OpenAIRE National Workshop

«OpenAIRE: uno strumento per la Scienza Aperta in Europa»

30-31 maggio 2016

CNR - Roma

The data pilot in practice: the experience of the University of Bologna

Marialaura Vignocchi, Irene Frascari, Roberta Lauriola, Milena Garofalo, Marilena Daquino, Raffaele Messuti

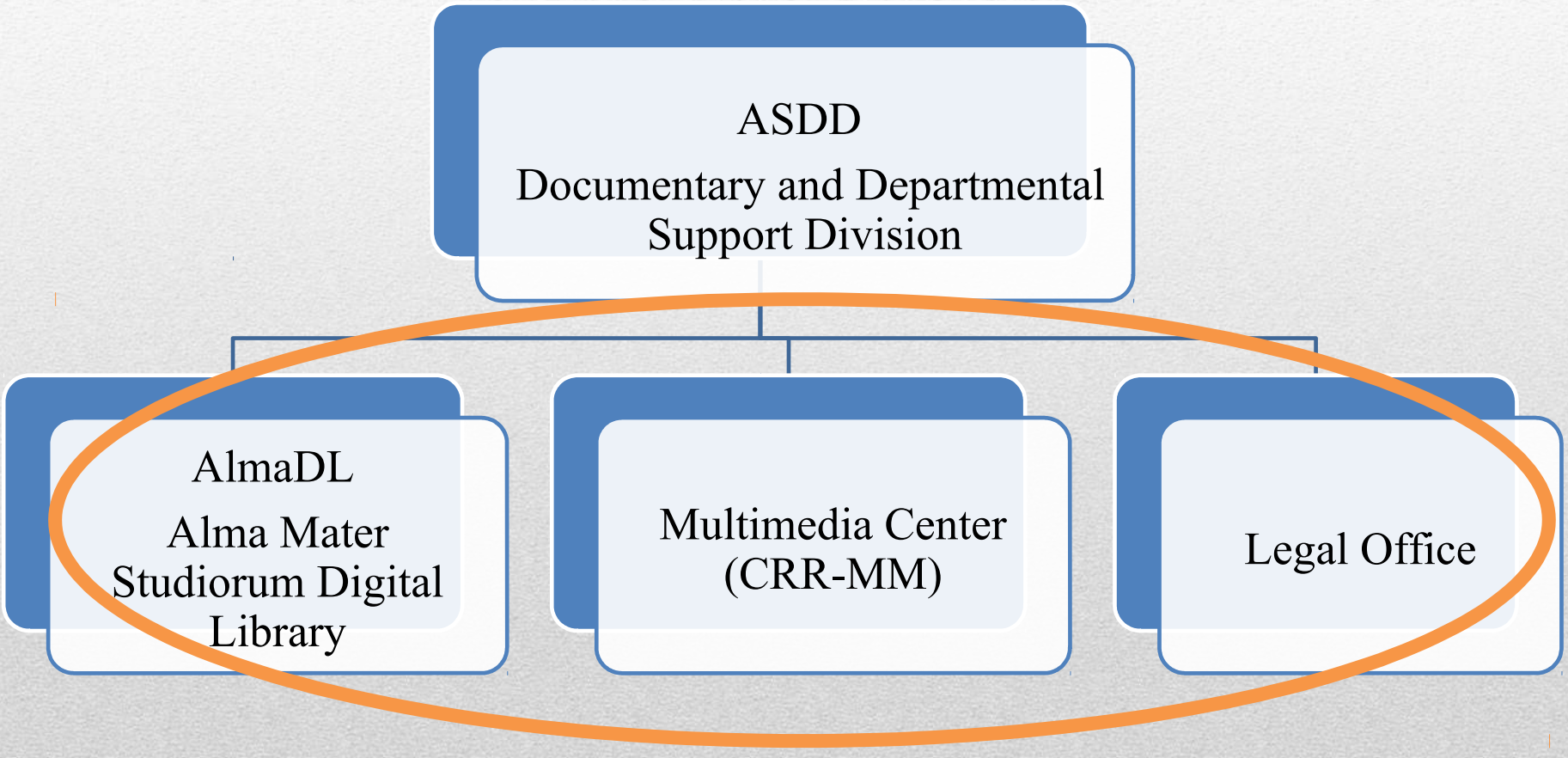
ALMA MATER STUDIORUM – Università di Bologna

Area Sistemi Dipartimentali e Documentali

Centro di Risorse per la Ricerca Multimediale (CRR-MM)

AlmaDL (Alma mater Studiorum Digital Library)

Who we are


ASDD
Documentary and Departmental Support Division
AlmaDL
Alma Mater Studiorum Digital Library
Multimedia Center (CRR-MM)
Legal Office

ALMA MATER STUDIORUM UNIVERSITÀ DI BOLOGNA
SISTEMA BIBLIOTECARIO DI ATENEO

HOME | CHI SIAMO | BIBLIOTECHE | SERVIZI | AlmaRE | AlmaDL | CATALOGHI

HOME /

ALMADL

AlmaDL è la Biblioteca Digitale dell'Ateneo. Raccoglie, conserva e rende disponibili in rete collezioni digitali a supporto della didattica e della ricerca.

