



# RESEARCH DATA MANAGEMENT SUPPORT SERVICES

*by libraries*

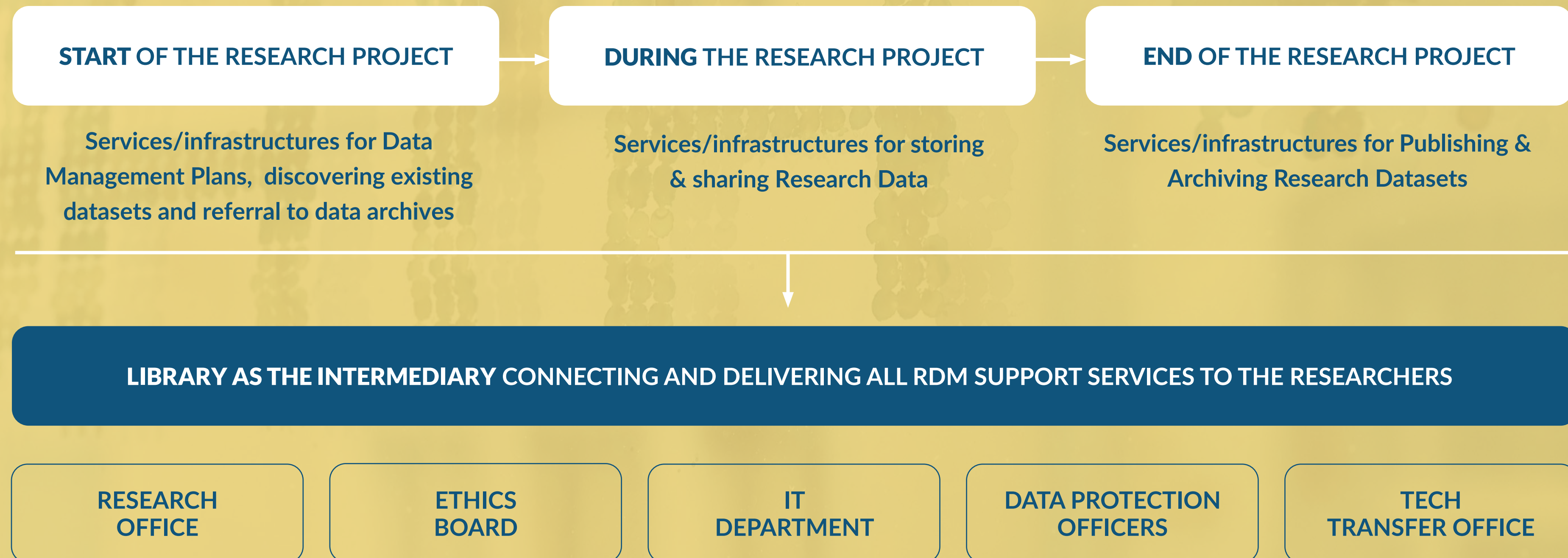
In collaboration with other stakeholders

TOOLKIT BY LIBER AND ADBU

# RESEARCH DATA MANAGEMENT SUPPORT SERVICES

*by libraries*

in collaboration with other stakeholders



# CONTENT

<b>INTRODUCTION</b>	<b>4</b>	<b>Job categories</b>	<b>15</b>
The rapidly increasing importance of research data	4	Data steward/data librarian	15
Many stakeholders are involved in the research support services with a special role for the library	4	Data manager	15
Purpose toolkit	4	Educator/trainer	15
Basis of this toolkit	4	Data curator	15
Four stages of development	4	Community manager	15
<b>1. INITIAL STAGE</b>	<b>5</b>	<b>[B] COMPETENCES</b>	<b>16</b>
Services	6	Three categories of competences	16
Organisational structures	6	Technical competences: knowledge and know-how	16
Competences	6	Interpersonal skills / soft skills	17
Remarks	6	A closer look at the competences of data stewards	17
<b>2. DEVELOPING STAGE</b>	<b>7</b>	Acquiring Competences for RDM support services	17
Services	8	<b>[C] RDM SERVICES</b>	<b>18</b>
Organisational structures	8	[1] Data Management Plans	18
Competences	8	[2] Dataset discovery tool or portal	18
Remarks	8	[3] Infrastructures and services for data sharing and data storage	18
<b>3. DEVELOPED STAGE</b>	<b>9</b>	[4] Data management support	19
Services	10	[5] Guidance and referral services to existing (Multi-disciplinary) data archives	19
Organisational structures	10	[6] Institutional data repository	19
Competences	10	[7] Registering datasets	19
Remarks	10	[8] Monitoring datasets	19
<b>4. ADVANCED STAGE</b>	<b>11</b>	[9] Community management	20
Services	12	[10] Research software	20
Competences	12	<b>[D] WHY THE LIBRARY?</b>	<b>21</b>
Remarks	12	<b>About this toolkit</b>	<b>22</b>
<b>[A] ORGANISATIONAL STRUCTURES &amp; JOB CATEGORIES</b>	<b>13</b>	About the study	22
<b>Organisational structures</b>	<b>13</b>	About this toolkit	22
Initial and developing stage: specialisation versus integration	13	Acknowledgements	22
Developed and advanced stages: first-line/second-line service structure	14		
RDM support ecosystem with the library in the middle	14		
Involvement of other library departments	14		



# INTRODUCTION



## THE RAPIDLY INCREASING IMPORTANCE OF RESEARCH DATA

FAIR research datasets are increasingly important as output of scholarly research. FAIR research data are research data that are Findable, Accessible, Interoperable and Reusable. This makes Research Data Management (RDM) support services particularly urgent.

## MANY STAKEHOLDERS ARE INVOLVED IN THE RESEARCH SUPPORT SERVICES WITH A SPECIAL ROLE FOR THE LIBRARY

At least 6 different departments within the University can be involved in research data with the library taking up a unique role as the intermediary connecting and delivering all these services to the researchers. See for further details [\[D\]](#).

## PURPOSE TOOLKIT

The purpose of this toolkit is to provide guidance to research libraries who want to start and/or expand RDM support services.

## BASIS OF THIS TOOLKIT

This toolkit is based on a study of research libraries that have developed RDM support services: [Open Science Services by Research Libraries: Organisational perspectives - a LIBER and ADBU report](#); Maurits van der Graaf; 2023

## FOUR STAGES OF DEVELOPMENT

This toolkit provides guidance to research libraries to develop their RDM support services in four consecutive phases. Per stage an overview of the services is given, followed by an overview of the organisational structures and the competences. The overviews per stage contain links to more detailed information about [\[A\] the organisational structures & job categories](#), about [\[B\] the competences](#) and about [\[C\] the services](#).



*Four stages of development*

# INITIAL STAGE

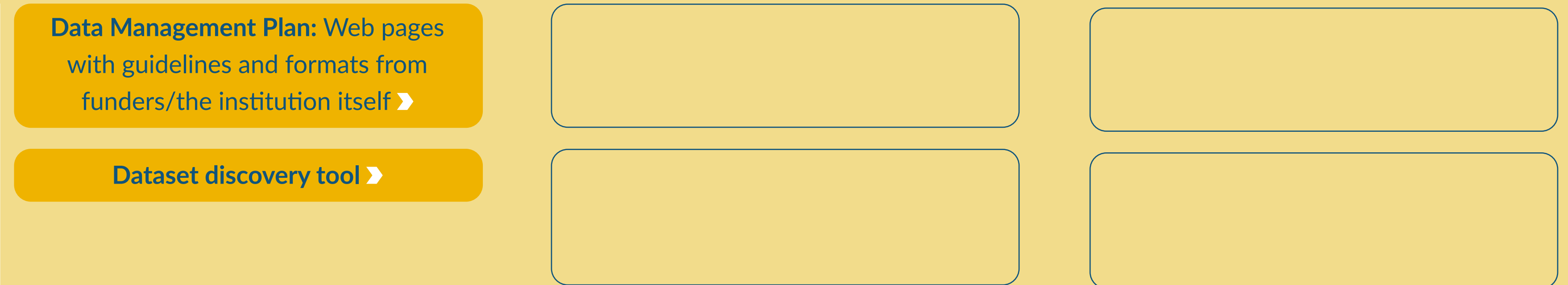


RDM SUPPORT SERVICES

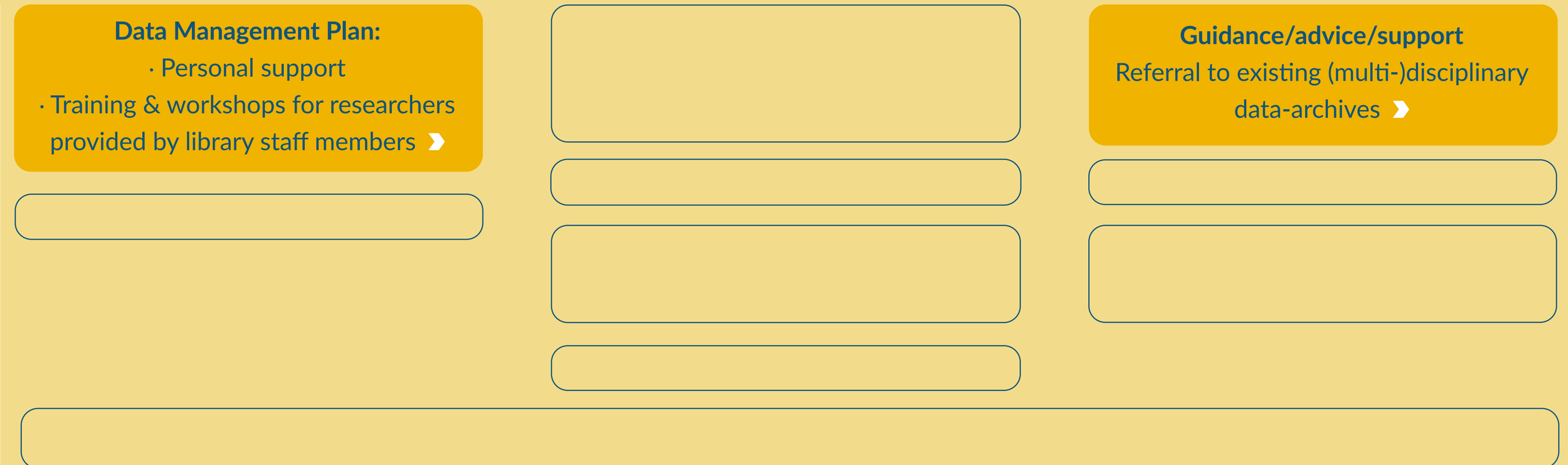


Research funders increasingly demand Data Management Plans for the research projects they finance. This often forms the trigger for research libraries to start developing RDM support services. Another starting point can be a survey among researchers about how they manage their research data now and what kind of obstacles they encounter. Such a survey can help to get a buy-in by the senior management for developing these library services. Some institutions start with formulating a policy on RDM, others develop the services first and formulate a policy later.

