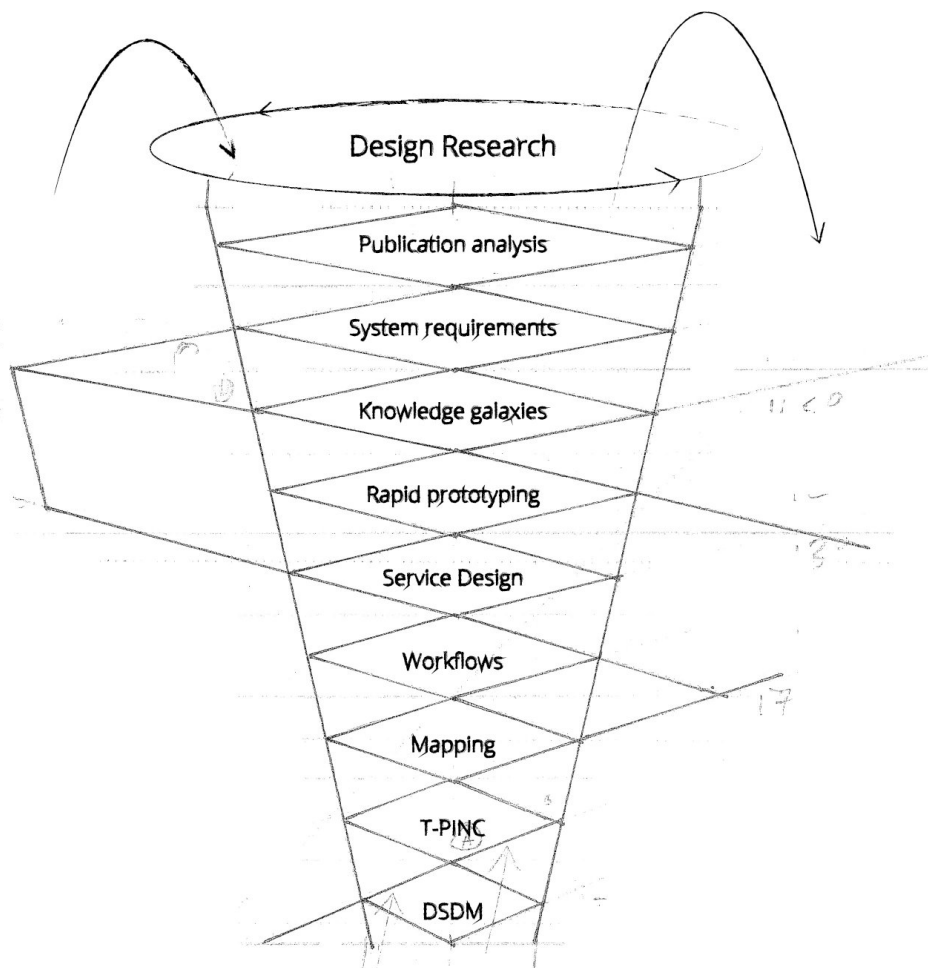# Dynamic Publishing

- *New Platforms, New Readers!*

Hybrid Publishing Consortium Research Plan.
Open source infrastructure for publishing.
August 2013, v1.1

Table of contents

Document link: http://consortium.io/research-plan

Contact: Simon Worthington, simon.worthington@inkubator.leuphana.de
–
Attribution: Author, Simon Worthington. Contributions from Minuette Le and Christina Kral on Design Research, workflows and consultations and Agata Królikowski on TPINK and system requirements. Editing: Anthony Iles, Post-Media Lab.
Thanks to Pauline van Mourik Broekman.

http://consortium.io/
http://cdc.leuphana.com/structure/hybrid-publishing-lab/
http://hybridpublishing.org/

1. Executive summary

Open infrastructures for publishing

- Dynamic publishing offers publishers one of the few routes to audience growth and financial sustainability in an age of digital disruption.

Dynamic publishing is a digital workflow where the publication processes are automated (layout, multi-format conversion, distribution, rights clearance, translation workflows and payments) and made available on request for reuse, to give access to new audiences and revenues. Currently, the majority of publishers are excluded from the dynamic publishing arena. The software systems on offer are either too expensive on the high-end of the product range or not 'fit for purpose' on the low-end.

- The objective of the Hybrid Publishing Consortium (HPC) is to significantly lower the price of professional digital publishing systems through the creation of open infrastructures for publishing based on open source software and open standards.

The infrastructures would be supported as a variety of industry partners, open source communities and research groups. HPC's contribution to this network is to connect the scholarly publishing community with the long standing work on open standards in digital publishing by the industry and the open source community.

Free at the point of reading!

The HPC supports open intellectual property rights (IPR), for reading and learning, not just in academia but for publishing in general. At the same time HPC also understands the need to economically support the skills authors and the publisher lends to the process. We see dynamic publishing playing a key role in a reconfiguration of publishing and knowledge institutions–such as the university–as it allows easy direct relationships with the readers, cutting out costly intermediaries. In a near-future where content is available for free, via file sharing, economic models that move payment away from the consumer and up the chain become essential–to the network provider, funder, collection agency or via some other route.

Bold policy moves are needed to accompany the development of the related technology. The first being to support technology infrastructures as part of the public domain–a part that can then support–independent enterprise, publishing, research and knowledge institutions.

2. Mission

Single source

In the consortium our focus is on single source publishing. Single source is the structuring of documents to make them machine readable and available for multi-format conversion. Over 2013 we will be releasing a tool set, called Typesetr, for single source conversion. We'll be running a series of rapid prototyping projects exploring how to best fit Typesetr into workflows, aiming to make the publishers lives a little more sane and enjoyable.

Publishing

Reliable figures on the size of the publishing industry are murky and hard to come by, but estimates put the industry size at in the EU at 64,000 companies, employing 750,000 people and valued at 0.5% of the EU GDP, which is currently 12.894€ trillion[1]. This estimate puts the value of the EU publishing industry alone at revenues at €640 billion per annum. But this lucrative industry is skewed, with approximately 20% of the companies taking 73% of the profits.

The scholarly and independent publishers that the consortium is focused on sit within the less profitable majority group of publishers, the 80%. When it comes to these publishers ability to experiment with new forms of digital publishing they are held back by limited budgets and similarly they are not supplied with services by the technology industry because of the unprofitable nature of providing services to a cash strapped industry.

The consortium has a two roles role to play in bridging this technology gap. First we can carrying out research into models of dynamic publishing that an individual small publisher could never afford. Secondly, by providing innovations in low cost technology infrastructure, that is free for all to use, we can make it affordable for tech and design companies to provide services to small publishers.

Dynamic publishing
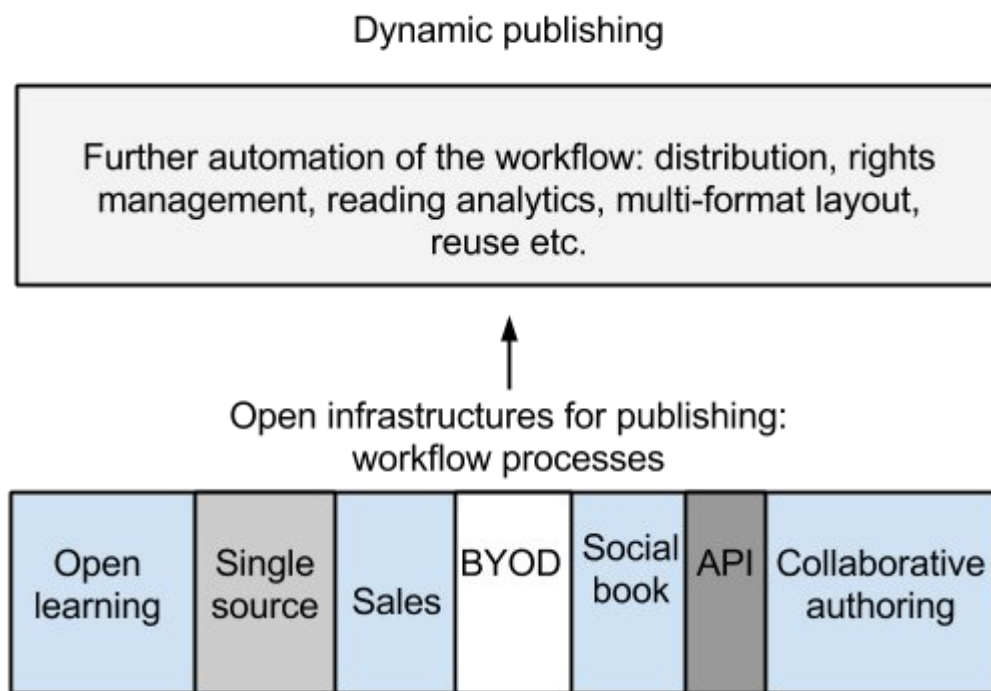
Figure: Dynamic Publishing. How the automation of workflow processes leads to dynamic publishing. Although this automation can only ever be partial.

Dynamic publishing means root and branch changes for publishing. The places for the availability of publication suddenly and dramatically increase, with new reading platforms and distribution channels. The need for new types of contracts and permissions for re-publishing to be as speedy as the flow of books around digital networks and providers. With the form of the

---

1   Eurostat http://epp.eurostat.ec.europa.eu/tgm/refreshTableAction.do?
     tab=table&plugin=1&init=1&pcode=tec00001&language=en

book broken up, from being a static bound object to becoming a set of 'content components' available for reuse, ways of authoring also need to adapt. Introducing these new dynamic publishing workflows into the publishers production environment is our core challenge.
The consortium will be exploring dynamic publishing models for new forms of scholarly dissemination with a series of rapid-prototyping exercises. For example in following the example of the US educational publishers Flat World Knowledge[2], who provide text books as free-to-read online and at 80% price reductions to institutions on 'bulk bring your own device' (BYOD) sales. Another example to emulate is how Pearson[3] distributes publications via an API, which essentially means publishing on request. With an application program interface (API), if another publishers wants to use Pearson content, the publishers systems make an automatic request and the publication is supplied on-the-fly, instantly. Previously such a re-publishing request would have gone through lengthy contractual arrangements, which may have taken several months.