CHE COS'È ALMADL

- Obiettivi e finalità
- Principali attività
- Staff tecnico
- Un po' di storia

ALMADL PER L'OPEN ACCESS

- Cosa significa Open Access
- Valorizzare la ricerca
- Condividere i dati di ricerca
- Depositare nell'archivio istituzionale
- Ottenere finanziamenti per il Gold Open Access

COLLEZIONI

- AMS Campus - Materiali didattici
- AMS Acta - Contributi di ricerca
- AMS Tesi di Dottorato
- AMS Tesi di Laurea
- AlmaDL Journals - riviste open access
- AMS Historica

SEGUICI SU

AMS HISTORICA - LIDEL

MULTIMEDIA CENTER

AlmaDL (Digital Library)

Mission and activities:

AlmaDL, together with the Web Technologies Office of the UniBo ICT Division (CeSIA), provides services and infrastructures for the preservation, dissemination and access of institutional collections of digital contents. It manages the **institutional repositories of:**

- **research publications and data (OpenAIRE 3.0 compliant)**
- doctoral theses, Master, Bachelor and student's theses
- educational resources
- digital editions of ancient and rare library collections
- peer-reviewed open access journals and series

It offers support for the ***7FP post-grant Open Access Pilot*** and ***H2020 Open Access mandate***

Staff: the AlmaDL team includes librarians, the institutional repository manager, ICT technicians, a research data specialist. The IP expert of our Division collaborates and provides legal advice. ICT technicians of the University ICT divisions support the technological infrastructure.

AMSActa
Institutional Research Repository
by AlmaDL University of Bologna Digital Library

AlmaDL
AlmaDL Search

it | en
Statistics Login
BROWSING TOOLS

Fall Risk Assessment Tools for Elderly Living in the Community: Can We Do Better?

Palumbo, Pierpaolo ; **Palmerini, Luca** ; **Bandinelli, Stefania** ; **Chiari, Lorenzo** (2015) *Fall Risk Assessment Tools for Elderly Living in the Community: Can We Do Better?* PLOS ONE, 10 (12). ISSN 1932-6203

Save citation | Share | Cited by

Full text available as:

PDF (Main text)
Licence: Creative Commons Attribution Non-commercial
Download (626kB) | Preview

Official URL: http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0146247

Abstract

Background Falls are a common, serious threat to the health and self-confidence of the elderly. Assessment of fall risk is an important aspect of effective fall prevention programs. Objectives and methods In order to test whether it is possible to outperform current prognostic tools for falls, we analyzed 1010 variables pertaining to mobility collected from 976 elderly subjects (InCHIANTI study). We trained and validated a data-driven model that issues probabilistic

ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

CENTRO DI RISORSE PER LA RICERCA MULTIMEDIALE - CRRMM

Cerca

- Home
- Centro
- Attività
- Competenze
- Portfolio

Contatti

Centro di Risorse per la Ricerca Multimediale - CRR-MM

Viale Filopanti, 7

40126 - Bologna

Scrivi una mail

tel: +390512094269

MULTIMEDIA CENTER
University of Bologna

Chi siamo

Il Multimedia Center CRR-MM è un Centro di Risorse per la Ricerca Multimediale che si occupa di aiutare docenti e ricercatori dell'Università di Bologna nell'ambito dell'attività di ricerca. Offre servizi per la produzione, preservazione, diffusione e fruizione di materiali multimediali.

In primo piano

Open research data: condivisione della ricerca e scienza aperta

Quale valore ha la condivisione dei dati nel contesto della Open Science? Come posso rendere aperti i miei lavori e i miei dati? Come si compila un Data Management Plan? Il 18 e 19 novembre 2015, presso il Centro Congressi del CNR a Bologna due giornate di formazione e dibattito sull'Open Research data.

EDITEI

Ambiente web per l'editing XML/TEI, inizialmente progettato per l'insegnamento, è un'agile soluzione per progetti collaborativi avanzati.

ASFE

Banca dati prosopografica della popolazione studentesca dell'Università di Bologna dal 1500 al 1800

Eventi

Non ci sono Eventi

Vedi tutto

Link utili

- Università di Bologna
- CRR-MM - english version
- AlmaDL
- Brown University

UNIBO Multimedia Center (CRR-MM)

Mission and activities:

CRR-MM provides services to the researchers of the University of Bologna for the production, processing, organization, storage, preservation and dissemination of digital contents such as collections of data, texts, images, audiovisual materials and sound recordings, animation, etc.

