should all innovative genres be peer-reviewed? Are there any differences?

2.1.3.2. What do you think about citing new writing forms when writing an academic publication?

2.2. Prestige

2.2.1. Do you think there are publication types that count more in your career assessment/academic profile than others? What are these and why?

2.2.2. Could writing for the non-scholarly audience be a source of academic prestige?

2.2.3. What are the elements that make up the prestige of a publication?

2.2.4. Are some innovative forms and genres of scholarly communication considered to be more prestigious than others?

2.2.5. What do you think about the prestige of OA publications?

2.3. Power structures

2.3.1. Which actors have currently the strongest influence on publishing: policy-makers, funders, research institutions, publishers, early career researchers, senior researchers. Why?

2.3.2. Who are gatekeepers in scholarly communications? Reviewers, editors, editor in chief? (Can the editor-in-chief make a decision contrary to the reviews?)

2.3.3. Do you think that peer review is effectively conducted by the best experts? Are they rather early career researchers or senior staff?

2.3.4. Are early career researchers and (or) scholars with no stable employment more vulnerable when engaging in innovative forms of scholarship (open data sharing, preprints sharing, open peer review)?

2.4. Peer-review

2.4.1. What are the main functions of peer review? (gatekeeping, improving scholarly work, filtering?)

2.4.1.1. Should the peer review be published together with the paper?

2.4.1.2. Is peer review always organised by journals and publishers or do you know any other forms of peer review that happens outside of the traditional publication workflow? Do you know any other practices of assessing and improving the output? (e.g. commenting on drafts, or code, participation in recommendation networks or twitter discussions etc.)

2.4.2. Have you ever heard about open peer review? Openness in peer preview can take many forms from open interaction between the reviewers, through publishing review reports anonymously or signing reports and openly publishing them.

2.4.2.1. What do you think about it?

2.4.2.2. Have you ever taken part in it as an author or reviewer? If so, what were your impressions?

2.4.3. In your opinion, is it easy or difficult to find researchers willing to participate in innovative peer-review practices? E.g. open-peer review, transferable (or portable) peer review, post publication review, cascade peer-review, open identity review.

2.5. **Publishing costs**

2.5.1. How does a budget influence or limit the horizon of choice with regards to publishing venue or format? What would a scholar with low-budget do?

2.5.1.1. Is there any institutional or national support for covering the publication fees.

2.5.1.2. Is it possible to obtain financial support from beyond academia, like crowdfunding?

3. **Synthesis**

3.1. What are the tools or services that you really miss from the current landscape, that would make the publication process much easier? (Here you can think of anything from writing, collaboration tools through services finding publishing venues for your work to post-publication or dissemination tools/services.)

3.2. What is the most important thing that should be changed in order to improve the current scholarly communication system?

3.3. Do you know any innovative publishing projects that we should examine?

Annex 3: List of case studies

- C1 New Panorama of the Polish Literature
- C2 Manifold Scholarship
  - C2A Metagaming. Playing, Competing, Spectating, Cheating, Trading, Making, and Breaking Videogames
  - C2B The Lab Book Situated Practices in Media Studies
- C3 Stanford University Press
  - C3A The Chinese Deathscape. Grave Reform in Modern China
  - C3B Black Quotidian: Everyday History in African-American Newspapers
  - C3C Constructing the Sacred : Visibility and Ritual Landscape at the Egyptian Necropolis of Saqqara
- C4 Aristotle's Topics / TOPOI
- C5 MELVILLE ELECTRONIC LIBRARY. A critical archive
- C6 Faust. Historisch-kritische Edition
- C7 The Pulter Project. Poet in the making

- C8 Living Books about History
- C9 PolEval
- C10 Platform for Responsible Editorial Policies (PREP)
- C11 NextBook
- C12 The Registry of Open Access Repository Mandates and Policies (ROARMAP)
- C13 Polona
- C14 POL-on. The Integrated System of Information on Science and Higher Education
- C15 PubPub
- C16 Open Library of Humanities
- C17 Overleaf: LaTeX Editor
- C18 Rebus Ink
- C19 Publons
- C20 ScienceOpen
- C21 ScholarLed
- C22 European Open Science Cloud (EOSC)
- C23 Language Science Press
- C24 Hypotheses
- C25 Scalar
- C26 Kudos Pro
- C27 GITenberg
- C28 Refereed
- C29 Protocols.io
- C30 Mukurtu
- C31 Muruca
- C32 Octopus
- C33 Editoria
- C34 Pundit
- C35 Authorea
- C36 Peerage of Science
- C37 #DHgoesViral "DH in the Time of Virus"
- C38 Peeriodicals
- C39 Jupyter
- C40 Fidus Writer
- C41 Recogito
- C42 HIRMEOS - High Integration of Research Monographs in the European Open Science infrastructure
- C43 dokieli
- C44 Electric Book Works
- C45 Janeway
- C46 From The Page
- C47 Omeka
- C48 SSH Open Marketplace
- C49 Ports, Past and Present: Cultural Crossings between Ireland and Wales
- C50 PubPeer
- C51 after.video
- C52 Gamer Theory (Future of The Book)
- C53 The disrupted Journal of Media Practice
- C54 FairCopy Editor

- C55 Content Mine
- C56 docloop