Open infrastructures for publishing

In the networked society there is an ongoing battle between private and public infrastructures. Unsurprisingly the consortium is in the public infrastructures camp. In the field of publishing we are researching dynamic publishing and how to make high quality technology public domain, allowing for open innovation, with reasonably priced service provision to the professional user. Many parts of the infrastructure have existed for a long time, provided by a wide spectrum of players from companies like IBM, research bodies such as the Hasso-Plattner-Institut and media activist service providers like DocumentCloud. The consortium's role is to speed along the process of addressing dynamic publishing by coordinating the existing projects into coherent and reliable workflows that publishers can trust.

Policy and technology

Publishing is in the midst of such fundamental changes due to digital disruption that the provision of infrastructural technologies to lower barriers to market entry and streamline production costs will not be enough alone. Accompanying policy developments are also needed in many areas. Significant moves have been made in the area of open licensing, for general users Creative Commons and in academia with Open Access (OA) publishing. A game changing move in policy is the EU main research fund for 2014-20, Horizon 2020[4], gives a directive on STEM[5] research publishing to becoming Open Access as the norm.
The areas that the consortium has identified for further policy reform are: IPR on content permissions, piracy (the need to differentiate personal use and commercial exploitation), pricing monopolies, new economic models that move payment away from the consumer, public service broadcast and media infrastructure provision. The consortium will be actively looking to see how its technology work can be connected to other policy researchers working in this area.

Values and objectives

- Support universal access to reading, learning and publishing.
- Re-imagine publishing - knowledge institutions, authoring and learning.

---

[2] Flat World Knowledge, open IRP commercial academic publisher
http://catalog.flatworldknowledge.com/
[3] Pearson publishing API http://developer.pearson.com/
[4] Horizon 2020 http://ec.europa.eu/research/horizon2020/index_en.cfm
[5] STEM fields–Science, Technology, Engineering, and Mathematics.

- Formulate efficient and low cost multi-format publishing workflows, including; rich media, social media, transmedia publishing and Open Learning, Open Education Resources (OER)/Massive Open Online Course (MOOC)+[6]
- Widen open IPR approaches beyond academia and to the independent publishing sector. With 'free at the point of reading' economic models of publishing.
- Using a mixture of dynamic publishing and open IPR models, enable vibrant economic models for critical and independent publishing.
- Address quality assurance in software design through design research processes.
- Advocate Open Access and Open Education Resources.
- Support the free culture movement of free/open software, open IPR, open standards, free/open hardware and the open web.
- Protect personal privacy and intrusive monitoring of scholars, researchers, students and readers.

Milestones

- Consortium to found key partnerships for publishing infrastructure building
- Launch the open research platform.
- Single source publishing software release (HPC project) - TypeSetr Open Source software release - the consortium's single source conversion platform for multi-platform publishing.
- Single source development and support infrastructure (HPC project) - create a community around the TypeSetr product and technology to develop single source technologies further.
- Prototype dynamic publishing - research example of dynamic publishing using the consortium's publishing infrastructure. Creating partnerships, case studies and rapid prototyping examples of our own.
- Facilitate a functioning consortium economy to support ongoing long term research on the publishing infrastructure.

---

[6]   MOOC+ refers to going beyond the MOOC. The reason to go beyond the MOOC is that it is a flawed model driven by the economic needs to institution. What is missed in the MOOC model is the potential of digital media for learning and reconfiguring of institutions of learning. See the book 'We´re all Game Changers Now!', http://linkme2.net/u1
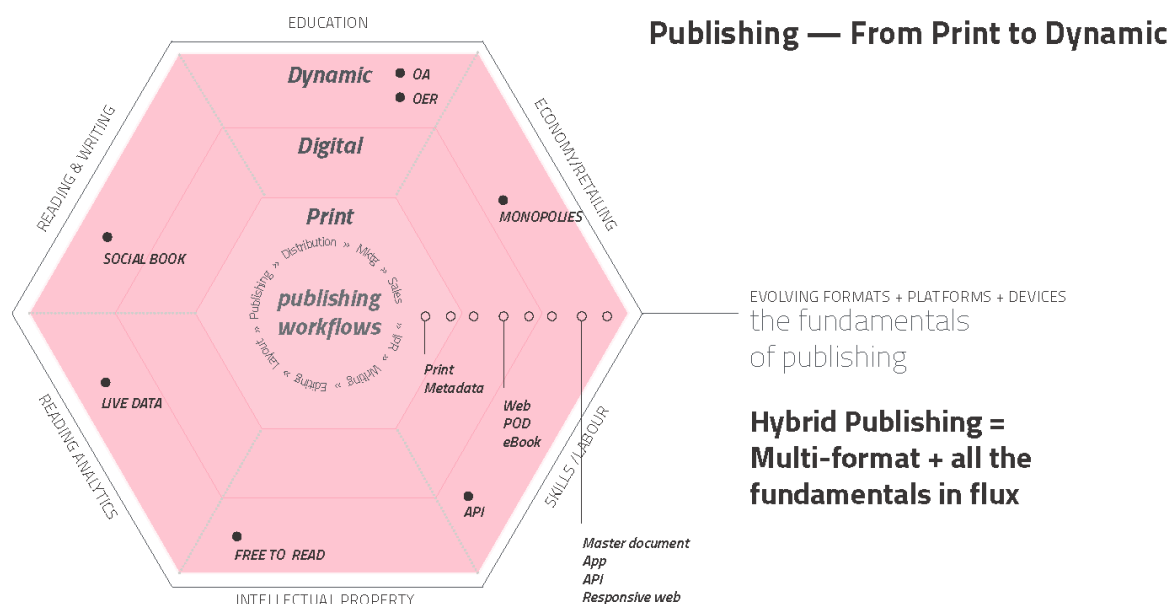
2.1. Defining Hybrid Publishing

Figure: Hybrid Publishing - From print to digital and onto dynamic publishing

Hybrid Publishing is where the book is broken apart, a move beyond the habitual mode of imagining digital publishing, which has consisted of copying the bound book in digital form. Hybrid publishing also reflects the challenges our age is facing–austerity, neo-liberalisation and digital disruption.

Hybrid publishing is the further innovation of the unbound book and the subsequent reorganisation of knowledge institutions, with an interesting twist that the cost of accessing knowledge lowers, so that the dreams of universal education can take one step forward. The development of the conventional book was always accompanied by parallel experiments with an unbound book form. In the 2011 book 'Paper Machines'[7] media historian Markus Krajewski traces a history of the European unbound book beginning as library records in the 16C created by a Swiss librarian Konrad Gessner, to the US Dewey Decimal System of the 19C and how the library record keeping system transferred to businesses in the early 20C as the card index system. This helps in part to define Hybrid Publishing. To gain full definition of Hybrid Publishing we need to add the agent of digital-disruption or as described by the economist Joseph Schumpeter, creative destruction[8]. Creative destruction is a process where a new economy emerges out of the destruction of a previous order. Technology innovation is the agent of this change, and Schumpeter describes the entrepreneur as the one who exploits this change.

Ironically it is the unbound book and its prodigy, the card index system, that led to the punch card, the early data packet of what are now packet networks. The packet network is where all media can be broken down into a common data packets and sent to any device. It is this technology that acts as the agent of creative destruction, that makes up basic internet and mobile networks, and have made the concept of the unbound book finally realisable.

---

7   Paper Machines About Cards & Catalogs, 1548-1929. 2011 by Markus Krajewski, Bauhaus University, Weimar. http://mitpress.mit.edu/books/paper-machines | http://linkme2.net/tu

8   Capitalism, Socialism and Democracy. 1942 by Joseph Schumpeter. http://www.scribd.com/doc/32097500/Capitalism-Socialism-and-Democracy-Copy

It is the scaling of this effect of packet networks which means that the fundamentals of publishing are in flux. The innovation of the unbound book and the replacement of print books means that economic models crumble, institutions of knowledge lose their relevance, copyright law becomes unenforceable. Readership data changes from being a bi-annual sales report to 'reader analytics' where a publisher can know what people read down to a granular level, following single readers to know which chapter they are reading at a given time and place. How people read and write changes. A re-skilling of an industry quickly becomes necessary.

Hybrid publishing innovation

The innovation process of digital-disruption brings a careful examination of the institutions and conventions of publishing, showing not only how new technologies can make improvements, but also how it exposes inherent failings in the status quo. Take for example Open Access academic publishing. The main barrier here is not technological, it is instead self interest, the vested interests of academics and publishers being resistant to the universal access agenda. Recently, what has been proclaimed as innovation within online educational publishing such as MOOCs, is rather an exemplification of how hard it is to innovate. Most MOOCs are simply replicated conventional educational formats that make the material available online, such as distributed textbooks and embedded videos. And it adds an automated student grading system (factory style) to the mix. There is neither rethinking nor innovation of new types education or forms of learning.

With this in mind, the consortium would advocate more exploratory approaches that use digital innovations and the capacities of information and communication technology (ICT) to rethink learning, education and publishing.

Fundamentals in flux

The following projects are tackling the changes to the fundamentals that underpin publishing, from new economic models to re-skilling of the industry and the reconfiguration of educational institutions.

The disruption caused to industry and institutional interests means it is necessary to explore and develop new forms of publishing and, along with it, the development of new policies in IPR, economic models and societal goals.

- **Knowledge institutions** - Open Access (OA) and Open Education Resources (OER)
  - OA - Registry of Open Access Repositories. http://roar.eprints.org/
  - OER - Learning Registry, sharing and analytics. Stanford Research Institute backed project. Video 10 minutes http://ustre.am/_WsQO:Hcd
- **Economics** - Monopolies - Flat World Knowledge, breaking college text book monopolies with free and low cost text books. http://flatworldknowledge.com/
- **Skills and labour** - Single source and new areas for re-skilling in publishing - The Darwin Information Typing Architecture (DITA) builds content reuse into the authoring process. http://dita.xml.org/
- **Intellectual property** - Free to read - Open Library - an Archive.org project to make many books free to read online. http://openlibrary.org/
- **Reading Analytics** - this area is so new as of 2013 that not many open systems are available. A case in point is that Amazon Kindle does not share any user reader statistics, although it collects the data.