It also supports the development of digital tools applied to humanistic inquiry.

It offers assistance for H2020 Open Research Data Pilot requirements.

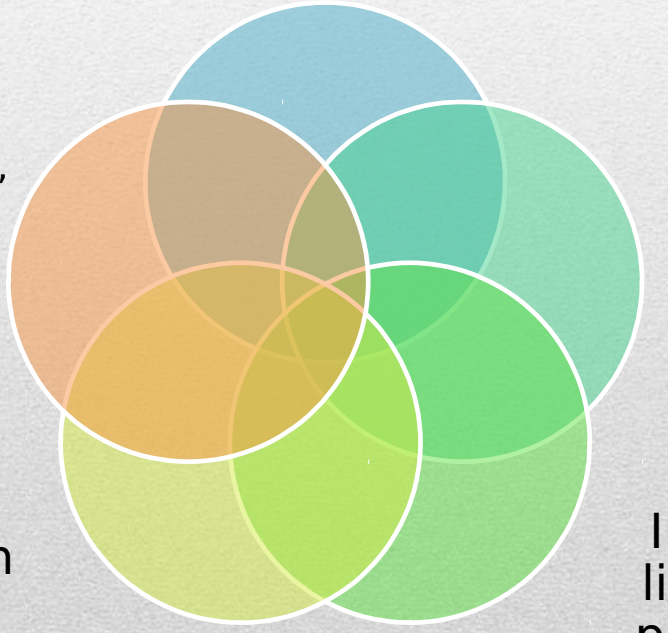
Staff:

The CRR-MM team includes a few dedicated staff and personnel belonging to the AlmaDL Office, working part-time for the Center: a librarian, the institutional repository manager, a metadata specialist and ontology engineer, an ICT technician, a research data specialist. The IP expert of our Division collaborates and provides legal advice.

Competencies and skills

Metadata and knowledge organization

Project management (organization, communication, monitoring, accountancy, problem solving)



Open Access (repositories, mandates, policies, compliancy with indexing services)

Data curation and preservation

IP, copyright licensing and personal data protection

How it all started

- March 2015:
  - meetings with the project managers of the UniBo European Research Office to understand H2020 requirements and the needs of our research teams
  - analysis of EU guidelines and documentations (OpenAIRE)
  - international benchmark: best practices and study cases (see bibliography); tools and guidelines
  - service offer
  - costing
  - meetings with research groups involved in Open Research Data Pilot

Service Offer

AlmaDL offers institutional repository services for free to UniBo researchers.

The Multimedia Center is self-sustaining and not all services are free of charge. Tariffs cover mainly direct and indirect costs.

Free services:

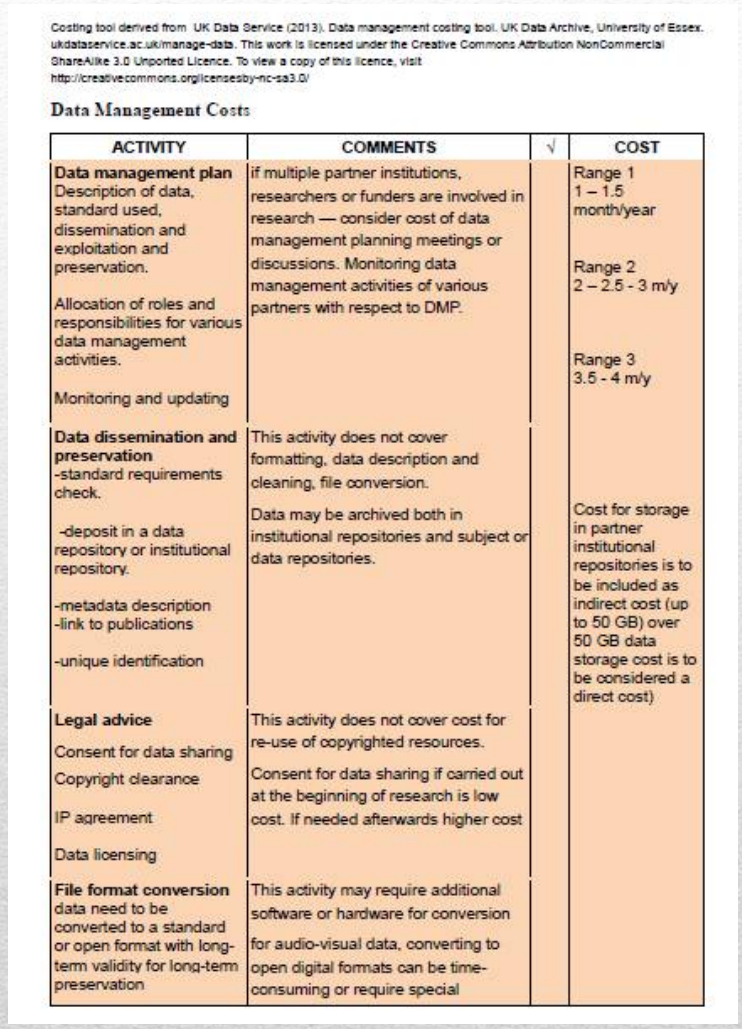
- assistance in pre-proposal and proposal submissions
- public presentations of the H2020 Open Research Data Pilot goals, Open Access mandate and OpenAIRE services
- guidelines and templates for research groups
- assistance for repository usage