INFRASTRUCTURES



SERVICES



## SERVICES

**Data management plans:** Guidelines and formats for data management plans (DMPs) of research funders and/or the institution are presented in web pages on the library website. This can be later developed into a tool for setting up DMPs. Training and advice supplement the service [C1].

### Referral to existing data repositories/data-archives:

Guidance (in the form of webpages and advice) to refer to existing data repositories and data-archives is often part of this initial stage and relatively easy to realise as this easily fits in other referral services from the library [C5].

**Dataset discovery tool:** A discovery tool or portal for datasets is often traditionally part of the services of libraries focusing on Economics and other SSH disciplines. In other cases, such a discovery tool often becomes part of the RDM support services in the developed/advanced stage [C2].

## ORGANISATIONAL STRUCTURES

At this stage, the RDM services are part of the portfolio of information specialists dedicated to supporting researchers. In other libraries, one staff member (data librarian) is dedicated to RDM services. See for further details [A].

## COMPETENCES

For the RDM services at this initial stage, generic, discipline-agnostic technical competences with regard to the entire data life-cycle are necessary:

- This means a gradual upskilling in line with the gradual development path of the RDM services.
- With regard to metadata production, the competences are already available in the library, it only requires an adaptation to other formats and standards for research datasets. See for further details [B].

## REMARKS

### Somewhat hidden RDM services are already there:

Various libraries have described that information specialists and faculty liaison libraries often already supported researchers in individual cases with research data management. Therefore, part of developing the RDM services was to spread knowledge and competences among more colleagues in the library and make it a more official, visible service to the institutional researchers' community.

**Senior management buy-in:** Many libraries stress the importance of a buy-in of the senior management of the institution: this gives the RDM services legitimacy vis-à-vis the other internal stakeholders within the University and vis-à-vis the researchers. It is the task of the library director to discuss the issues around research data management in the institution with the senior management in order to get support to start services in this area. The buy-in by the senior management is often seen as a prerequisite for developing RDM support services. However, some libraries experienced at the start little resonance with regard to the importance of research data management among the senior management

but over the years the management became convinced, and now fully support the RDM support services and the role of the library in it.

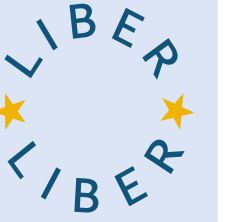
### Added value of RDM support services:

- *The added value from the perspective of researchers is:* (1) research becomes more efficient; (2) continuity in research is better guaranteed by adequate research data management.
- *The added value from the perspective of the institution:* (1) properly conducted research based on research data that can be retrieved in the future; (2) security of research data to prevent data leaks which can lead to reputational damage.

### Institutional data policy at the start or later?

A number of RDM services have started based on the formulation of the institutional policy on research data management by the senior management. This policy statement is often initiated by the library and once an institutional FAIR data policy is in place, this gives the library RDM activities more leverage vis-à-vis researchers. However, others say that with adequate services in place, the implementation of a data policy by researchers is more professional and sustained.

# DEVELOPED STAGE



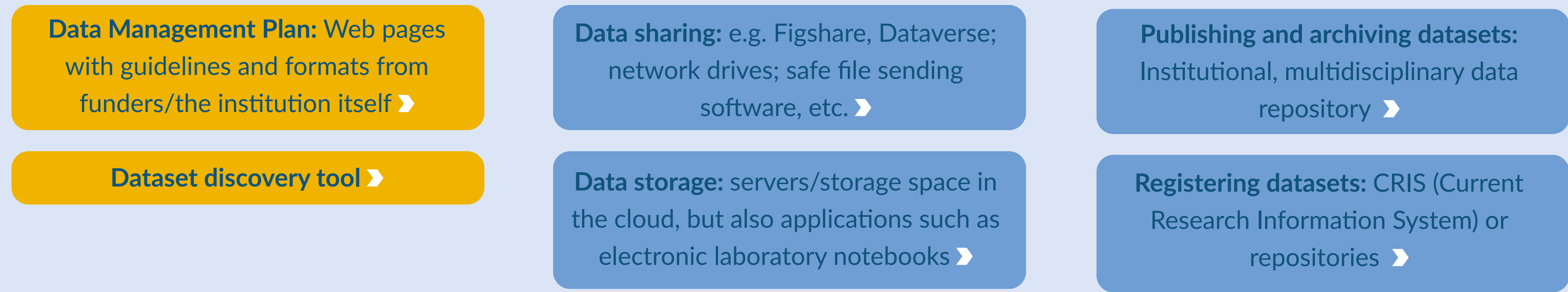
## RDM SUPPORT SERVICES



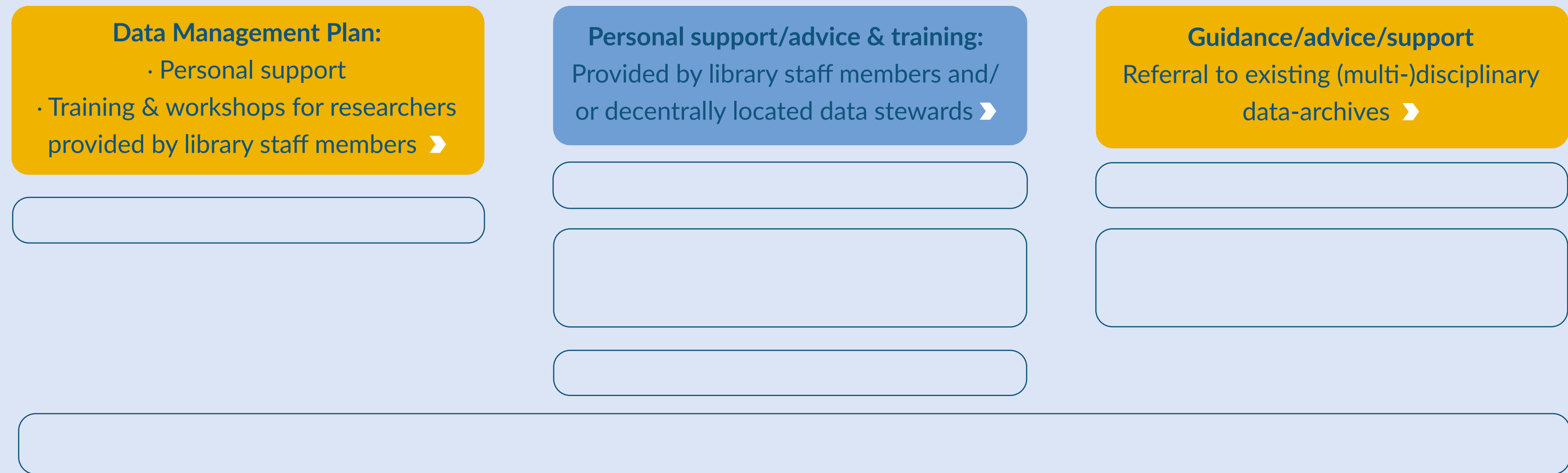
At this stage, the RDM support services are covering the entire data life cycle. Research libraries – in close cooperation with other internal stakeholders such as the IT department – deliver services with regard to data sharing & storage during the execution of the research project and services when the project is finalised.

For the publishing & research datasets many libraries have started an institutional data repository or have an institutional space in a shared data repository.

## INFRASTRUCTURES



## SERVICES



## SERVICES

**Infrastructures and services for data sharing and data storage during the research project:** Without specific RDM services, universities, faculties, and research departments have infrastructures in place for data storage and data sharing for researchers. At this stage of the RDM support services however, a clear overview of the (central) infrastructure is given, in some cases improved, and combined with training and advisory services [C3].

As these infrastructures are often managed by the IT department of the institution, these services are generally delivered in tight collaboration between the library and IT department: sometimes the library is in the lead, sometimes the IT department. See for further details [A].

**Institutional data repository:** An institutional data repository (or an institutional space at a shared data repository) is an important element of this developing stage. With a data repository, the library can offer a solution to the researchers, which triggers the development of various services around the data repository. One of these services involve data curation of the research datasets that are deposited [C6].

· Many European research libraries have a Current Research Information Systems that is used to register and report the research output of the institution. These CRIS systems are adapted so that research datasets can be registered as well: when a research dataset is registered in the CRIS, the dataset can be simultaneously deposited in the data repository [C7].

## ORGANISATIONAL STRUCTURES

The organisational structure for the RDM services in this phase is characterised by a central location in the library and an intensive collaboration with the IT department, the Research Office, Data Protection officers, the Ethics Board, and the Technology Transfer Office. The support professionals of the library often have a coordinating role here to connect all the stakeholders.

Most libraries choose specialisation with the RDM support professionals grouped into units (e.g., Research Data Centre or Digital Competence Centre), or combined with Open Access services in units (e.g., Open Science team, Centre for Digital scholarship, Research Services). Some libraries however use integration as the organisational route. See for further details [A].

## COMPETENCES

In this stage and particularly in the following stages, specialisation in (groups of) disciplines by the RDM support professionals is increasing, meaning that although most competences are still generic and discipline-agnostic, increasingly discipline-oriented RDM competences are needed. See for further details [B].