- **Reading and writing** - Rich Annotation – Document Cloud and Backbone.js its subsidiary. https://www.documentcloud.org/ | Scalar, multi-media annotation. http://scalar.usc.edu/

2.2. Making the case

The Consortium has a two part strategy to address the goals of its research remit to support digital scholarly publishing.

- Firstly, addressing the inefficiency of current digital publishing software with a single source approach, which is the key to dynamic publishing.
- Second, to quickly deliver innovation to support publishers undergoing digital disruption and bring them the benefits of dynamic-publishing. We do this by supporting the development of open infrastructures for publishing.

Single source. Why here? Why now?

There is an undeniable case for the publishing industry moving towards digital and multi-platform distribution so the challenges of Hybrid Publishing have to be faced, head on. All of the affordable commercial and open source software available are neither cost effective, efficient, nor error free (up to a reasonable tolerance) to carry out multi-format publishing. There are industry-level tools and providers who can carry out this work, but entry prices are restrictive.

Current multi-format publishing systems
- Not fit for purpose - software cannot efficiently produce publication ready documents
- Systems that do work have an entry prices start at 20k€ per annum per-publication
- Current market forces labour to be offshored to make publications profitable
- Offshoring encourages de-skilling in local publishing professions
- Technology environment prevents new publishers entering the market and holds back publishing product innovation

Single source

Single source means you have one master document document, that is available for dynamic publishing meaning it can be automatically converted to multi-format. With further automation in styles, reuse–broken up and recombined, with rights management, remuneration, sychronisation, revisioning and reading metrics, long term preservation and distributed, and converted into any new format–eBook, Print-on-Demand (PoD), app, library system, OER, etc. This means only one edit and proof, in current systems each format needs a new edit and proof.

Why here? why now?

Single source is not new. The issue of taking/turning unstructured paper and digital documents into a machine-readable format is a large part of the history of computing. Take for example Charles Babbage, computing pioneer, in the 1820s, looking to remove human reading error from manual logarithmic calculations, for example used for nautical navigation, with his computing device the Difference Engine[9] – an early concept for a computer. As a footnote it is worth mentioning that Babbage also recognised in the 1820's that scholarly publishers were even then creating restrictive access to knowledge with unfair practices, and ran his own

---

[9]   Difference Engine http://en.wikipedia.org/wiki/Difference_engine

Open Access campaign of sorts, resulting in his books being censored by publishers and banned from sale[10].

Schemas have been built for structuring documents over the last forty years but now with the consumerisation of technology with the Bring Your Own Device (BYOD) phenomenon the grip of the paper books has been broken and digital reading has taken off.

The consortium

Open innovation is now the default in the blue chip technology sector. The setting up of the consortium is our way of doing open research by creating a network of industry, open source communities, research groups and other stakeholders with a shared purpose of implementing open infrastructures for publishing.

The consortium looks to cover many aspects of publishing with the aims of supporting reliable, high quality software, with a vibrant economy to drive the technology innovators, developers and providers.

The consortium is based on open source software and open standards that form the foundations for innovation on the net and across new platforms, like mobile phones and tablets.

Many parts of the open source infrastructure for publishing are in place, for example in journalism there is Document Cloud[11] for referencing documents and for eReading there is *Calibre*[12], a desktop reader and shareable home library. We use Android as a mobile, tablet and eReader operating system.

The consortium contributes to these communities and acts as a meeting place for the different stakeholders, media activists, educators, publishers, technologists, creators, universities and entrepreneurs.

2.3. Roadmap

The roadmap is based on the development of our core technology, TypeSetr, for 'single source publishing' which is aimed at efficient multi-format conversion, layout and distribution. The final goal being transmedia publications, including rich media, social media and open learning environments. This goal is achieved by moving up through layers of increased complexity, starting with TypeSetr Academic in Google Docs and ending in open learning systems.

A series of rapid prototyping exercises and workshops with partners will be taking place in an accompanying timeline.

Milestones

Phase A - Conversion infrastructure

1. Open Research Portal - Dec 2013
In partnership with a number of technology providers, research groups and publishers an open research portal will be launched.

2. Inkubator Company Formation - InfoMesh Technologies UG[13] - Dec 2013
InfoMesh Technologies UG founded in Lüneburg by Simon Worthington as part of the Innovations Inkubator EU funded programme. The company will carry out commercial

---

[10]   On the Economy of Machinery and Manufactures. Charles Babbage, 842. Chp. 31. 'Combinations of Masters Against the Public', pp 312.
https://archive.org/stream/oneconomyofmachi00babbrich#page/312/mode/2up
[11]   Document Cloud https://www.documentcloud.org
[12]   Calibre http://calibre-ebook.com/
[13]   UG Unternehmergesellshaft - German equivalent of a Limited Liability Company.

development and client support for digital presses with Universities, museums and publishers, seeking further business to generate revenues for product development.

3. TypeSetr Academic - private beta - Feb 2014
A web base software platform release for multi-format conversion, with accompanying user and developer support environment. A Google Docs based toolset with multi-format publishing, collaborative writing, customisable templates, full academic markup to accommodate research papers, journals and monographs.

4. TypeSetr Academic Open Source Release - Feb 2014
Typesetr Academic will be released as a fully documented and supported open source compliant release, with source code on a software repository and publishing lab staff carrying out community support as part of their research process. The core algorithmic engine of TypeSetr Academic will also be open source compliant and so new inputs and outputs can be added.

Phase B - prototyping and workshops

5. TypeSetr Academic (TA) integration into Open Journal System (OJS) and Open Monograph System (OMS) - May 2014
TA document creation workflow integration into OJS and OMS with additional document input and output types.

6. TypeSetr - publishing partner prototypes and hacklabs
All of the prototypes would have the theme of 'Bring your own device (BYOD) distribution', to academia, institutions, the wider public and in search contexts.

BYOD is key in mutli-format publishing as it is the phenomena that is driving development in publishing, with readers moving to new devices, to accompany the book and library.

Each workshop compliments the related prototype and involved research content and videos being published.

Prototypes
1. Merve - search/BYOD distro   **2014 Q1**
2. Leonardo - academic BYOD distro   **2014 Q2**
3. AEE/Mute/Booksa/xm:labs - BYOD distro to OER   **2014 Q3**
4. Public Library/McLuhan - public BYOD (iPad/tablet) distro   **2014 Q3**

Hacklabs
1. Merve - search/BYOD distro/post-processing book scanning   **May 2014 Q2**
2. Leonardo - academic BYOD distro   **July 2014 Q3**
3. AEE/Mute/Booksa/xm:labs - BYOD distro to OER   **Sept 2014 Q3**
4. Public Library/McLuhan - public BYOD (iPad/tablet) distro   **Nov 2014 Q4**

Phase C - post HP planning and partner projects

7. Hybrid Publishing Consortium Stiftung - Feb 2014
Based in Lüneburg, Mute Publishing and partners will form a Stiftung (non-profit organisation) to develop the long term support of the consortium. It is important that the consortium is formed as a vendor and institution independent organisation to allow open and fair collaboration. Further grant funding for the project will be sought.

Phase D - further development

9. Distribution API - June 2014

The Distribution API would add distribution functionality to TypeSetr Academic. API publication distribution allows for the automatic distribution of whole documents and specific sections of documents. The automated distribution is important here for Open Access repositories. It also allows for synchronisation of document across different systems, as well as interaction with features like annotation and usage analytics.

Other projects, workshops, publications and prototypes

Completed

- Workshop - Transmediale publisher consultation, Jan 2013
- Workshop - Video Vortex. FLOSS MOOC survey, Feb 2013
- Book publication - Open learning and MOOC research, Coventry University/Mute/Ted Byfield (New School). OHP, May 2013. 'We're All Game Changers Now': Open Education - A Study in Disruption. http://linkme2.net/tq
- Research data publication - Academic Publishing, publication taxonomy, June 2013.
- Consultation - Future university press exploration with Coventry University - digital scholar consultation, costs and resource use. July 2013
- Meteor and Backbone.js experimentation (Re: Milestone 1. Open research portal)
- Publisher consultation and interviews
- Open Research Portal https://consortium.io
- Single source and dynamic publishing software requirements.
- Market analysis
- Mapping dynamic publishing and the unbound book examples

In progress

- Designing a stakeholder consultation
- Post-Digital Scholar
- Dynamic publishing research - technologies and standards frameworks

Postponed or shelved

- Indy Portal, direct selling and e-commerce – Nätverkstan/Mute

3. Design Research (Hacking methodologies)

Design research can be summarised as learning through a process of making. Our special version of Design Research also combines cutting edge Open Innovation and Service Design. A version of Open Innovation being borrowed from the Free Culture[14] activist movements and Service Design being from hyper-capitalist product development, with the need to differentiate products in an ever increasing 'product flood', and mix online, offline consumer relations.
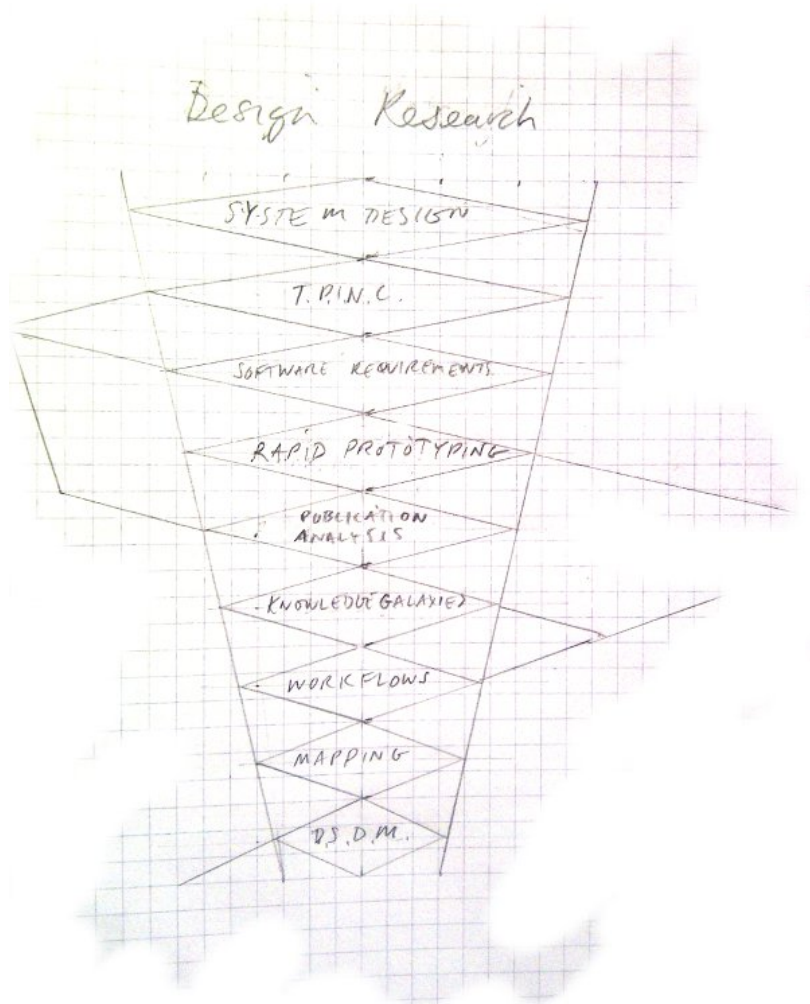
Figure: The diagram lists the methodologies or design processes that the Consortium uses in its research, technology prototyping and product development work.