Paid services:

- DMP deliverables for coordinated projects
- customized technical services: data format conversion, transcription, anonymization, data storage and transfer …

Costing…

- Delivery of the Data Management Plan is charged yearly according to the Consortium size.
- DMP costs are included in the project budget.



Costing tool derived from UK Data Service (2013). Data management costing tool. UK Data Archive, University of Essex. ukdataservice.ac.uk/manage-data. This work is licensed under the Creative Commons Attribution NonCommercial ShareAlike 3.0 Unported Licence. To view a copy of this licence, visit http://creativecommons.org/licensesby-nc-sa3.0/

Data Management Costs

ACTIVITY	COMMENTS	√	COST
Data management plan Description of data, standard used, dissemination and exploitation and preservation. Allocation of roles and responsibilities for various data management activities. Monitoring and updating	if multiple partner institutions, researchers or funders are involved in research — consider cost of data management planning meetings or discussions. Monitoring data management activities of various partners with respect to DMP.		Range 1 1 – 1.5 month/year Range 2 2 – 2.5 - 3 m/y Range 3 3.5 - 4 m/y
Data dissemination and preservation -standard requirements check. -deposit in a data repository or institutional repository. -metadata description -link to publications -unique identification	This activity does not cover formatting, data description and cleaning, file conversion. Data may be archived both in institutional repositories and subject or data repositories.		Cost for storage in partner institutional repositories is to be included as indirect cost (up to 50 GB) over 50 GB data storage cost is to be considered a direct cost)
Legal advice Consent for data sharing Copyright clearance IP agreement Data licensing	This activity does not cover cost for re-use of copyrighted resources. Consent for data sharing if carried out at the beginning of research is low cost. If needed afterwards higher cost		
File format conversion data need to be converted to a standard or open format with long-term validity for long-term preservation	This activity may require additional software or hardware for conversion for audio-visual data, converting to open digital formats can be time-consuming or require special		

Costing tool derived from UK Data Service (2013). Data management costing tool. UK Data Archive, University of Essex. ukdataservice.ac.uk/manage-data. This work is licensed under the Creative Commons Attribution NonCommercial ShareAlike 3.0

Our experience…so far

- Horizon2020 Open Research Data Pilot Unibo coordinated projects:
  - CATCH-EyoU
    - Area: Social Sciences & Humanities (YOUNG-5a-2014)
    - 9 partners
    - 36 months
    - DMP: Initial, Final and updates
    - status: initial DMP delivered (Februay 2016)
  - MADFORWATER
    - Area: Agro-food, Bio-economy, Energy and Environment (WATER-5c-2015)
    - 18 partners
    - 48 months
    - DMP: Initial, Mid-term, Final and updates
    - status: meetings in preparation of kick-off

Institutional Support Data Management Service

Pre-approval stage | **Approval of the project** | **Before starting the project**

PROPOSAL SUBMISSION AND EVALUATION

technical and legal advice to outline the project **data management policy** required in RIA or IA proposals (2.2. *Measures to maximise impact* under section 2 *Impact* of the template):