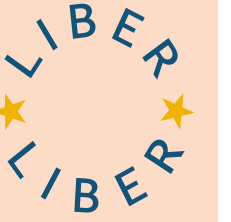
## REMARKS

**Internal Stakeholders – who does what?** RDM support services along the entire data life-cycle require the collaboration between the various stakeholders within a research institution: the IT department, Research office, Data protection officers, Ethics board and Technology Transfer office (TTO). The library often has the coordinating role and has the lead in the RDM services related to preparing and finalising a research project. With regard to the RDM services supporting the execution phase of the research project, at some institutions, the library is in the lead, while at other institutions the IT department is in the lead.

**The infrastructures that underpin the RDM services have to be managed:** The management of the data sharing and data storing infrastructures during the execution of the research projects are often carried out by the IT department, the management of the institution data repository is often in the hands of the research library.



# DEVELOPING STAGE

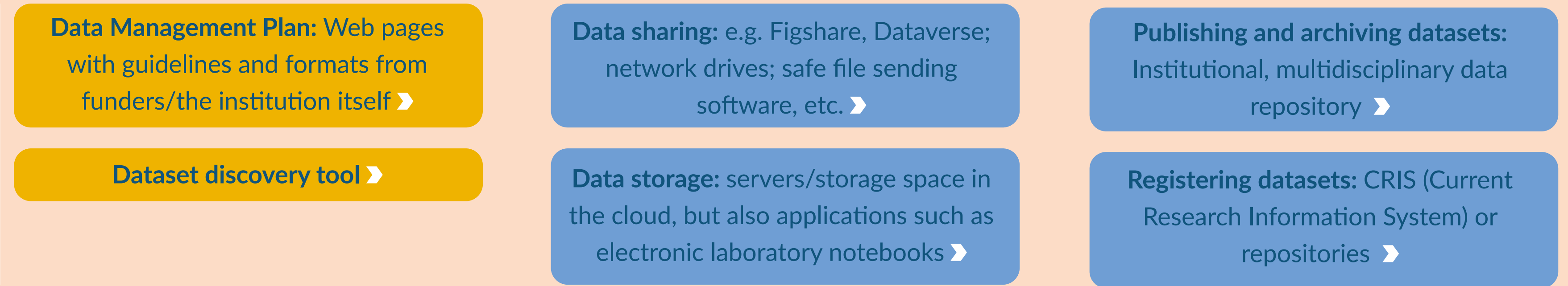


RDM SUPPORT SERVICES

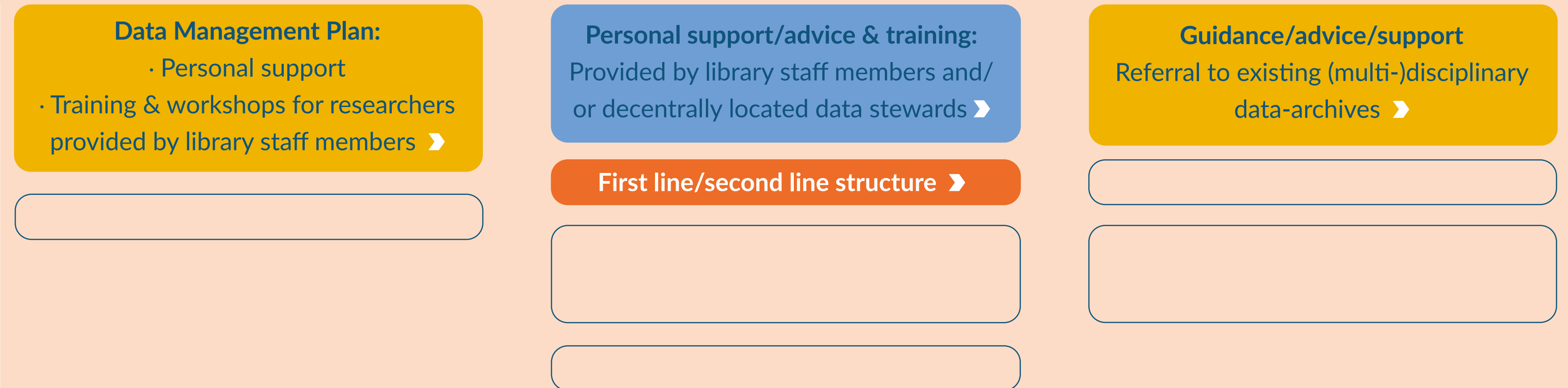


This stage is characterised by developing increasingly discipline-oriented RDM support services. At the larger institutions, this has led to a first-line/second-line service structure: a first-line support with decentrally located data stewards dedicated to a faculty or research department and a second-line support by a centrally located team with a coordinating role. With a growing number of RDM support professionals, community management ensures communication and information exchange.

INFRASTRUCTURES



SERVICES



## SERVICES

**Increasingly discipline-oriented training and advice:** The training and advice services become in this stage discipline-oriented with major implications for the organisational structure of the services as well as the competences needed to deliver these services. At this stage, the discipline-oriented services are mainly limited to advice, while in some libraries, the services at the advanced stage also include operational services.

**Community management:** As the number of RDM support professionals and related individuals within an institution is in this stage growing, a number of research libraries have started to take on a community manager in order to connect and manage the community of these professionals within the institution [C9].

## ORGANISATIONAL STRUCTURES

**First-line/second-line service structure:** Research libraries in larger institutions have generally set up their RDM support services in this development stage with a first-line/second-line service structure:

- **A first-line support** with decentrally located data stewards, often dedicated to and/or embedded a faculty or research department.
- **A centrally located second-line support team** which has a coordinating role and other tasks. See for further details [A].

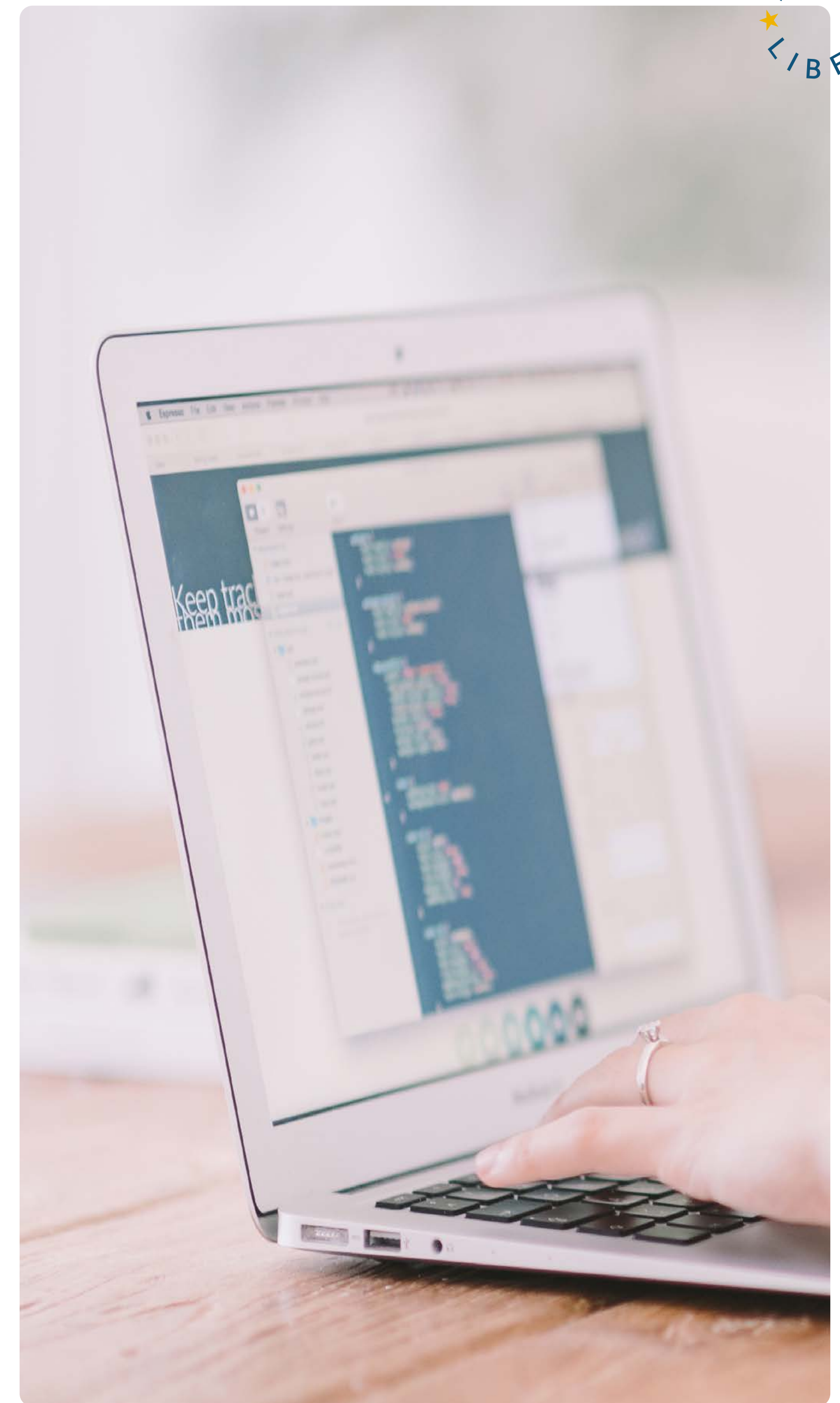
## COMPETENCES

In this first-line/second-line service structure, the staff members in the first-line need discipline-oriented competences and (depending on the nature of the support function) some operational competences. These discipline-oriented competences are rather different and usually do not fit in the competences of 'traditional' librarians. See for further details [B].

## REMARKS

**Embedded data stewards:** The role of data steward embedded at a faculty often takes some time to be accepted and recognised. In the experience of one embedded data steward, there was in the beginning not a good understanding of the role of data stewards among the managers and researchers of the faculty. Therefore, a lot of outreach activities took place to heads of departments and researchers. This led gradually to a mutual understanding of the role, which after more than a year is generally recognised and accepted.

**A more organic approach to embedded data stewards exists as well:** At some institutions, one observes that researchers within research teams get data stewardship tasks. Therefore, these institutions see less need to implement an official organisational structure with decentrally located data stewards. However, as these researchers with data stewardship tasks often are PhD students or early career researchers, the continuity of such an informal data steward network is not guaranteed.