The consortium mantras

- **Rapid prototyping** - allows us to test out dynamic publishing opportunities and suits HPC model of partnership building
- **Seamless integration into user workflows** - our findings demonstrate the need for a technical tool that lets publishers' workflows adapt easily to the demands of multi-format publishing[15]

---

14 Free Culture Movement http://en.wikipedia.org/wiki/Free_culture_movement
15 Based on publisher consultations with six cultural publishers in 2012 by Mute and LShift as part of a UK funded project by The Technology Strategy Board (Metadata strand) for automatic eBook creation.

Two key goals for the project have been in place since the outset of the project
- Firstly ´**single source and dynamic publishing**'
- Secondly **open publishing infrastructure**. Building a consortium for the open source community to collaborate on building the open publishing infrastructure

Goals

Single source and dynamic publishing

Single source and dynamic publishing is important, as it enables efficient workflows, brings new publishers into the market and allows new types of hybrid publishing to easily emerge: e.g., social reading, fluid curricular building, new editorial grouping etc. Without single source publishing the failures of multimedia age will be repeated in high-cost and time intensive authoring environments for products that quickly become redundant.

The Hybrid Publishing Consortium

The consortium has to balance collaborative transdisciplinary research with the rigors of a product-development. The single source can be seen as an infrastructural component, that then allows many products to be plugged in, emerge from or build on top of it. The consortium is a place for these cooperations, conversations and dialogues to take place, with experts in other parts of the publishing chain.

What can be achieved within the timeframe?

We have divided the tasks into four areas put together from the team's combined expertise:
1. **TypeSetr software product** - TypeSetr is the software tool that the consortium will be developing and implement as a single source publishing architecture performing multi-format publication conversion. The program will be developed, tested and launched during 2013.
2. **Qualitative research** - Carry out studies on scholarly research behavior patterns on how stakeholders collect, organise, and annotate their work. What tools are they using, what measurements are available? What kind of metrics might be useful for authors?
3. **Prototype dynamic publishing, future scenarios of publishing** - Single source publishing architecture allows publications to be easily distributed and have further algorithmic computing applied to the content, such as language translation and analytics.
4. **Open research** - a consortium web portal with open research and community outreach to industry, the open source community, publishers and other research groups. The objective of the open research is to foster long term partnerships to address the numerous technical publishing infrastructure components that need improvements and share vision/inspiration and latest developments.

Hybrid Methodologies for Transdisciplinary Research and Design

The Hybrid Publishing Consortium is currently a group comprising publishers, academics, artists, designers, and programmers working across multiple disciplinary fields and research interests, while also weaving across theory and practice. These activities will necessitate mixing, matching, and hybridizing methods as appropriate. However there are a few overarching principles and methods we will employ:

Systems Thinking. Simply put, systems thinking is the process of understanding how things, regarded as systems, influence one another within a whole.

Participatory Design[16]***.*** In the design of new open-source tools for publishing the HPC will rely on, and be informed by, users as co-creators and end-users, following participatory design's history of Democratizing Innovation[17] (von Hippel, 2005). The group will be carrying out a wide variety of research activities, events, workshops with our stakeholders to develop future scenarios and 'publishing futures.' Due to the complex nature of our research area, we are aiming to make the design process open and transparent to interested parties via our *Open Research Platform*.

Design Thinking[18] ***and 'Research through Design'.*** The process will integrate various tools and methods of Design Thinking to inform our process and outcomes. Complexity mapping, modeling, and rapid prototyping are simply means for us to learn about our given subjects through the process of making.

Service Design. The design process will entail the development of new product-service-systems with the intent of developing future business endeavors. Various tools for developing redefining areas for innovation will be used to develop future publishing scenarios, business models, and relationships within a given system.

These overarching processes may be combined with additional theoretical frameworks, tools and methodologies we have identified below:

1. **Systems mapping** - a tool to understand relationships within any given system, i.e. standards, stakeholders, comparators, tech stack etc.
2. **Publication analysis** - creates a taxonomy of publications, traditional and emerging. Breaking down all academic documents into constituent parts as content, structure and layout. This allows us to find commonalities between documents and ensure that documents are machine readable and thus computable.
3. **Knowledge galaxies**[19] - as publishing technologies change, from print, digital and to hybrid the technologies and conceptual models change. Part of our work is to map out the new forms of publishing in digital and hybrid publishing.
4. **Workflows** - workflows are the steps needed to create, use and read a publication. This can cover conventional publishing–commissioning, editing and layout–to new platforms like Bob Steins' 'Social Book'[20] project featuring group annotation. Since we consider workflows to include the creation and use of published materials, so we can include a reader or a librarian in workflows. At a certain point workflows tip into life-cycles and the circle is complete. Publishing turns back into new textual creation with the work of the post-digital scholar.

---

[16]   DiSalvo, C. and LeDantec A. (2013), Infrastructuring and the formation of publics in participatory design, Social Studies of Science, Sage UK, 3-5. http://ledantec.net/wp-content/uploads/2013/02/Le-Dantec-sss-2013.pdf

[17]   Eric von Hippel, E. (2005). Democratizing Innovation. Cambridge: MIT http://libros.metabiblioteca.org/bitstream/001/183/7/0-262-00274-4.pdf

[18]   Buchanan, Richard, "Wicked Problems in Design Thinking,"*Design Issues*, vol. 8, no. 2, Spring 1992.http://www.jstor.org/discover/10.2307/1511637?uid=3737864&uid=2&uid=4&sid=21101950270017

[19]   The Gutenberg Galaxy: The Making of Typographic Man, McLuhan. http://linkme2.net/tv

[20]   Social Book – http://livemargin.com/

5. **T-PINC** - **Technology, Power, Ideology, Normativity and Communication**[21] - discourse analysis related to how technology is developed and used.
6. **System and Software Requirements** - a process to build up clear requirements for software development, primarily based on user journeys. The system being the overall process that the software plays a role in, eg. printing or user skills etc.
7. **DSDM** - **Dynamic Systems Development Method**[22] - Agile software project management methodology.
8. **Project management** - aligning mission/objectives with projects and available resources and constraints.
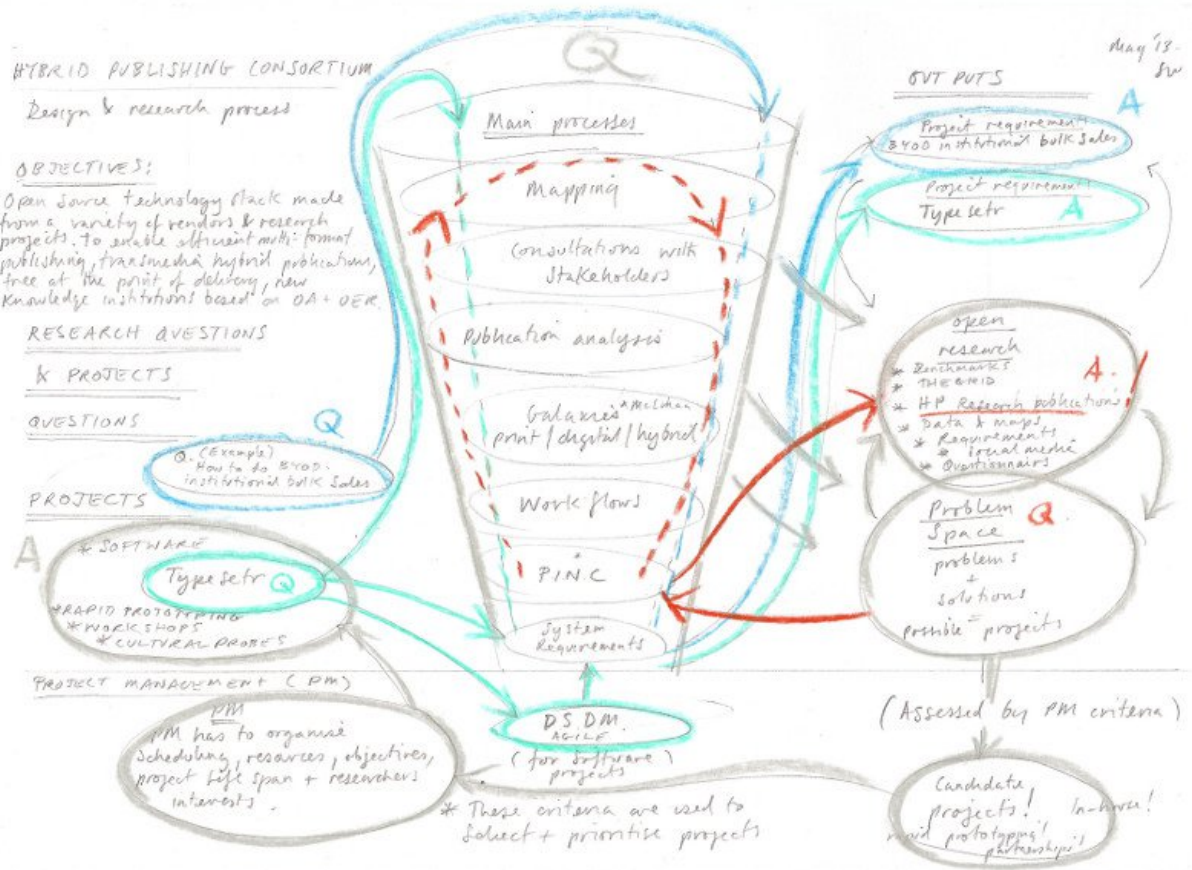
Figure: Design research processes. The core processes can be repeatedly queried with new project questions. May 2013.