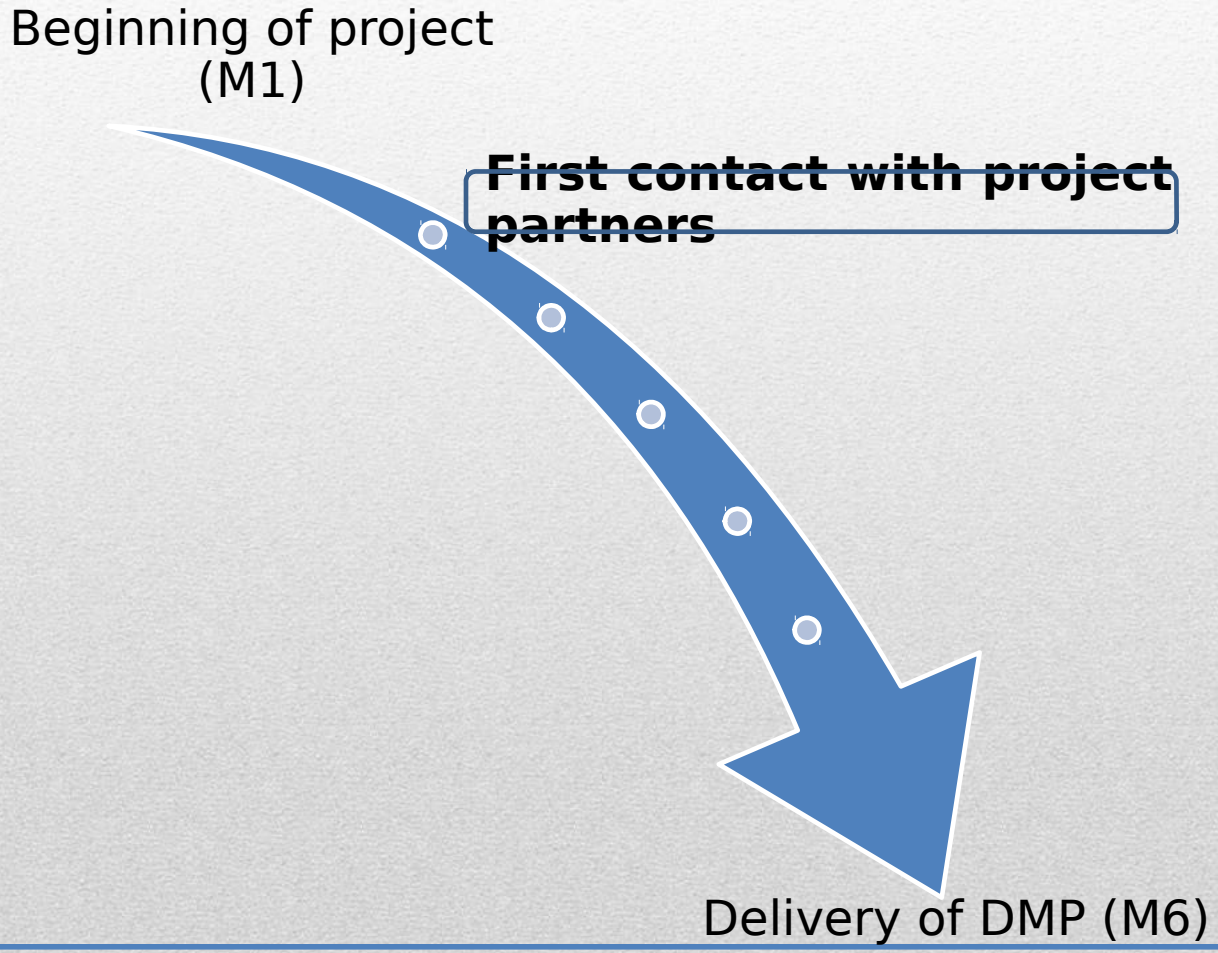
- meeting with research team, presentation of H2020 ORD Pilot, role of CRR-MM, quick analysis of the project research data
- revisions of the proposal also after the officers' review

- internal official designation of the Institutional Support Service (CRR-MM)
- study of the project included in the approved Grant Agreement (GA) and first sketch of the project structure and of the data sets to be described
- meetings with project coordinator

Pre kick-off meeting with project coordinator

- Explain goals and requirements of H2020 OA mandate and ORD pilot
- Discuss possible critical issues:
  - copyright and permissions for reusing third-party data sets
  - personal data treatment and confidentiality issues
  - data regarding possible patent applications
- Discuss list of expected data sets
- Ask for project data sharing facilities in the course of research
- Identify partners responsibilities about data management and archiving infrastructures for open access dissemination and long tem preservation
- Define roadmap: schedule deadlines timetable

Roadmap to Data Management Plan (DMP)