# ADVANCED STAGE



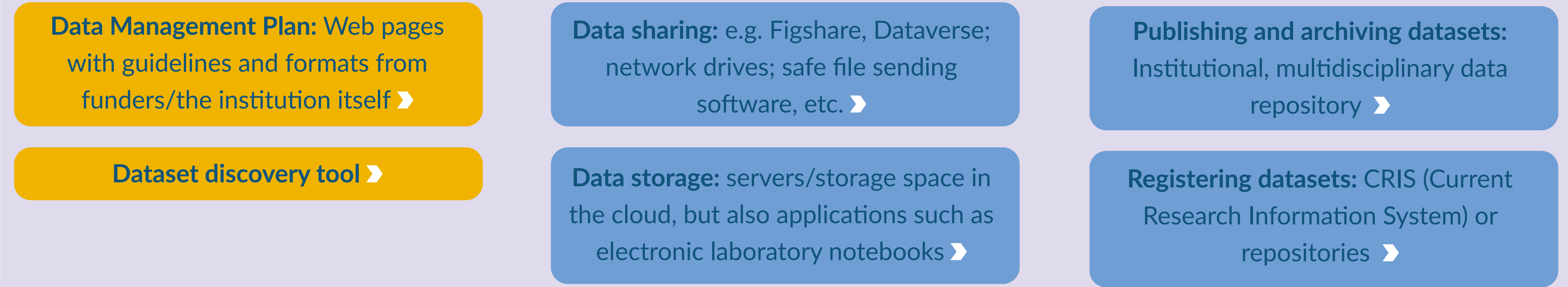
RDM SUPPORT SERVICES



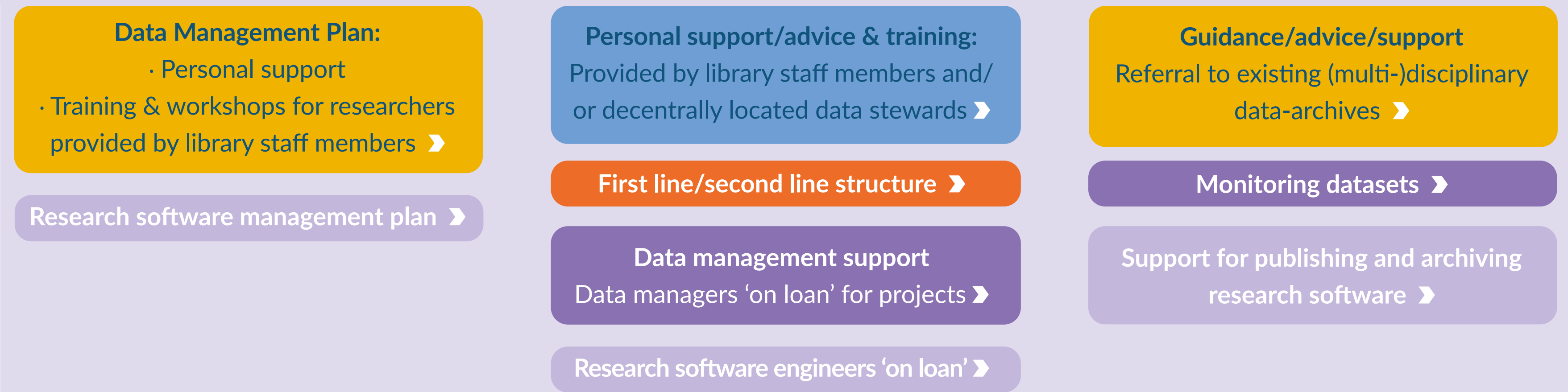
The advanced stage is characterised by:

- Offering operational RDM support services (data management).
- The extension of the RDM services to research software management.
- Monitoring of which data archives the research data sets of the institutional researchers are deposited in.
- At some institutions, the role of the libraries expands to coordinating all Open Science activities.

INFRASTRUCTURES



SERVICES



## SERVICES

**Operational RDM support services:** If the library has services in the domain of research data management, ‘you also get questions from researchers about research data themselves’. This has led to the development of data management services at some libraries with, for example, data managers ‘on loan’ for research teams [C4].

**Monitoring datasets:** There are tools in the market that help identify research datasets deposited in the wide array of data archives by the researchers of a specific institution. This monitoring aids the registration of datasets in a CRIS system but also gives an understanding of to what extent these research datasets are FAIR, which informs the educational activities of the library [C8].

**Research software management:** Services regarding research software management is the ‘next border to cross’ for libraries. Research software that is developed for a specific research project (such as codes or scripts) form an important part of the reproducibility of a study. Therefore, research funders are starting to pay attention to this field, while some libraries and IT departments in the forefront have developed some services in this area [C10].

## COMPETENCES

The data manager’ competences involve extensive data skills and operational know-how regarding research data management [B]. In an institution that also has research software engineers on offer, both work directly with

research teams and deliver practical, operational support to the researchers. Both functions also have a mentorship role for researchers and a role in training them.

## REMARKS

**Bottom-up approach:** A number of libraries in the advanced stage of RDM support services emphasise the importance of a bottom-up approach as part of the community management that is often part of the developed stage. There are several options: one library created a network of researchers who function as data champions who spread the word to their colleagues. Out of this network came the need for an e-lab notebook: an electronic version of laboratory notebooks, which researchers can use to document everything in it from the start of the experiments. Another approach is to connect with researchers who form grassroots Open Science communities themselves, such as groups of researchers who organise Reproducibili Teas at the University of Edinburgh.

**Open Science coordination offices:** At some institutions, the role of the libraries have expanded to coordinating all Open Science activities – top-down initiatives as well as grassroots activities in order to implement Open Science in all its aspects. The aim is to have the Open Science office function as a ‘spin in the web’ which means building networks and participating in networks in the often devolved structures of these universities.

**Future of research support services:** How will the future of research support services develop and what capacity is

needed for it? At one library they have visualised this with the researcher in the middle and in circles the various support functions, such as ethics, software, research data management, policies et cetera. Although researchers will certainly develop more data skills and competences, research support in these various specialised domains will always be necessary and should be offered in an integrated manner instead of via various separate silos.



## Support Services

# [A] ORGANISATIONAL STRUCTURES & JOB CATEGORIES

## Organisational Structures

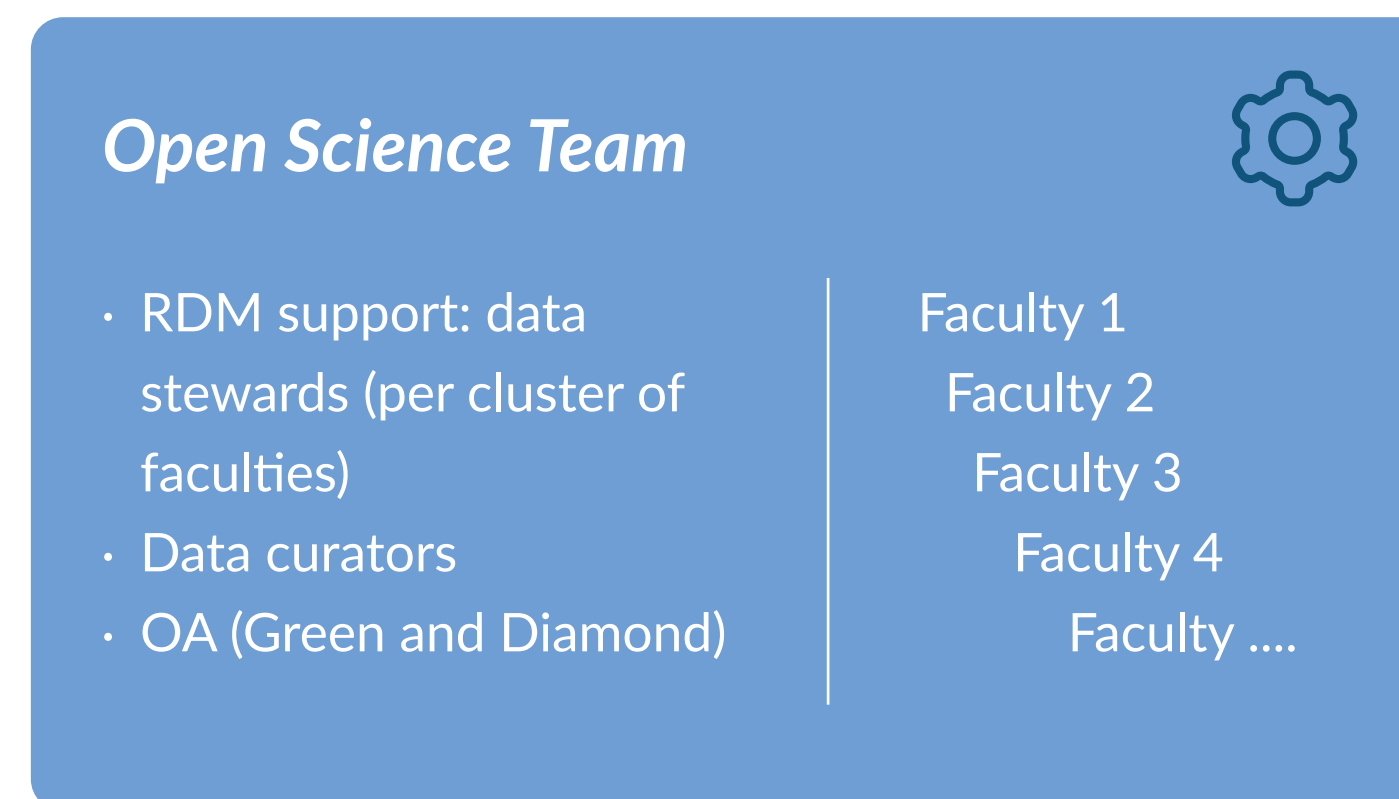
### INITIAL AND DEVELOPING STAGE: SPECIALISATION VERSUS INTEGRATION

Research libraries have organised RDM-services according to the stage of development:

- **In the initial stage**, the RDM services are part of the portfolio of services delivered by librarians who deliver all kinds of research support or by one or two data librarians focusing on the RDM services.
- **In the developing stage**, there are two options for a library organisation:

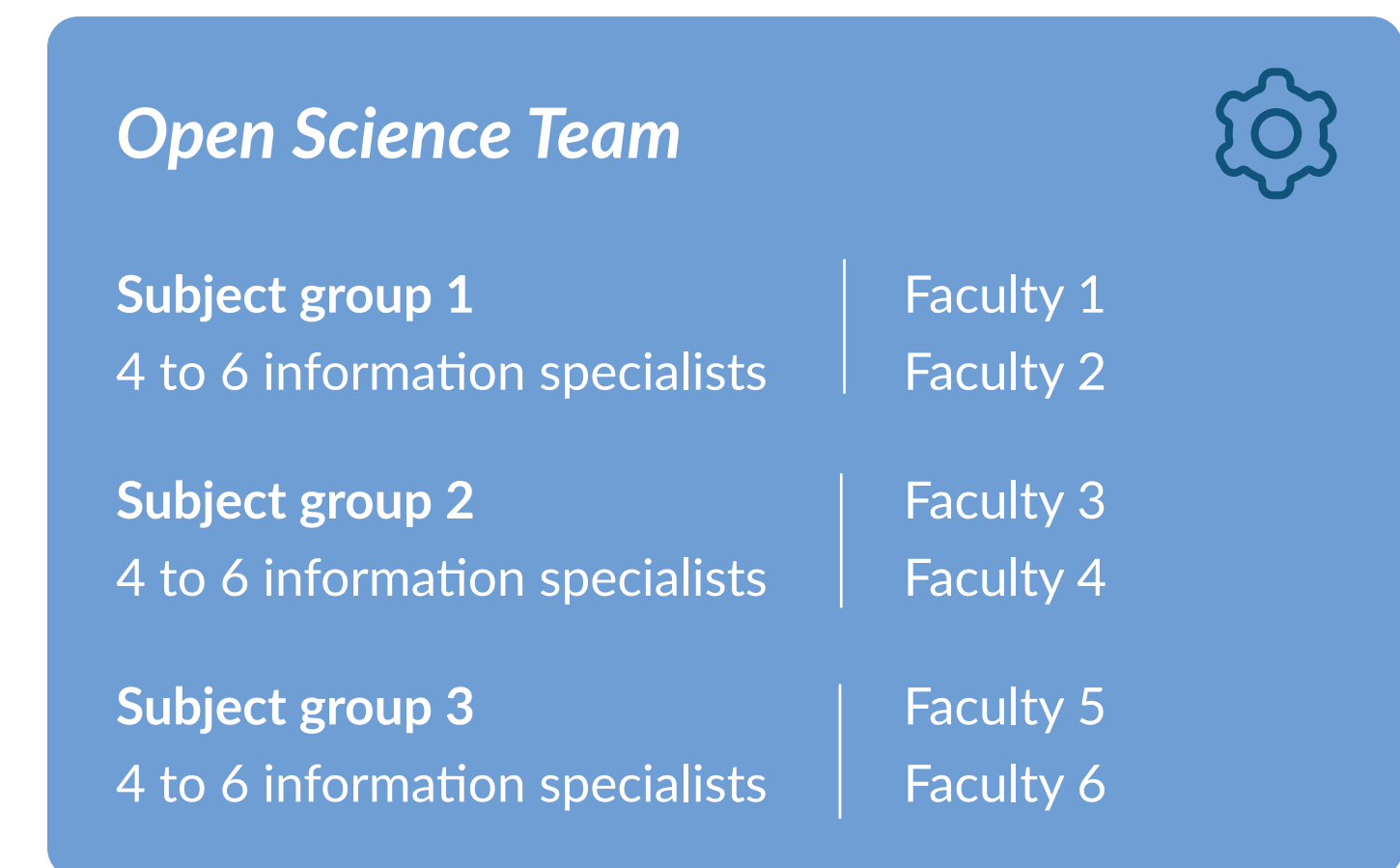
### SPECIALISATION IN A TEAM

The RDM support services are delivered by a team of data librarians/generic data stewards that is centrally located in the library. These data librarians/data stewards are grouped in a unit with names like Research Data Centre or Digital Competence Centre or are combined in units with librarians who deliver Open Access services in units with names like Open Science team, Centre for Digital scholarship, or Research Services. This option seems to be the organisational route chosen by most library organisations.



### INTEGRATION

Another option is to choose integration as the organisational route: in one library, they had a team of 4 to 5 FTE dedicated to Open Science support, but the other staff members of the library were not involved and many of those did not really believe Open Science was a task for libraries. A review of the library organisation led to the recommendation to spread Open Science knowledge among the greater part of the library organisation. This led them to choose to integrate the Open Science services with the other library services: they have formed four subject groups, each around a cluster of disciplines. Each unit has 4 to 6 information specialists who deliver the entire portfolio of library services to the faculties. In other words, Open Science is now integrated in the work of all information specialists, some have developed a deeper knowledge while others only have a basic knowledge.



## DEVELOPED AND ADVANCED STAGES: FIRST-LINE/SECOND-LINE SERVICE STRUCTURE

In the developed and advanced stage, the RDM services become more discipline-oriented, which in the larger institutions has led to the development of a first-line/second-line service structure. This consists of:

### A FIRST-LINE SUPPORT

with decentrally located data stewards/data librarians, often embedded in/dedicated to a faculty or research department. Sometimes these embedded data stewards are employees of the central library, sometimes they are employed by the faculties themselves.

### A SECOND-LINE SUPPORT TEAM

– centrally located – which has a coordinating role and takes care of complex questions that involve many stakeholders. In addition, the central team guarantees the continuity by training newly arrived data stewards, following the evolving data policies and requirements from research funding organisations and creating generic training modules to which faculty data stewards can add disciplinary information.

#### *First line support*

- Embedded data steward(s) at Faculty 1
- Embedded data steward(s) at Faculty 2
- Embedded data steward(s) at Faculty 3
- Embedded data steward(s) at Faculty 4
- Embedded data steward(s) at Faculty 5
- Embedded data steward(s) at Faculty 6



#### *Second line support*

##### RDM support:

- Coordinator
- Generic data stewards
- Data managers



## RDM SUPPORT ECOSYSTEM WITH THE LIBRARY IN THE MIDDLE

The RDM services by libraries in the later stages of development are based on an intensive collaboration with the IT department, the Research Office, Data Protection officers, the Ethics Board, and the Technology Transfer Office. The support professionals of the library often take a coordinating role here to connect all the stakeholders and have to answer complex questions with the aid of a range of specialists from the other internal stakeholders. This collaboration is the key to develop an adequate RDM support ecosystem.

## INVOLVEMENT OF OTHER LIBRARY DEPARTMENTS

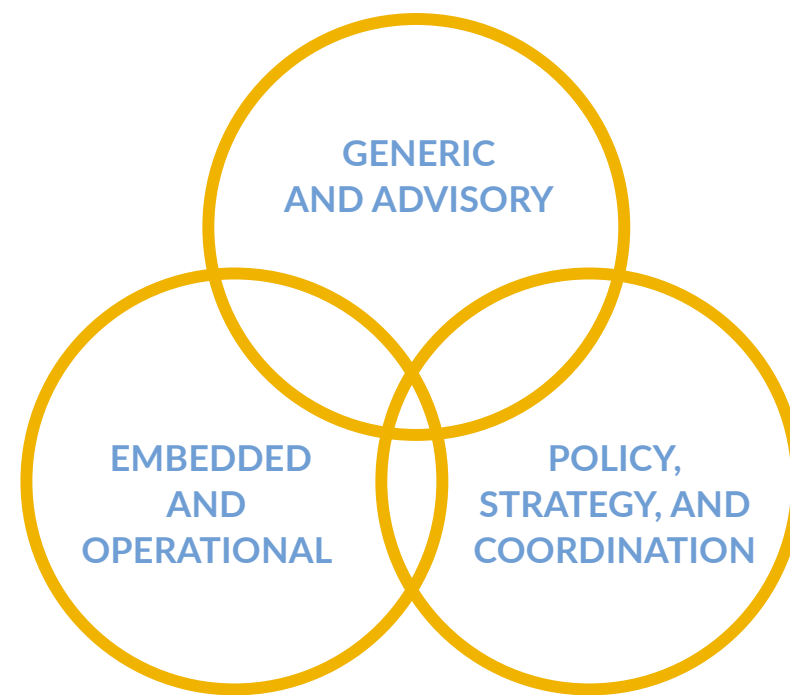
The management of the data repository and the metadata production for the data repository are often carried out by other library departments.



# Job categories

## DATA STEWARD/DATA LIBRARIAN

The roles and tasks of data stewards or data librarians can be divided in three areas:



In practice, the generic and embedded data stewards are often differently positioned in the organisation:

- **The generic data steward or data librarian** is generally positioned within the (central) library and helps researchers with all kinds of data related questions and refers them to others if necessary. He/she supplies information and training with regard to policy requirements and guidelines and helps to draw data management plans. In other words, the generic data steward is a centralised knowledge and communication hub for researchers. The coordinator of those centrally located data stewards generally is more involved in policy-making, while the other data stewards are more involved in delivering consulting services.

- **The embedded data steward** is familiar with the specific needs of fellow researchers within the research unit/ the relevant domain and translates generic data policy to practical implementation. The embedded data steward is often embedded within a faculty or research unit and has knowledge about the type of research data used within the disciplines of that faculty.

Not all libraries have labelled the positions with these tasks as 'data steward' or 'data librarian': they use the more traditional label 'information specialist' or other labels.

## DATA MANAGER

The above-mentioned embedded data stewards have mainly advisory and educational tasks, but in some cases the library services are expanded with data managers in order to take on an operational role within a research project. The aim is to take research data management tasks out of researchers' hands.

## EDUCATOR/TRAINER

Training, workshops, and other educational activities are often carried out by the above-mentioned data stewards. However, at some universities a specific position has been created for this: for example, one library has a coordinator of Digital Skills who organises professional training on a broad area of digital skills, data management, data science and software development. These training programmes are sometimes developed by the library themselves, but also by third parties such as [The Carpentries](#).