3.1. T-PINC (Technology - Power, Ideology, Normativity and Communication)

T-PINC is a method of discourse analysis developed by Jochen Koubek, University of Bayreuth and based on PINC originated by Werner Patzelt, Technische Univerität Dresden. T-PINC was developed to examine decisions about technology development, usage and context. Where the stakeholders involved in a given situation are examined on the basis of the T-PINC vectors (Technology - Power, Ideology, Normativity and Communication) which are

---

21 Originated by Jochen Koubek, U of Bayreuth and based on PINC Werner Patzelt, TU Dresden. Get URL
22 Dynamic System Development Method http://www.dsdm.org/

used to see how and why those stakeholders take part to effect a technologies use and development.

T-PINC forms part of the design research processes of the consortium and is employed at varying stages; to identify who the varying stakeholders are, to examine more specific questions and to communicate the project's overall understanding of Hybrid Publishing.

T-PINC has been chosen to establish the questions about technological infrastructures and what they can enable or facilitate, and how this translates into an effect of change on society. There are such systematic failures in the institution of publishing, knowledge and learning, failures in terms of ICT (Information Communication Technology), economics, media diversity, skills and efficacy.

As an example, focusing on the institution of the university, one failure is in the realm of ICT technology. Universities have become trapped by IPR concerns and controlled by software vendor lock-in resulting in restrictive content approaches to content use. The alternative strategy of open and exploratory approaches to IPR and ICT has been neglected, making their systems–libraries, Online Learning Systems (OLS)–almost redundant in a Bring Your Own Device (BYOD[23]) environment or what can be described as a consumerisation of infrastructure, with smart phones, tablet and the abandoning of the desktop computer. Separately in the research role of the university, in publishing universities have allowed themselves to be exploited in a double-dip payments to publishers, paying their own scholars to contribute to publications, then buying back these same publications via the university library. University publishing has primarily narrowed its focus on a select range of peers, not addressing digital opportunities or new audiences. To sharpen the focus even further, if we look at Open Access (OA) publishing the barriers to take-up is not technological but to a large part the individualised self interest of scholars: their personal or departmental positions and standing being prioritised over a greater dissemination of their works.

Finally. In university staff training, scholars have near to zero 'digital scholarly skills'.

Hence T-PINC is needed to help unpick these complex relations to identify how technologies role can be strategised to play an effective role with working institutions.

---

[23]   BYOD, Bring Your Own Device - http://www.zdnet.com/topic-byod-and-the-consumerization-of-it/

An example T-PINC exercise:

The question: 'Open Access v closed publishing?'

- the failure of publishing by knowledge institutions (universities) and actors/stakeholders
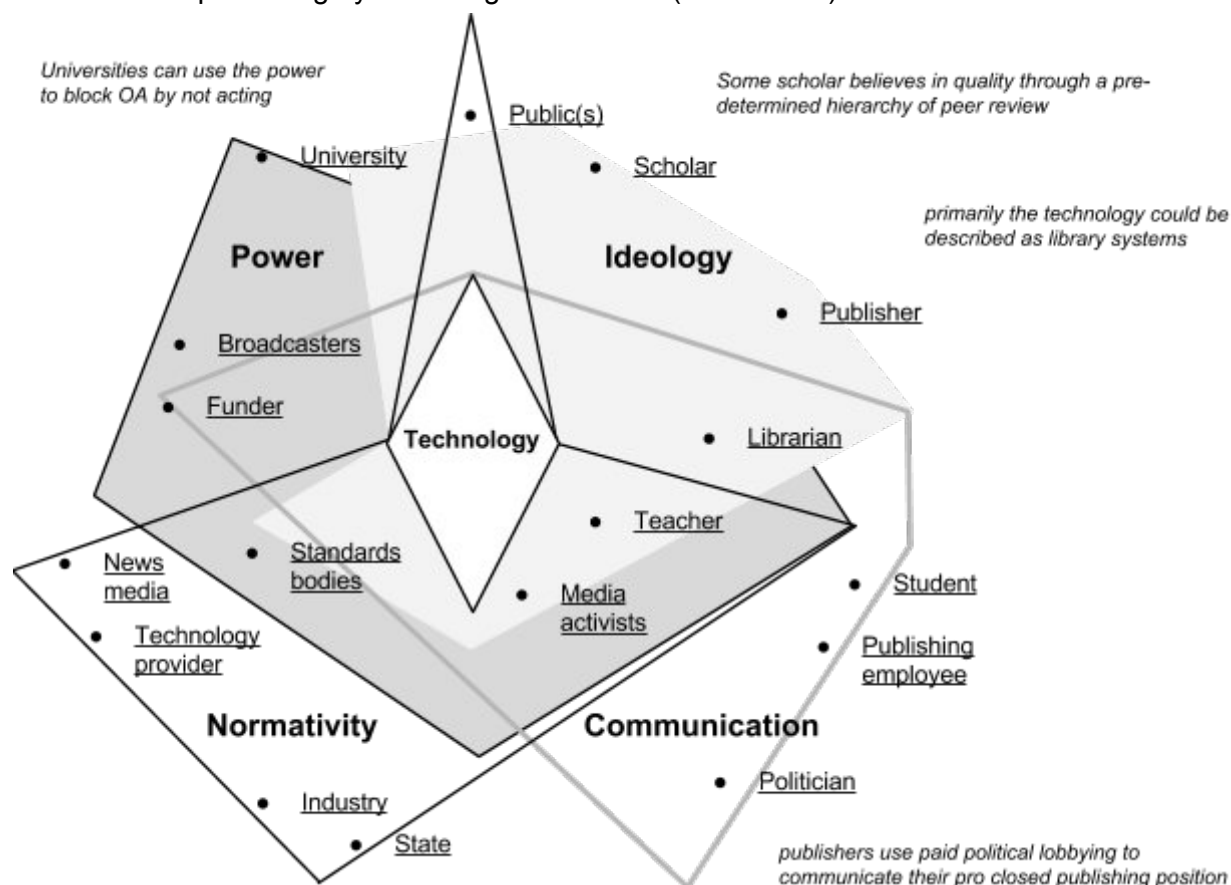

Figure: T-PINC Schema - 'Open Access v closed publishing'

The actors/stakeholders

Scholar	Publisher	Librarian	Student	Teacher
University	Funder	Technology provider	State	Politician
Publics(s)	Media activists	News media	Standards bodies	Publishing employee
Broadcasters	Industry			

Mapping the actors on the TPINC schema

- **Technology** - primarily the technology could be described as library systems. Then we have device hardware, BYOD, as well as general ICT. Packet networks technologies. Reading and learning analytics. Writing, formatting and knowledge management. Social networking, social graph technologies and learning analytics.
- **Power** - politicians, states can act to assert power by passing OA supportive bills or policies. Media activists and scholars can go on strike or petition institutions to

withdraw from commercial closed publishers services. Universities can use their power to block OA by not acting.

- **Ideology** - Media activists believe in the free circulation of knowledge. Some scholars believe in quality through a predetermined hierarchy of peer review.
- **Normativity** - copyright law is used to make the use of scholarly material illegal. The employment conditions of scholars determines decisions they will make in favour of closed publishing.
- **Communication** - publishers use paid political lobbying to communicate in favour of their closed publishing position. Media activists use the internet, OA and the media to deliver their message. Education expert bodies also communicate the OA message.

3.2. Workflows and stakeholders

Workflows in the publishing process are at the heart of our research. Examining workflows has a threefold benefit, they show what software requirements need to be addressed, what efficiencies can be made and most importantly the preferences and habits of the user.
In order to articulate the preconditions and potentials of dynamic publishing we will engage with a variety of stakeholders ranging from individuals, to scholarly institutions and indy publishers. We will build a series of engagement activities to strengthen–tools, resources, skills, and knowledge–for the single source document structuring and multi-format, Open Access publishing infrastructures.
We start out by analysing the workflows of our partners and stakeholders through a mixture of formal and informal processes. We want to find out how knowledge is circulated in a collaborative editorial group and how the group is organised around its publishing goals. By workflow we mean all aspects of a publication's life from the research process before authoring to the publication going into circulation and use, finally its long term preservation and use in new textual production in the emergent workflows of the post-digital scholar. Particularly for software building (Typesetr) the study of workflows is an indispensable activity. Typesetr needs to integrate a chain of interlocking processes, technical, real world situations and personal habits of the user, causing minimal disruption. And in a broader sense the studying of workflow will provide us with an overview of skill levels, economics and efficiencies of our chosen market. Beyond the workflow of our stakeholders we will investigate the changes and challenges of post-digital scholarship and the transition from conventional book publishing to unbound, hybrid forms of knowledge dissemination.
Our approach is designed to focus on each stakeholders singular engagement in processes that allow for co-creative development of new tools and strategies. This will simplify the work process, making works available in multiple formats, envisioning new forms of publishing and finding new readers. The design research process will introduce a reciprocal knowledge exchange platform, hosting a number of workshops and prototyping sessions. Aiming to discover spaces of opportunity: e.g., alternative tools, resources, networks, skills etc. We will facilitate the imagination of a speculative publishing landscape, its potential futures and practical implications. The outcome of our research will guide the development of our single source conversion platform: Typesetr as well as other research projects within the Hybrid Publishing Lab. We will make our research results open and transparent, publishing outputs on the Consortium's open research platform.