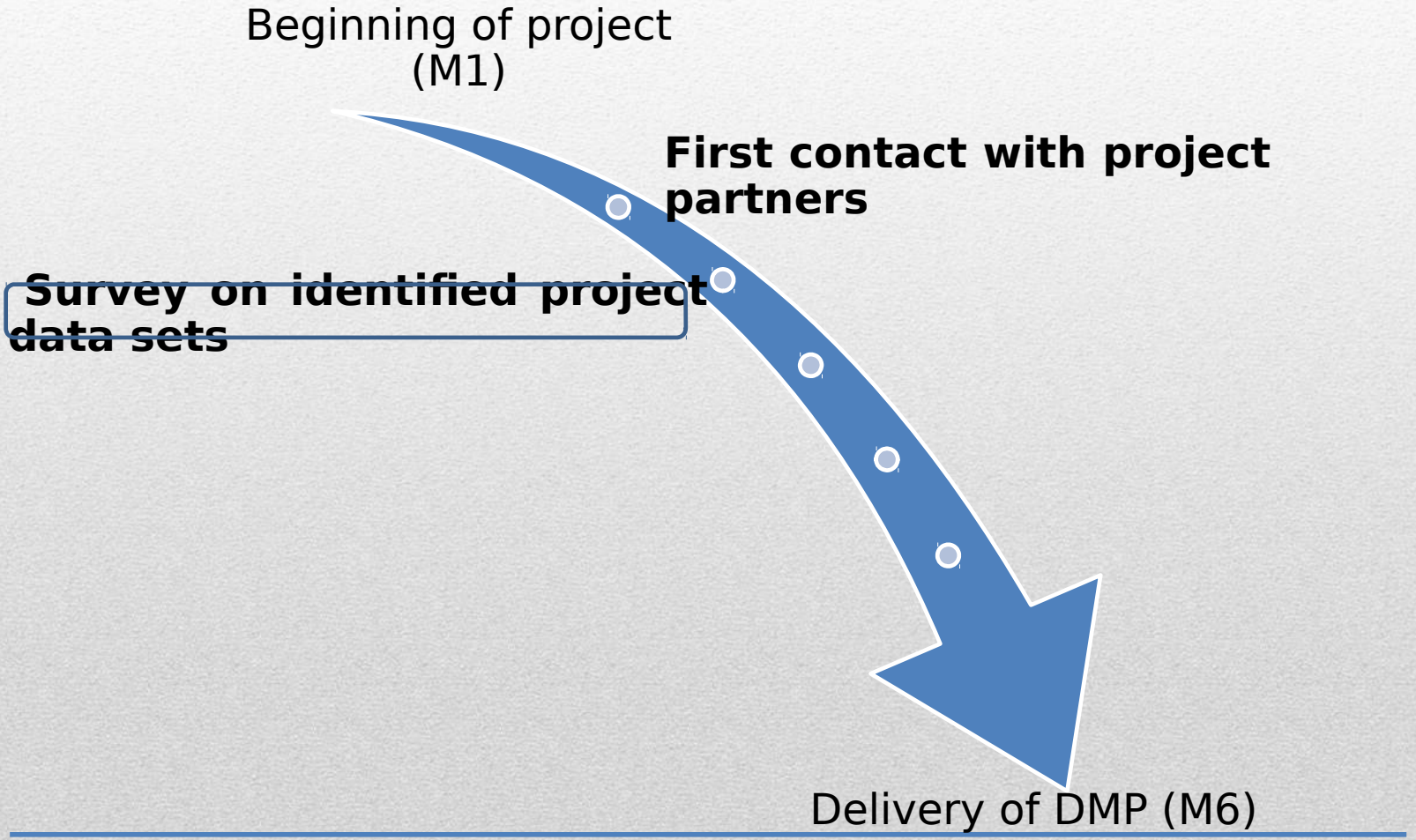
Kick-off meeting

- presentation of CRR-MM as the institutional DMP support service to project partners:
  - explain staff role as agreed with coordinator
  - remind the requirements of the open research data pilot
- presentation of roadmap steps to prepare DMP, work schedule and deadlines
- Focus on a check list of important issues that partners should detail or take into account in the DMP

Kick off meeting: main issues check list

- available repository infrastructures and archiving policies (Are they compliant with OpenAIRE? Are they trustworthy? Community or institutional based? ...)
- licenses and permissions required by commercial databases of scientific literature and textbooks for mining activities;
- terms and conditions regarding reuse of third-party databases protected by copyright or confidentiality agreements;
- databases compliancy with personal data protection legislation and consistency between terms and conditions included in informed consent forms used to collect personal data and the further treatment and dissemination of the same data
- datasets produced by the Consortium that can be commercially exploited;
- possibility to save data in open formats when using commercial software for data processing

Roadmap to Data Management Plan (DMP)



Survey on project data sets

To team leaders and/or WP coordinators:

- the data set list as derived from GA
- a questionnaire to be filled in with state of the art details about each expected data set:
  - introduction
    - aims and requirements of the pilot
    - how to fill in the questionnaire
  - four sections questionnaire covering the 5 questions required by H2020 ORD template
  - guidance and examples

1. Data set reference, name and description				Guidance
1.A. Data set name	Name _____			Provide a name or a title for the data set. A data set can be defined as a collection of similar research data, sharing a structure. In accordance with the project coordinator, we suggest WP leaders to provide the descriptions of the cross-national data sets pertaining their WP and resulting from the aggregation of individual data subsets collected or generated locally by the different research teams. We also suggest team leaders to provide descriptions for the individual data subsets collected or generated by their research teams. As a suggestion the data subsets may take the first part of their name from the name attributed to the data set resulting from their aggregation. In addition to this common name a specification must be added.
1.B. Data set ID	Id: _____ Id type:_____			If available, provide an identifier for your data set. In addition, indicate the identifier type, such as for example: DOI, Handle, PURL, URL, URN. Persistent identifiers are generally associated to a data set at the moment of archiving for dissemination or preservation in a repository. If the data set has no persistent identifier at the moment, just indicate "not yet available".
1.C. Data set origin, WP and current status	1.C.1. Indicate if the data set is: □ Collected □ Generated	1.C.2. Indicate the WP and the task: WP____.____	1.C.3 Indicate the data set current status (choose one option from the guidance): Status: _____	1.C.1. Collected data sets are data set that use existing data sources; generated data sets are based on data created by the project. 1.C.3. To indicate the data set current status use one of the following: 1) available 2) in progress 3) not yet available