## DATA CURATOR

This position is related to a data repository or data archive and also depends on the level of quality assurance of the data repository. After a researcher has uploaded his dataset, the data curator will conduct a number of quality control checks before the database is accepted and available in the data repository. Checks are carried out on the file format, the completeness of the files, the data file contents and structure (such as the English language, tabular datasets with legible headers and labels et cetera), if the data contain sensitive information (personal information et cetera). In addition, the metadata will be reviewed, and additional metadata will be suggested. Finally, the data curator will check and advise on the licence of the dataset. In practice, the position of a data curator is sometimes a separate position, while in other cases these tasks are carried out by data stewards.

## COMMUNITY MANAGER

Especially within the larger research institutions with numerous different specialists within the RDM support ecosystem, libraries have created a position of a community manager in order to engage various groups related to RDM within the institution. The focus of a community manager is to engage researchers and data support professionals about data management and to bring discipline-specific communities together to stimulate the creation of FAIR research data.

GENERAL	1. Knowledge about the structures of higher education and about how research and researchers work (a background in research might be required)
PLANNING AND DESIGN	2. Composition of research datasets 3. Data policies (institutional, funder, national) 4. Fair data principles 5. Data life cycle 6. Data Management Plans 7. Research integrity, ethics and GDPR regulations 8. Pseudonymization and anonymization of data 9. Knowledge of the landscape of data repositories and data archives 10. Metadata of datasets
DATA COLLECTION AND MANAGEMENT	11. Data creation and interoperability 12. Searching data sources
DATA DESCRIPTION	13. Metadata for digital collections and datasets
DATA FORMATTING AND STORAGE	14. Institutional infrastructures for storing & saving data
DATA QUALITY ASSURANCE	15. Knowledge about cleaning data 16. Data curator competences [related to (2) and (10)]
DATA PROCESSING AND ANALYSIS	17. Data analysis and visualisation
DATA ARCHIVING	18. Archiving and preserving data
PUBLISHING AND DISCOVERABILITY	19. Copyright and intellectual property 20. Licensing of research datasets 21. Reproducibility and reuse of data 22. Data discovery tools

## THREE CATEGORIES OF COMPETENCES

Generally, one can distinguish between three categories of competences:

- Knowledge
- Know-how
- Interpersonal skills/soft skills

## TECHNICAL COMPETENCES: KNOWLEDGE AND KNOW-HOW

In the table the technical competences related to RDM support services offered by libraries are listed. One can distinguish three levels in these technical competences:

- **Knowledge – generic, discipline-agnostic competences:** For the RDM support services in the initial and developing stages, generic and discipline-agnostic knowledge of the research data management topics as mentioned in the table. In the developed and advanced stages of the RDM support services, these generic and discipline agnostic competences relate specifically to the centrally located RDM support professionals, who are often part of the second-line support if there is a first-line/second-line service structure. Their tasks are mostly advisory and educational.

- **Knowledge and a bit of know-how – discipline-oriented RDM competences:** In the developed and advanced stages, the RDM support services also involve discipline-oriented knowledge with sometimes a bit of know-how of the research data management topics as mentioned in the table. These competences are relevant for the first-line RDM support professionals (embedded data stewards) who work for specific faculties and/or research departments. Their tasks are mostly advisory and educational but sometimes do involve some operational tasks: these technical, operational skills relate mainly to metadata and depositing procedures.
- **Know-how – operational competences:** These operational competences are crucial for the RDM support professionals / data managers who are (temporarily) part of a research team and carry out operational tasks. The data curators who curate the research datasets for an (institutional) data repository also need technical, operational skills in metadata production and technical insight in the composition of research datasets.



## INTERPERSONAL SKILLS/SOFT SKILLS

COMMUNICATION SKILLS (INCLUDING A CERTAIN ROBUSTNESS WHEN ENCOUNTERING RESISTANCE/RELUCTANCE FROM RESEARCHERS)
ADVISORY SKILLS
CONNECTING ROLE, ABILITY TO BUILD AND MAINTAIN/MANAGE NETWORKS
SERVICE ORIENTATION
PEDAGOGICAL COMPETENCES
ABILITY TO TRANSLATE PRACTICAL ISSUES INTO PROTOCOLS, SERVICES, AND POLICIE

The table lists Interpersonal skills/soft skills that are important for those categories of RDM support professionals who are in regular contact with researchers and other internal stakeholders.

### A CLOSER LOOK AT THE COMPETENCES OF DATA STEWARDS

In practice, the generic data stewards are often library staff members who have grown into this role and gathered the discipline-agnostic knowledge about the data life-cycle of research data by 'learning on the job'.

The discipline-oriented competences needed to fulfil the role of embedded data stewards appears more difficult for 'traditional' library staff members to fulfil. For this job category, libraries recruit personnel outside the library domain and look for:

- 1) Research background, for example a PhD
- 2) Communication and interpersonal skills
- 3) Data skills & experience/knowledge about research data.

This combination is difficult to find as there are few data steward courses available at the moment. In order to fulfil the role of embedded data steward, 1) a research background, and 2) good soft skills are essential: without a research background and good soft skills, it is difficult to make a good start in this position (you need to understand how research and researchers work) and you cannot adequately communicate with the researchers. However, the more technical competences mentioned under (3) can be learnt on the job if necessary.

### ACQUIRING COMPETENCES FOR RDM SUPPORT SERVICES

An often-expressed concern is that it is difficult to get the competences for RDM support services as there is hardly any formal training because this domain is relatively new. However, many librarians involved in developing these services are practical about this:

- The competences needed for discipline-agnostic RDM support services can be learnt on the job by librarians if the will and the curiosity is there. One librarian stated that 'you have to make steps', but 'it is not a different world'. They gradually developed their competences by attending professional conferences and liaisons with other libraries.

At another library, they started the RDM support services with self-study, group discussions and meetings before outreaching to the researchers. The role of LIBER and ADBU in exchanging experiences among professionals is seen as important in this.

- For the discipline-oriented RDM support services, a research background in the discipline is generally seen as essential while knowledge and skills regarding RDM could eventually be learnt on the job (see above).
- Increasingly, manuals, handbooks, and training courses on RDM developed:
  - How to be FAIR with your data. [A teaching and training handbook](#) for Higher Education institutions [FAIRSF AIR]
  - [Practical Guide](#) for Open Badges in RDM/Open Science [OBERRED]
  - The project [Skills4EOSC](#)
  - LIBER Research Data Management [Working Group](#) for exploring and developing RDM services and exchanging Best Practices
  - The [commission](#) Recherche et Documentation of ADBU.

# [C] RDM SERVICES



## [1] Data Management Plans

- **DMP development:** Research libraries help researchers develop data management plans as these are increasingly mandatory by research funding organisations and sometimes by the institution itself. Guidelines and formats are presented at dedicated webpages. When the service in this domain develops further, a data management plan tool is provided: with such a tool, the researcher has to fill out questions and the tool provides the researcher with (an outline of) a relevant data management plan.
- **Training and advice:** After developing the web pages, libraries generally follow up with providing introductory courses about data management plans. In next steps, the library starts providing personal advice, guidance, training, and consultancy on data management plans.
- **Quality control:** Some libraries have developed workflows to ensure that the proposal's data management plan of the researchers is first discussed with a data librarian or data steward before the submission. The aim is to ensure a good quality of data management plan in the grant proposals so that there is a higher chance of obtaining the grants.

## [2] Dataset discovery tool or portal

- **Economics:** In some disciplines such as Economics and other Social Sciences and Humanities, libraries traditionally have acquired licences to datasets and therefore offered a discovery tool or portal to find these datasets. For example, a research library specialised in Economics traditionally offers access to 90 datasets, about 80% of them are licensed datasets and about 20% are public datasets. This library also acquired experience in negotiating access to specific datasets for one researcher or research group. In short: for these disciplines, datasets for researchers were traditionally part of the library services and a portal or discovery tool for datasets already existed before the development of RDM support services.
- **Other disciplines:** In other research libraries however, offering a discovery tool for research datasets is often part of later stages in the development of RDM services.

## [3] Infrastructures and services for data sharing and data storage

It is important to note that in an institution without RDM support services, infrastructures for data storage and data sharing are of course provided by the central IT department, by faculties or research departments themselves. However, with the start of RDM support services, a clear overview of the (central) infrastructure is given and in some cases improved with services such as Figshare, Dataverse or electronic laboratory notebooks. These infrastructures are combined with training and advisory services.

In many cases, the infrastructures are provided by the IT department of the institution and therefore, the services around these infrastructures are generally delivered in tight collaboration between the library and IT department: sometimes the library is in the lead, sometimes the IT department.

## [4] Data management support

If the library has services for RDM, you also get questions about data. Some libraries have therefore expanded their services with data support or data managers 'on loan' for research projects. The data managers carry out operational tasks regarding research data management for specific research teams.

## [5] Guidance and referral services to existing (multi-disciplinary) data archives

In the first stages of RDM support services, a referral service consisting of webpages and advice with an overview of disciplinary and multi-disciplinary specialised data archives where researchers can deposit their research datasets can be relatively easily developed, based on existing knowledge of existing library staff members such as liaison or faculty librarians, of the institution researchers and on existing overviews of other research libraries.

This service easily fits in other referral services of the library. See also [8] Monitoring datasets.

## [6] Institutional data repository

- **Publishing and/or archiving:** After the research project, the FAIR or Open research datasets can be published, archived, and registered (see [7]). Increasingly, institutional data repositories are set-up and offered to researchers alongside existing, often disciplinary data-archives. There are also alternatives to an institutional repository, such as an institutional space in a national repository (for example recherche.data.gouv in France) or an institutional space in a generic repository such as Zenodo. Libraries without an institutional data repository or an institutional space in a generic repository offer generally a referral service, advising researchers to find their way in the landscape of data archives.
- **Data vault:** In addition to a data repository, institutions develop special scrutinised data repositories for sensitive research data, called a data vault or a data safe haven.

## [7] Registering datasets

Many libraries maintain a Current Research Information System (CRIS) with the aim to register all the research output of the institutional researchers. In order to register the research datasets as well, these CRIS systems had to be adapted. In most CRIS systems, the registering of datasets can now be done alongside the registration of the publications. In most cases, the dataset registration involves only the metadata and/or a description of the dataset, while the dataset itself is (in many cases automatically) deposited in the institutional data repository.