In concrete terms:
- One to one interview sessions with stakeholders, analysing workflows (via Skype and face-to-face)
- Personal relationship and document analysis

- Site visits to get to know stakeholders and their practice
- Templated notations documentation
- Script for key questions to be asked
- Hosting micro workshops with stakeholders to identify issues and potentials of hybrid publishing
- Envisioning alternative forms of dynamic publishing through rapid prototyping
- Process for moving findings over to research findings
- Sharing our approach and results on the online Open Research Platform, inviting stakeholders to get involved in the process

Stakeholders

- Stakeholders allow us to research publishing with a wider view outside of the confines of production and distribution issues and in doing so determine the influences directing how publishing is presently developed.
- What has been learned from previous infrastructure building projects like Transmission.cc[24] for activist video production is that merely producing tools does not answer the issue of media democracy. Instead, the accompanying issues of uptake, outreach, policy, design and economics need to be taken into consideration.
- Our stakeholders are identified and validated through the TPINK process.

3.3. Open research - a digital strategy

The Hybrid Publishing Consortium will partner with Merve Verlag, Leonardo Electronic Almanac, Eurozine, Mute Publishing, Data Futures, xm:labs and a variety of other groups to produce a shared open research platform. The platform is a way for the consortium to carry out open innovation, with the objective of supporting the creation of an open source infrastructure for publishing.
Address: http://consortium.io/
Start date: Dec. 2013

Digital strategy

A digital strategy is based on clearly defining your values, message and audience. It determines how you communicate to and engage your audience with these values. The audience being made up of staff/members, publics, community and stakeholders. A digital strategy then helps select what tools, platforms and content is produced, as well as methods of measurement and evaluation that are used to continually develop the strategy.

Open research

By its nature the Hybrid Publishing Consortium needs to partner with a variety of vendors and research groups to enable the expansive technology stack needed for open publishing systems.
An open research strategy allows us to create a greater profile for our research, gain contributions to our research through questionnaires and enter into collaborations with commercial vendors, open source projects and research institutions.

Examples of open research platforms

- Mozilla Labs - https://mozillalabs.com/
- Learning registry - http://www.learningregistry.org/documents/starter-resources
- Open Knowledge Foundation - http://okfn.org/projects/

---

[24] Transmission - video activist open source infrastructure building network http://transmission.cc/

- WikiMedia Foundation Labs - https://wikitech.wikimedia.org/

The portal would publish the following
- Project plan
- Project collaborative areas
- Roadmap
- Maps
- Research papers
- Presentations
- Our research methods

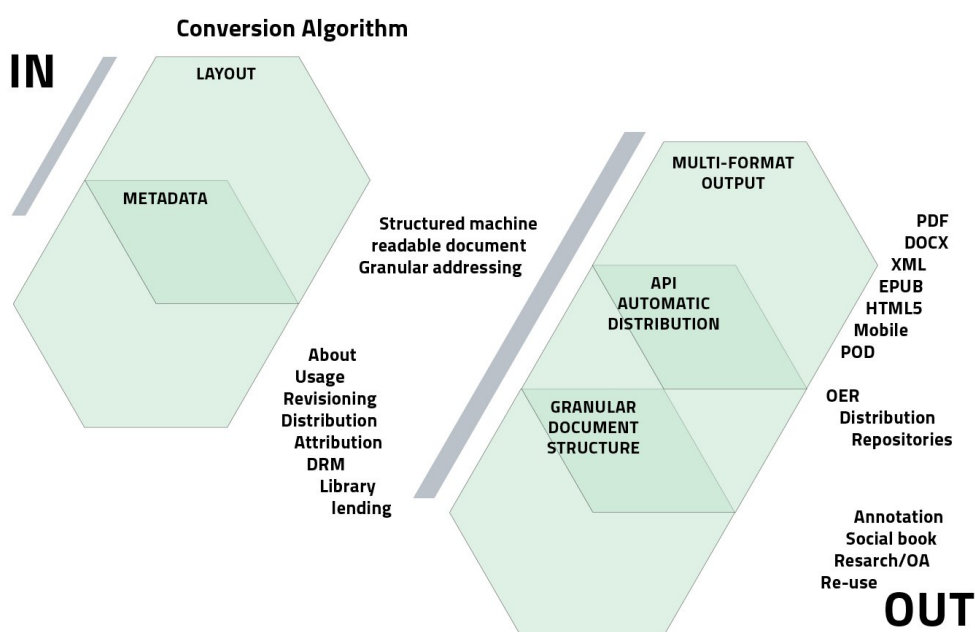## 3.4. Single source - system requirements

Single-Source Publishing

Figure: Single source workflow diagram

Multi-format publishing based on a single source publishing technical model?

The system requirements process is one of our design research processes, which also includes TPINC process used to help identify stakeholders then used in the system requirements process. The stakeholders being vitally important for gathering the constraints placed on software to solve real-world problems.

The process has three phases that are applied repeatedly to produce a finished requirements document. After this process is complete the 'requirements document' then enters into an agile development process, where once again the design of the software is reworked at varying stages.

Three repeated phases
- Requirements gathering
- Requirements specification

- Requirements validation

These three phases are applied as we move through stages of planning refinement.
Starting with;

1. Identifying a real-world problem that can be solved and testing this assumption - in our case we begin from the problem that current multi-format publishing systems are 'inefficient, costly and error prone'.
2. Consulting with stakeholders/users
3. Feasibility study - gathering other areas of knowledge to ensure a functioning system can be made, these include; establishing our goals, domain knowledge and stakeholders.
4. A variety of modeling exercises are undertaken at different stages to address different issues; conflicts of interest in requirements, identify problems to be solved, sort our system requirements from software requirements etc.
5. Translation and validation - final documentation has to be made, including schedule, costs and risks
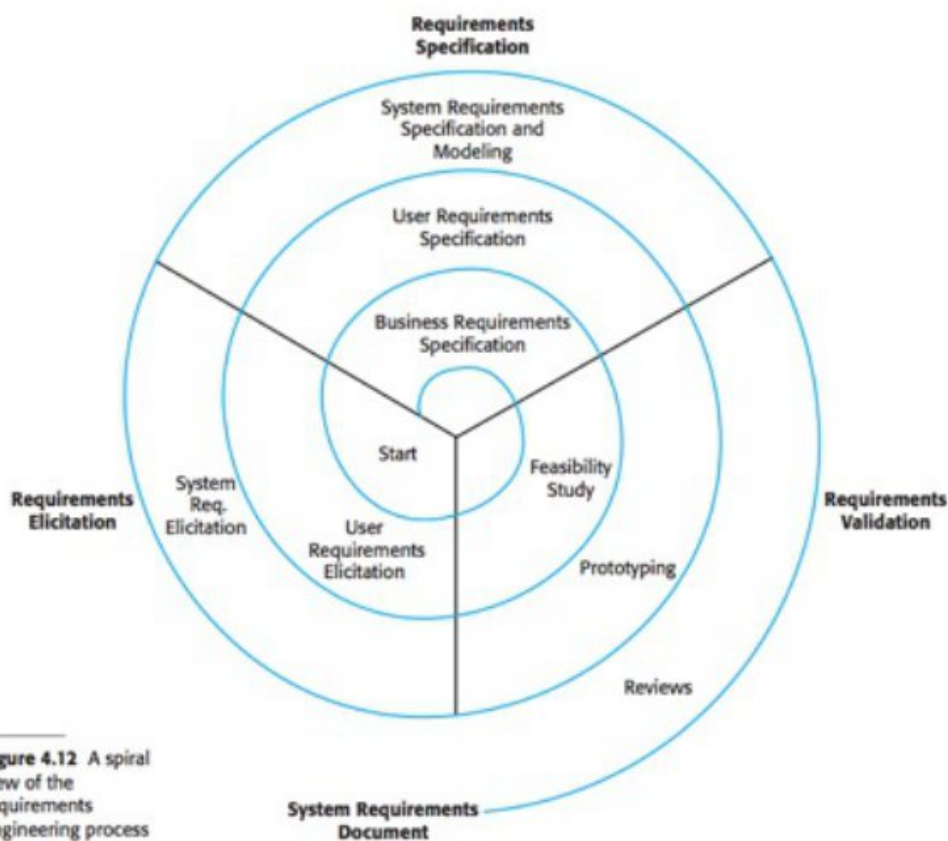
Figure 4.12 A spiral view of the requirements engineering process

Sommerville 2011, p. 99

Figure: Sommerville, 2011, requirements process diagram[25].

---

25 Somerville, Software engineering, ISBN: 978-0-13-703515-1 — 9th ed - http://neerci.ist.utl.pt/neerci_shelf/LEIC/3%20Ano/2%20Semestre/Engenharia%20de %20Software/Bibliografia/Software%20Engineering%209th%20ed%20(intro%20txt)%20-%20I. %20Sommerville%20(Pearson,%202011)%20BBS.pdf | Also see: http://www.softwareengineering-

1. **Outputs formats** - ePub, mobi, designed PDF (screen and print ready, with covers for print ready version, HTML5, TEI, DITA, XML and master document.
2. **Inputs formats** - Google docs, docx, database, InDesign, Scribus, MS Sharepoint, realtime text editors such as OX Documents[26] or other Operational Transformation algorithm based editor and other API inputs.
3. **Metadata** -
   a. **Distribution metadata** - Have ability to make master document, an authoritative unique source with all the distribution metadata for publications and output metadata required by each output format, channel or platform.
   b. **Usage metadata** - consider requirements for usage metadata such as learning and reading analytics.
4. **Metadata inputting** - allow metadata to be inputted at any time by different workflow members with relevant parts of knowledge they hold.
5. **Markup** - academic markup, styling, citation, referencing management.
6. **Structured document** - machine readable document, semantic markup, TEI, OAI, DITA.
7. **User feedback for structured document editing** - with comments on the doc prompting corrections for re-input into system by the user. These comments should be two tier, generated by two separate methods, automatic/machine made and by an admin operator.
8. **Comments and annotation** - what standard do we use for user/reader annotation? W3C Annotea, Framework 7 EU projects, or industry standards in use by (Amazon, MS, OO, Adobe).
9. **Checking-in/out (round tripping), versioning** - the system can be treated as a versioning system, where user can check-in and check-out documents, rollback documents to any version. Different document types could be round-tripped; docx, Indesign, GoogleDoc etc, accepting the changes made in any of these editing environments?
10. **Editing templates** - edit LaTex for example with Sharelatex. https://www.sharelatex.com/. With multiple templates needed if we're outputting multi-format publications. With permissions system for restricting to groups and users.
11. **Heuristic conversion features** - the system to learn about formatting and conversion for specific publishers and users, as well as for the common share rule sets.
12. **API distribution** - automatic distribution as well as translation to different formats, for example from TEI to OAI.
13. **Workflow, users/group access control** - because documents will used by a variety of different users, for example for outsourcing proofing or layout, we need a workflow user/group access control set of functionality.