Survey Questionnaire

section 1: Data set id, name and description

- data set name and ID
- data set origin, WP and current status
- data set creators and contributors (including contact persons and right holders, if relevant)
- data set description (abstract)
- related publications, if already available

Survey Questionnaire

section on Data Standards and Metadata

- methodologies for data collection or generation, data processing and quality assurance
- typologies, content and quality of data (numerical, textual, ..., raw, derived or secondary, ...)
- data format (used for raw data, data processing, sharing and preservation) and scale

---

- Metadata standards

Survey Questionnaire

section 3: Data Sharing and dissemination

- data accessibility (measures to share and disseminate the data, versions or portions of data sets not accessible under motivations)
- licenses
- repository for the dissemination
- data reuse (documentation, tools and software for enabling re-use of dataset)

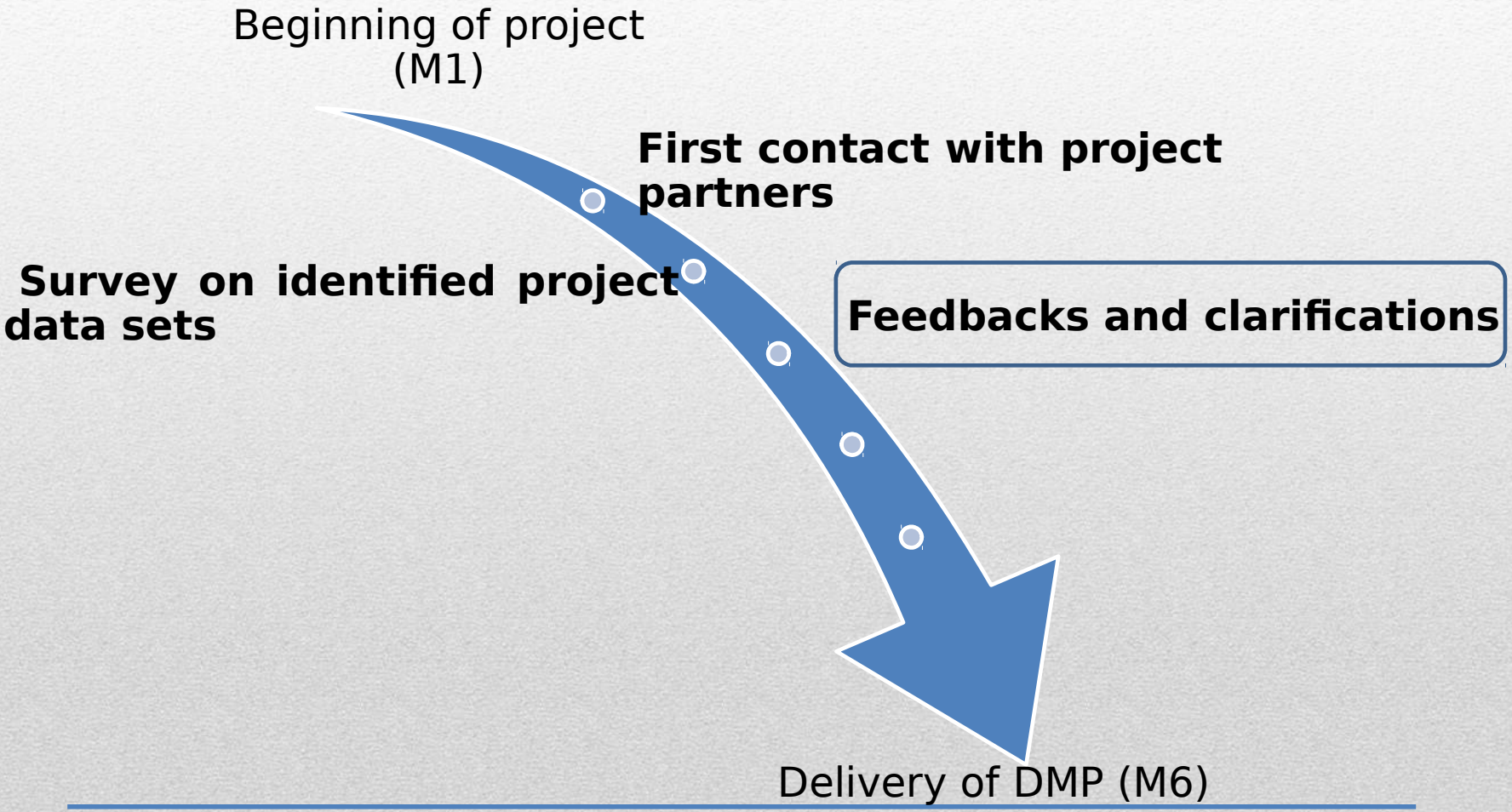
Survey Questionnaire

Section 4: Data set archiving and preservation

- Long-term preservation of the data set: *are there portions or versions of the data set that are not archived for preservation?*
- Data processing for long-term preservation (conversion to open formats, data anonymization, metadata preparation, documentation…)
- Repository for long-term preservation (responsible for data curation and management)