## [8] Monitoring datasets

One library uses a data monitor in order to identify datasets by their researchers that are deposited in data archives and/or data repositories. With this data monitor, they could increase the number of datasets from researchers of the University that were documented in their CRIS system by a factor eight.

In addition, the results gave them the following insights:

- The data repositories that are most relevant for their researchers
- Insight into what extent the deposited datasets are FAIR. This informed their FAIR education strategy.
- A good overview of datasets produced by research groups and a good accessibility of those datasets is seen as a plus in evaluations of research groups and researchers.
- The information gathered from this helps them to form their future institutional data policies.



## [9] Community management

Within the larger research institutions with numerous different specialists within the RDM support ecosystem, libraries have started-up community management. Sometimes this is part of the task of one RDM support professional, sometimes the position of a community manager is created. The aim of community management is to engage researchers and data support professionals about data management and to bring discipline-specific communities together to stimulate the creation of FAIR research data. This involves both top-down and bottom-up approaches.

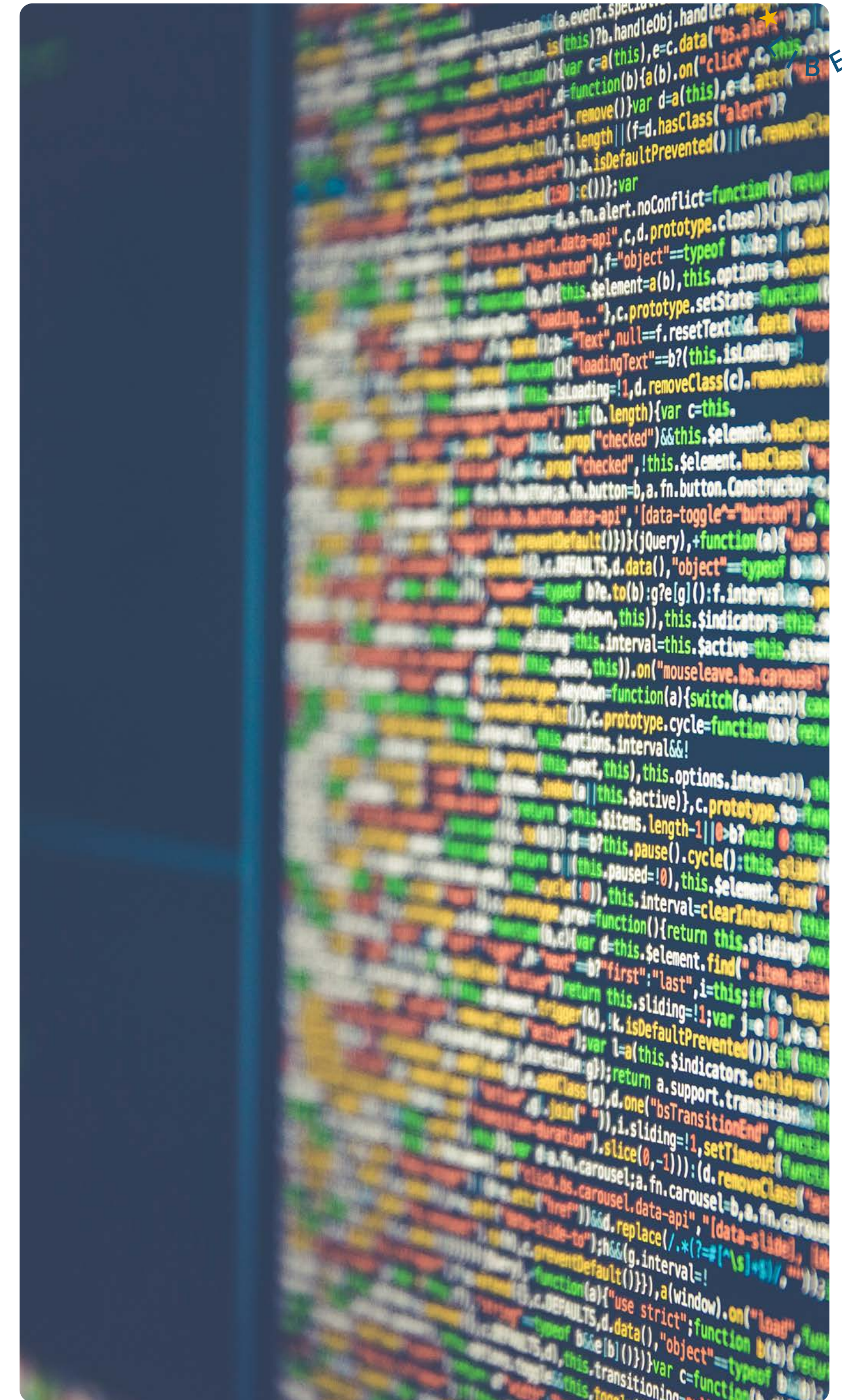
With regard to the bottom up approach, one library has created a network of researchers who function as data champions who spread the word to their colleagues. Another approach is to connect with researchers who form grassroots Open Science communities themselves, such as groups of researchers who organise Reproducibili Teas at the University of Edinburgh.

## [10] Research software

Services regarding research software management is the 'next border to cross' for libraries. Research software that is developed for a specific research project (such as codes or scripts) form an important part of the reproducibility of a study. Therefore, research funders are starting to pay attention to this field, while some libraries and IT departments in the forefront have developed some services in this area.

Services in this area are the 'next border to cross' for libraries: these services have to be developed in close cooperation with the IT department and are still in an experimental phase. At a few institutions, one has made a start:

- A template for a [research software management](#) plan has been developed.
- Research software engineers 'on loan' for research teams. These software engineers focus on the coding of software, scripts used, et cetera. In this institution, the research software engineers are employed by the IT department and their activities are coordinated with the data managers 'on loan' for research teams who focus on research data management tasks.



# [D] WHY THE LIBRARY?

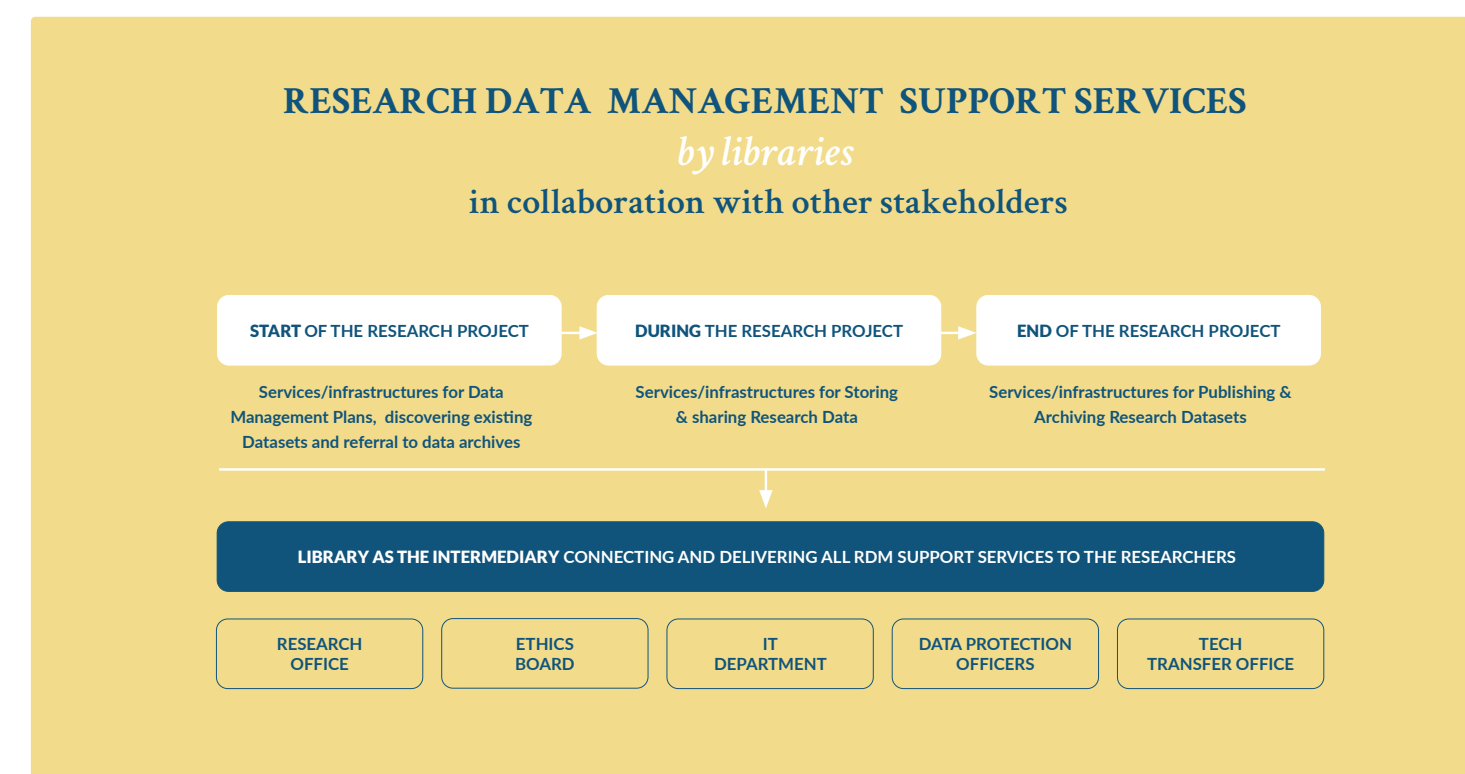
RDM support services cannot be delivered by one internal stakeholder: there are specialists involved from the IT department, Research Office, Data protection Officers, the Ethics Board, and the Technology Transfer Office. This means that libraries can play a coordinating and intermediary role between these departments on the one hand and the research community.