4. Technology

4.1. Open infrastructure for publishing

A technology stack is a layered set of components and services needed to perform a given task. A classic technology stack for a web server and in open source terms is referred to as LAMP (**L**inux [Operating System], **A**pache [HTTP Server], **M**ySQL [Database] and **P**HP

9.com/

[26] Real time OX Documents https://ox.io/ox_text

[Scripting language). In a stack components can be swapped out for a different version or vendor, for example Linux could be replaced with Windows. Obviously you would prevent your stack being open source by using Windows.

In the case of the consortium the technology stack is core to what we do as it is how an open source public infrastructure for digital publishing is delivered. The consortium looks to identify, support and partner with third parties to enable an open infrastructure for publishing and this is where we would need to choose a stack and then examine how it can be applied to the scholarly publishing.

The open source publishing stack that we look to provide is vital for scholarly publishers, since nearly all of them sit in the majority bracket of the 80% of publishers who are priced out of the opportunities that access to the stack offers, which is access to new markets and revenues. An annual cost of running a single source, XML first workflow is in the region of 100,000€ in end-to-end costs. Hence it is the consortiums top priority to ensure publishers can access this technology at an affordable price. The dynamic publishing market means that publishers can offer content in bulk subscriptions to institutions, distribute title listings to online repositories and distributors, as well as carry out a wide variety of publishing options to reach readers. Currently most publishers use off-shore cheap labour conversion providers to make digital publications like eBooks. This is satisfactory as a low quality, short term measure, while the digital publishing market matures, but unless publishers make use of the single source publishing model they will continue to see profits fall and risk closure.

The stack is provided by a variety of vendors, research groups, open source communities, industry and standards bodies. The consortium's contribution is in the specialisation of single source multi-format conversion.

Publishing stack - pricing comparison - proprietary v open source

This stack shows the basic requirements for multi-format publishing. The pricing is based on three years of end-to-end costs. The costing does not include publication production costs. The overall publishing stack can be extended in many ways, depending on the specific workflow, for example if social reading or OER are introduced. And as was mentioned before with the example of the LAMP web server stack, vendors and services can be swapped out for alternatives.

The stack can have various workflows applied to it to make custom stack selections, or paths through the stack.

Publishing stack - pricing – 3 year period/ small publisher. Proprietary v Open source						
	Proprietary			Open Source		
Software	Product/Vendor	Cost	Dev/staff costs	Product/Vendor/Maintainers	Cost	Dev/staff costs
Office suite	MS Office	4,480 €		Open Office, Apache	0 €	
File Server	Windows server	1,159 €	18,000 €	Ubuntu Samba Serve	0 €	18,000 €
Project management	Basecamp	1,764 €		Mantis	0 €	
Knowledge Management System	BackPack	1,764 €		MediaWiki	0 €	
DTP/PDF/Artwork	CS 5.5	4,527 €		Images GIMP, Graphics InkScape, PDF annotation Okular	0 €	
Fonts	Various	500 €		Open Font Library	0 €	
Conversion Auto DTP/eBook/Tablet	Woodwing	6,000 €	30,000 €	Typesetr	0 €	4,000 €
BookData	ISBN-Agentur für die	1,500 €		-	1,500 €	
Content Management System (CMS	Kentico	1,400 €	30,000 €		0 €	30,000 €
CCMS	Easydita	90,000 €	30,000 €	N/A		
Application Program Interface (API)	Apigee		10,000 €	MetaMeta, Mozilla	0 €	5,000 €
Semantic markup	Custom build		5,000 €	-		5,000 €
Ecommerce	Actinics	500 €	12,000 €	x com, eBay	0 €	12,000 €
Payment provider	PayPal	1,200 €		-	1,200 €	
Analytics web/ social	Trackboard	2,484 €		Pikwik	0 €	
Customer Relationship Managemen	SalesForce	24,480 €		Sugar CRM	6,385 €	
Translation	n/a			Worlwide Lexicon	n/a	
Accountancy	Xero	684 €		GnuCash	0 €	
Digital Asset Management Services	AssetBank	12,968 €	6,000 €	Nuxeo	3,000 €	6,000 €
App – Apple/ Android OS	Mobile Roadie	3,958 €	6,000 €	PhoneGap, Adobe	0 €	6,000 €
		159,368 €	147,000 €		12,085 €	86,000 €
			306,368 €			98,085 €

Figure: proprietary v. open publishing stack price comparison

4.2. Single source and dynamic publishing

There are a variety of competing approaches to single source publishing, below is one route that we have identified as being tried and tested in the field of 'Technical Writing'. Technical writing is help systems and engineering documentation. Although the technology from these forms of writing cannot be directly transferred to the humanities, there are still many lessons to be learned.

XML first workflows

Single source publishing or 'XML first workflows' enable what is known as dynamic content publishing. Dynamic content publishing is the vital key for publishers to enter new markets and access their target audience in a time of digital disruption. It allows publications to become fully digital so they can be automatically converted, styled, broken up into components and distributed.

Dynamic publishing is where content can be queried and specific components brought together and published in automated workflows in multi-format to a variety of channels. An example convention in dynamic publishing is known as 'topic based authoring'. For example if you had a set of essays for a reading group, then a set of sections, chapters and book sections could be automatically collated and published as full text multi-format publication and as a data feed of titles and links for a mobile app. Topic based authoring is commonly used in things like customer help systems or large technical systems like aeronautics engineering.

The 'XML first workflow' approach to single source has been around for over ten years and was preceded by other languages such as SGML. The objective is to structure the content into

content components, as opposed to leaving documents as single unstructured entities in the way that a Word or PDF document are commonly used.

Large companies and publishers have in the majority of case adopted this XML first workflow model, but the costs and resource use are way beyond the majority of publishers, academics or individual universities budgets. System costs start at the hundreds of thousand euros and take in the region of a year to build, then there are the annual running costs. Since 80% of publishers in the publishing market have turnovers under 4€ million and not enough titles per annum the entry point cannot be justified these types of Dynamic Publishing systems and have stayed out of their reach.

This XML first architecture involves the following building blocks. Having a CCMS (Content Component Management System), XML editing/authoring environment, content legacy system, choice of XML output and templating engine.

Architecture component	Example provider	Example standards
CCMS (Content Component Management System)	Afresco (overlaid) or DitaToo http://ditatoo.com/	Not applicable
XML editing/authoring environment	DITA tool kit (open source) http://en.wikipedia.org/wiki/DITA_Open_Toolkit List of editors http://en.wikipedia.org/wiki/Comparison_of_XML_editors	Not applicable
Content legacy system	Stilo - Migrate tool http://www.stilo.com/products/conversion/	Custom pipelines
XML output standard	DITA tool kit for publishers http://dita4publishers.sourceforge.net/	DITA 3.0
Templating engine	Freemarker http://freemarker.org/fmVsXSLT.html	XSLT

Structured writing and DITA

Darwin Information Typing Architecture (DITA) is a framework and tool set made open source by IBM in 2005 and appears to have taken over from older formats such as DocBook. DITA allows for what is know as structured writing, where content is marked up as its created, either by the author inputting into templated fields or by some follow up process. With structured writing, content can then easily be reused, opening up exciting new publishing opportunities.

DITA community http://www.ditawriter.com/
History http://www.digplanet.com/wiki/Darwin_Information_Typing_Architecture

Our target users and the consortiums contribution to this architecture

For the smaller publisher these technologies have not been affordable or might seem irrelevant to their needs. The task of the consortium is to make these architecture affordable for the majority of smaller publishers, which makeup in the region of 80% of the market. The

case for making this approach relevant to publishers is by our research project exploring ways that these small publishers can be enhanced by dynamic publishing.

5. The Hybrid Publishing Consortium

The objective of the hybrid publishing consortium is to support the building of an Open Source compliant software infrastructure for digital publishing. An infrastructure for publishing would cover core issues as well adopt new components as it develops. The consortium would be formed of a number of partnerships with components of the infrastructure being provided by a variety of publishers, technology companies, open source communities, research groups and industry bodies.

Core publishing infrastructure issues are;
- Multi-format digital document conversion
- Multi channel distribution and sales
- API automatic distribution
- Document conversion and single source architectures
- Dynamic publishing, structured writing and topic based authoring
- Open IPR and 'free at the point of reading (use)' models of publishing
- Open Access and Open Education Resources
- Rich media and Transmedia publishing
- Long term preservation
- New forms of reading, writing and research
- Developing new types of digital presses for universities, small publishers, Galleries, Libraries, Archives and Museums (GLAM) sector
- Sales and reader analytics

Quality assurance, open source software, open standards and open IPR

As well as being open source compliant, the infrastructure would be based on open standards and advocate open IPR content models.
Within the area of open source the consortium has the priority of encouraging quality assurance and with partners we would look to make use of our own research design processes and partner expertise to make a contribution to this important issue.

Open innovation

The infrastructure project acts as part of the consortium's open research strategy allowing for sharing and partnership between different partners and the wider community. The consortiums specialisation of single source and dynamic publishing also finds a place within the consortiums infrastructural approach. This is especially valuable as the technology we are focusing on has many opportunities to be applied in different areas and it would be our interest to have partners create a variety of products based on our technology.

Governance and long term plans

The consortium has to balance the interests of the different partners and this requires that the consortium is a neutral body and has transparent decision making and organisational processes. Independence from institutions and vendors can be ensured by constituting the consortium as a Stiftung, or non-profit, in early 2014.
The consortium is currently based at the Lüneburg Innovation Inkubator, a major EU project within Leuphana University of Lüneburg, financed by the European Regional Development

Fund and co-funded by the German federal state of Lower Saxony. The Inkubator project comes to an end in July 2015, over this period the consortium will develop plans for continued operation. Part of these long terms plans is to create spin-off companies and commercial collaborations in the Lüneburg region and internationally as part of our business incubator funding remit.

History

The consortium had its beginnings as a Mute Publishing project lead by Simon Worthington, Mute Publishing's Director of Digital and Pauline van Mourik Broekman, Mute Publishing Director when Mute magazine adopted an OA policy on 2004. Subsequently in 2007 the issue of single source publishing was first approached with the help of Mute's systems administrator Darron Broad at a meeting with the Open Source Publishing[27] organisation in Brussels. The work was continued with support from the Open Society Institute (OSI)[28] and later the London Develop Agency and ANGLE plc[29] in 2009 launch a POD production company PicoPress providing digital press services to publishers. Then later in a Technology Strategy Board[30] funded metadata project over 2012 for an RnD project for eBook conversion. As of 2012 the consortium project became part of the Hybrid Publishing Lab, Innovations Inkubator, Leuphana University.

5.1. Lab and people

The Hybrid Publishing Lab

The Hybrid Publishing Consortium is a project of the Hybrid Publishing Lab. With the wider lab being staffed with eighteen people, made up of professors, visiting professors, doctorates, post-doctorates and research associates. The Lab covers research areas of Open Access, open learning, the future university press as well as creating a press for the Centre for Digital Cultures.
The consortium is an integrated part of the lab's research with all lab members giving input and feedback on the consortium's work. Lab team members periodically move between project or work on joint collaborations like workshops and colloquiums.
It is also important to emphasise that the lab site within the Innovations Inkubator and Centre for Digital Cultures, with over ninety staff, specialising in digital media and combining research areas that cover industry topics such as gamification and public service broadcast, with an approach which is informed by the humanities and critical discourse.

People

Working team
Simon Worthington - project leader
More than twenty years of experience as an independent publisher. Co-founder of Mute magazine. As Mute director of digital Simon has been involved in managing numerous independent and activist media technology projects, funded by London Development Agency, UK - Technology Strategy Board and Open Society Institute. Experience covers business, organisational and innovation management. Most recently until 2012 leading 'Art of Digital

---

27 Open Source Publishing, http://osp.constantvzw.org/
28 Open Society Institute, now Open Society Foundation (as of 2011) http://www.opensocietyfoundations.org/
29 ANGLE plc, http://www.angleplc.com/about-us.asp
30 Technology Strategy Board, https://www.innovateuk.org/

London (AoDL)', a digital strategy training project for Arts Council England, two hundred and seventy-five London clients. Simon studied at UCL, London and CalArts, California.

Agata Królikowski - information architect

Agata Królikowski studied law and informatics at the Humboldt University in Berlin, where she subsequently worked in the field of informatics in education and society. In 2012 she joined the Innovation Inkubator at the Leuphana University, Lüneburg. As a PhD student her research focuses on the intersection between informatics and laws regarding privacy protection, copyright and the control of information in general. She is the spokesperson for the working group "Internet and Society" of the German Informatics Society.

Minuette Le - strategic design

Minuette is a designer and researcher with an interest in open innovation systems and re-thinking public services. She has a background in both business and design and completed her MFA in Transdisciplinary Design at Parsons the New School in New York, an experimental curriculum focused on design-led research, systems thinking, and service design. Her practice involves design across disciplinary fields ranging from community-based projects, public services, to emergent digital futures. Her work at the Hybrid Publishing Lab is researching the Post-Digital Scholar and shifting research and learning paradigms.

Ulrike Gollner - interaction designer/researcher

Ulrike is an interaction designer and researcher with a background in computer science and media arts. She completed her MA in Interface Cultures at the University of Art and Design Linz and BSc in Engineering in Media Technology and Design at the University of Applied Sciences Upper Austria. She brings many years of experience from Ars Electronica Linz and the Design Research Lab of the University of the Arts Berlin, most recently designing and developing hardware and software prototypes in the context of disability-inspired interaction design. Her work at the Hybrid Publishing Lab explores emerging scholarly practices in today's paradigm of 'open everything'.

Johannes Amorosa – developer

Christina Kral - stakeholder coordinator/blue sky facilitation strategist

As an artist she focuses on forms of participation and utopian idea development. Within the Leuphana University she researches the future of learning, concentrating on alternative forms of knowledge production and visual, as well as contextual, translations of existing knowledge to reach new publics. The approach raises the questions of how these forms will look like and how will this influence the way we learn and ultimately alter educational institutions.

5.2. Partners

The consortium is able to carry out publishing research that is beyond the scope of many publishers and academic because of their workloads and resource constraints. And with our Innovations Inkubator we look to share this research using open research methods and open source IPR models to provide real-world solutions to the publishing issues these publishers and academic face.

In relationship to industry groups who are much more advanced in the development of single source and XML first workflows our work plays a role in translating how these approaches can be used by the smaller scholarly publisher.

The publishing infrastructure project of the consortium is by its expansive nature made of collaborations by domain knowledge specialists and so partnership forms part of this work.

Current partnerships and associations

Merve Verlag

Centre for disruptive media, Coventry University
Leonardo Electronic Almanac
Mute
xm:labs, Experimental Media Lab, Academy of Fine Arts Saar
Hogeschule - Amsterdam and Rotterdam
Data Futures, Westminster University

Publishers

The consortium is working with a series of publishers that cover a wide spectrum of academic publication types to ensure that we can capture the requirements of their workflows. We work with the publishers in a variety of ways, from carrying our analysis of a specific publication type workflow, like a academic monograph. Over to in-depth collaboration on issues like BYOD bulk sales to institutions.

The consortium also works with a number of publisher support agencies such as Eurozine, Nätverkstan and Independent Art Publishers, who have a combined publisher list of over three hundred publishers to consult with.

Current publisher list

Autonomedia, b_books, Dienadel, ISEA Leonardo Electronic Almanac, Liquid Books, Merve Verlag, Mute, Post-Media Lab, r0g_agency, Kuda, Booksa and Berliner Gazette.

5.3. Comparators

Single source publishing has its origins in technical two worlds. Firstly technical documentation, engineering, legal documentation and help system publishing. Secondly the creation of the networked computing and the development of Internet, as personified by the Web as a universal document system. The application of single source in the humanities and literature is less common. This makes our job of finding comparators a little more difficult and requires a degree of translation and interpretation of relevant projects.

We have chosen two categories of comparable projects, single source and publishing infrastructure consortiums. Then dividing these into industry groups and research groups.

Single source

Industry

Commercial providers
Woodwing http://www.woodwing.com/
Stilo http://www.stilo.com/
Le-tex http://www.le-tex.de/en/

Open source, open standards
DITA Toolkit, IBM http://en.wikipedia.org/wiki/DITA_Open_Toolkit
Sphinx - python document generator http://sphinx-doc.org/
LaTex http://www.neodoc.biz/fr/calenco/index.html

Research

- Public Knowledge Projects http://pkp.sfu.ca
- Hasso Plattner Institut, Semex - Semantic Media Explorer - http://www.hpi.uni-potsdam.de/meinel/knowledge_tech/semex.html

- University of Amsterdam Press https://www.aup.nl/

Consortium

Industry

IDPF (Industry Digital Publishing Forum) http://idpf.org/
OASIS (Organization for the Advancement of Structured Information Standards)
https://www.oasis-open.org/
W3C (World Wide Web Consortium) http://www.w3.org/

Research

LGM (Libre Graphics Meeting) http://libregraphicsmeeting.org/2013/
Digital Publishing Toolkit http://digitalpublishingtoolkit.org/
IfBook - The Institute for the Future of the Book http://www.futureofthebook.org/blog/

Related

There are a variety of project which do not exactly correlate with the Hybrid Publishing Consortium but instead their values or leadership in their field builds a strong affinity.

Open Archive Initiative, Cornell University http://www.openarchives.org/
Open Knowledge Foundation http://okfn.org/
Data Liberation Front http://www.dataliberation.org/

Appendix

Glossary

See a full Glossary at http://consortium.io/glossary and
https://en.wikipedia.org/wiki/Book:Dynamic_and_Single_Source_Publishing_Glossary

(Below is a limited sample)

Dynamic publishing - Dynamic publishing is a digital workflow where the publication processes are automated–layout, multi-format conversion, distribution, rights management, file transfer, translation workflows, document updates, payments and reading metrics. Secondly dynamic publishing to give access to new audiences and revenues, due to the fact that the publication is available on request, re-use and recombination and can be automatically supplied to libraries, repositories and Web 2.0 reading platforms. Dynamic publishing is enable by machine readable, structured document formats, such as IBMs FLOSS framework Darwin Information Typing Architecture (DITA).
Publishing API - An Application Process Interface (API) is how one computer talks to another computer. For publishing the API acts as an automated place where distribution and its related activities happens-orders, file transactions, metrics, payments etc. A publishing API allows for free and paid content to be distributed. With an API a publisher can connect directly into an library information system, or a new Web 2.0 reading service can initiate their own connection to the API and start interacting with content. Example publishing APIs are Pearson and The Guardian, see: http://developer.pearson.com/ and http://www.theguardian.com/open-platform. API listings site http://www.programmableweb.com/apitag/?q=publishing

Single source - single source is a technical architecture for multi-format digital publishing. It means you have one master document document that is automatically converted into any required digital publication format eg eBook, POD, PDF etc. This type of architecture for publishing is required as current parallel formatting workflows are inefficient and low quality. With single source any document update can be automatically distributed to all the output formats. In current digital workflows each output format needs to be individually updated, making digital publishing prohibitively expensive and almost impossible.

XML first workflows - is an alternative technology approach to document management from the dominant models of word processing technologies like MS Word. What differentiates XML first workflows and makes them superior is that the document is structured and machine readable. This allows for speedier and greater information and knowledge use. Most large publishing companies use XML first workflows.

Reading analytics - Reading analytics means the collection of detailed reader activity and then the methods to make sense of this data. Traditionally a publishing would only have one dimension of readership, for example, a quarterly book sales report. Currently the knowledge about reading activity is undergoing a revolutionary explosion, with the dimensions of data on the ready vastly increasing. The level of data about a reader can go down to location, time and even a wide spectrum of behavioral characteristics inputs via mobile devices–heart rate, body temperature and more.

END

–