On the other hand: Libraries have a connecting, coordinating, and intermediary role for RDM services because:

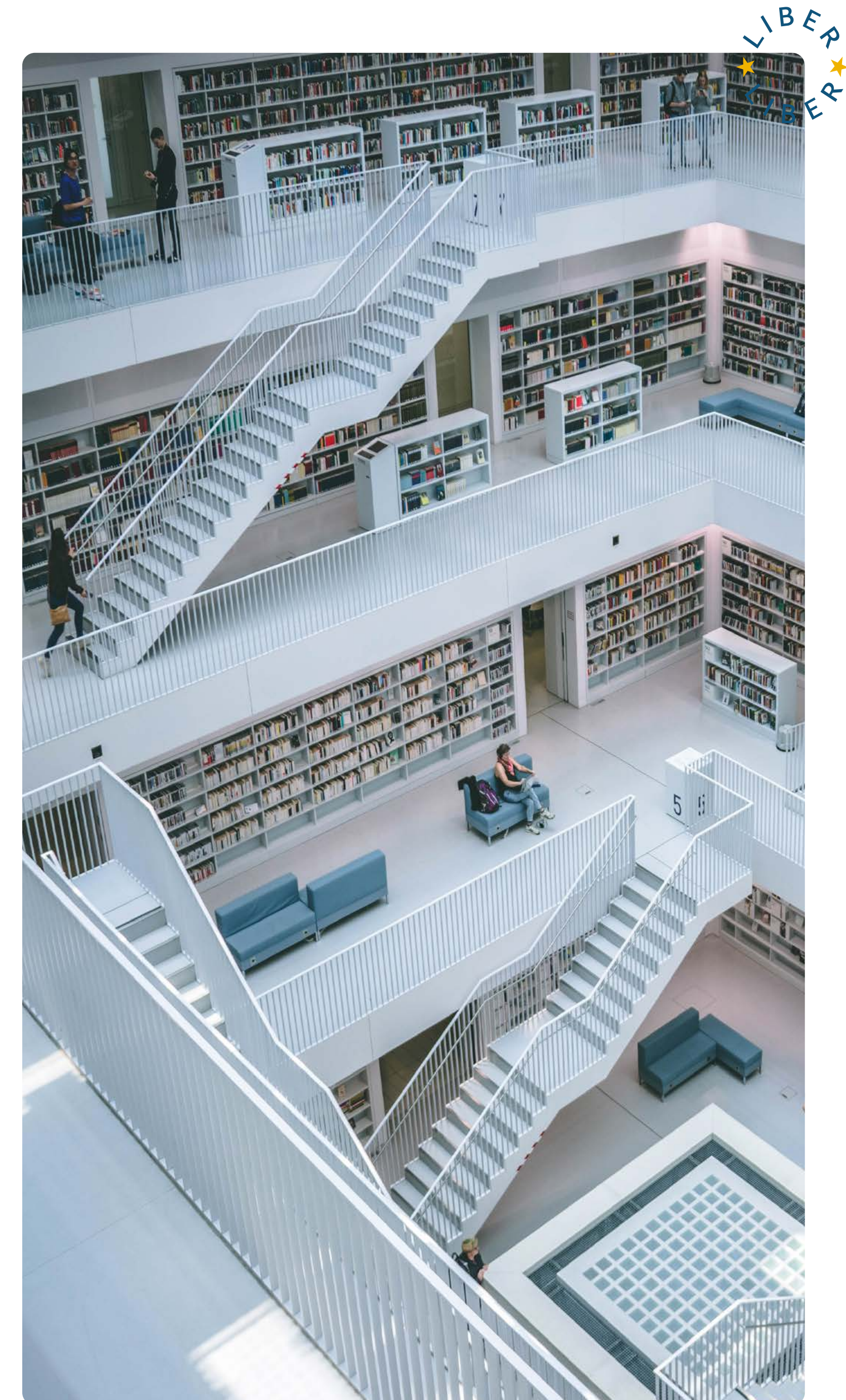
- Libraries have a good understanding of the context in which research and researchers operate.
- Libraries are the logical contact point for researchers
- Libraries are in the middle of things and 'know a bit of everything'

In addition, libraries bring relevant expertise to the table:

- Libraries have a comprehensive overview of the regulation and policy environment of institutions and research funders with regard to Open Science, Open Access, FAIR research data, research integrity, ethics et cetera.
- Libraries have technical expertise regarding digital publishing, metadata standards, ontologies, licenses, copyright, and identifiers.
- Libraries already support research output in the form of publications, which makes them a suitable party to support research output in the form of research datasets.
- Libraries have expertise with regard to long-term preservation of digital research outputs.



[Click for bigger version](#)



# ABOUT THIS TOOLKIT

## ABOUT THE STUDY

Many research libraries in Europe deliver Open Science services in the field of Research Data Management support services and Open Access services. However, it is estimated that up to half of the European research libraries deliver only limited services in these domains. This LIBER/ADBU study focuses on understanding the organisational structures and competences needed for these services. The aim of the study is to support an acceleration of the adoption process of these two Open Science services by research libraries throughout Europe.

The study was commissioned by [LIBER](#) (the Association of European Research Libraries) and by the [ADBU](#) (the Association of directors of University libraries in France) and took place in the period May 2022 to March 2023.



The study was carried out by Maurits van der Graaf (Pleiade Management and Consultancy), accompanied by an expert group consisting of Christine Okret-Manville (Université Paris-Dauphine-PSL); Thomas Chaimbault (Enssib); Thorsten Meyer (ZBW); Minna Niemi-Grundstrom (Helsinki University); Göran Hamrin (KTH Royal Institute of Technology) and Astrid Verheusen (LIBER) during the entire study period as well as the participation of Sophie Valade (ADBU), Sophie Forcadell (SciencesPo), Veronique Stoll (Observatoire de Paris) for parts of the study period.

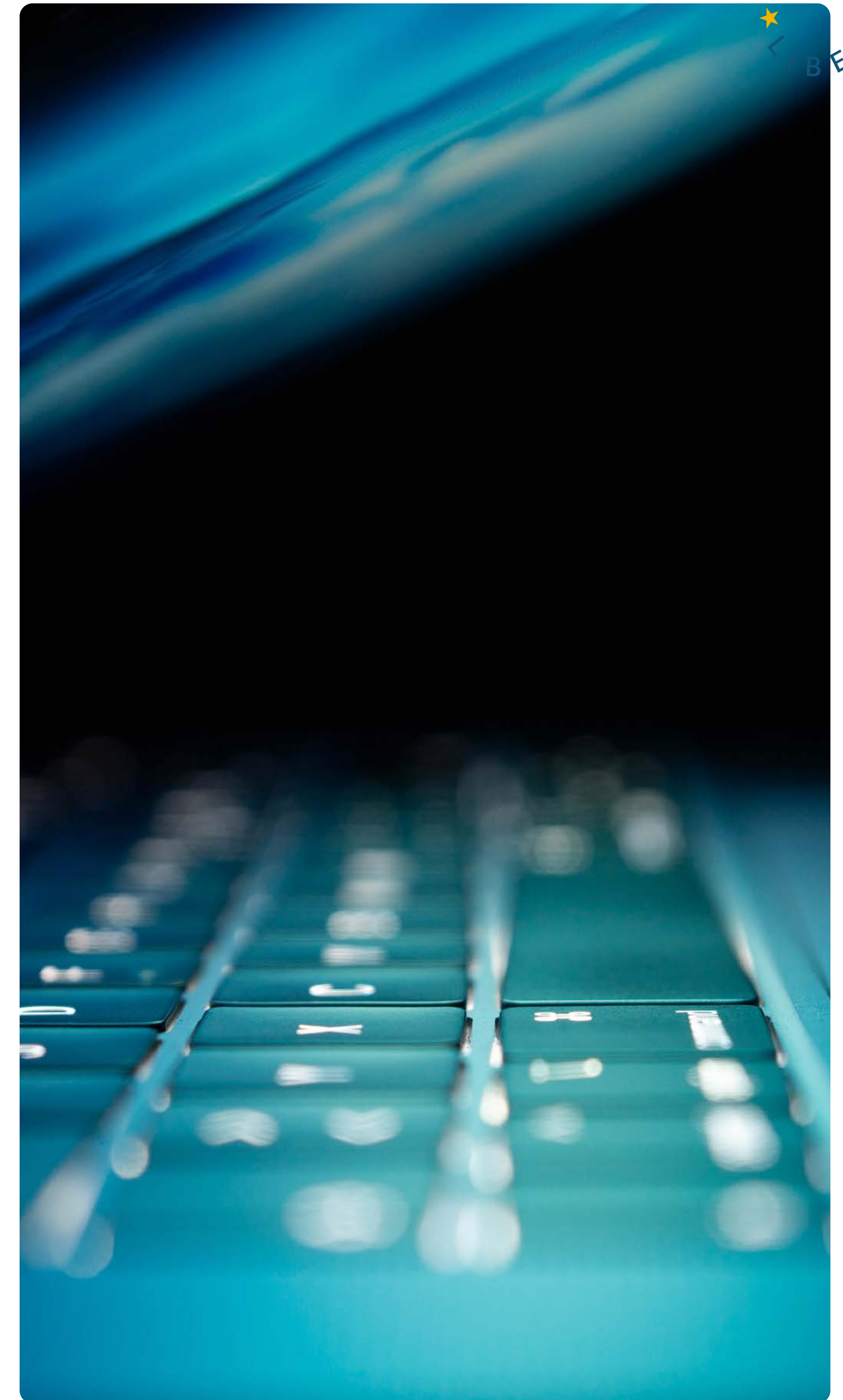
The results of the study are published in the report: Open Science Services by Research Libraries: Organisational perspectives; a LIBER and ADBU [report](#); Maurits van der Graaf; 2023.

## ABOUT THIS TOOLKIT

This toolkit has been developed based on the results of the above-mentioned study. The toolkit has been discussed within the expert group and has been tested in a webinar with ADBU members. This interactive PDF has been designed by Spresso.

## ACKNOWLEDGEMENTS

We would like to thank all participants in the ADBU webinar of 13 April 2023 for their feedback.





Ligue des Bibliothèques Européennes de Recherche  
Association of European Research Libraries

[www.libereurope.eu](http://www.libereurope.eu)



[www.adu.fr](http://www.adu.fr)