- Approximate data volume

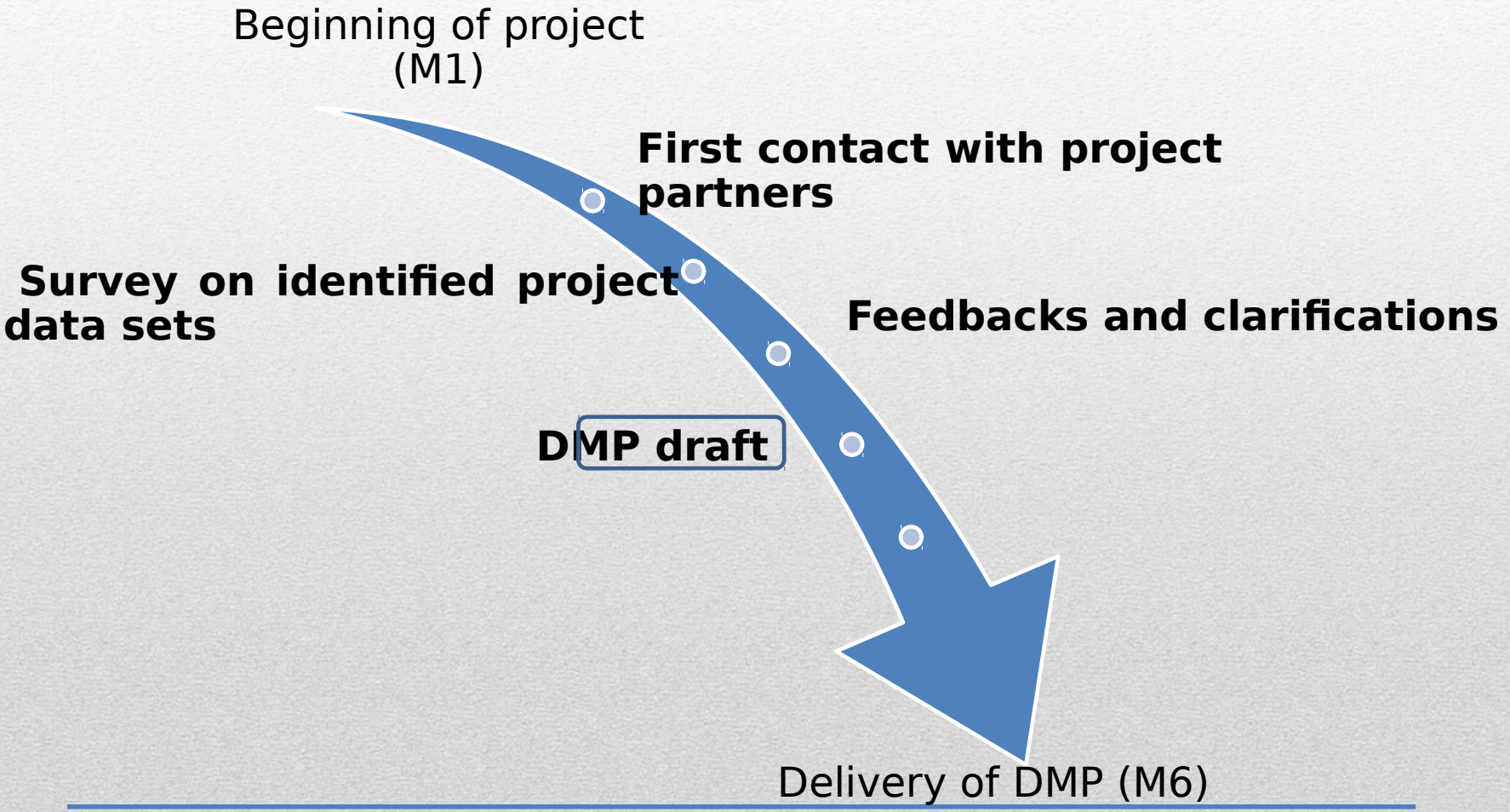
Roadmap to Data Management Plan (DMP)



Details collection and feedbacks from partners

- Monitor deadlines and send reminders to partners
- Offer customized help and further guidance for filling out the questionnaire
- Analyze the filled-out questionnaires
- Ask partners for missing information or clarifications
- Answer partner’s info inquiries about the questionnaire and pilot requirements
- Provide solutions for specific issues in accordance with project coordinator

Roadmap to Data Management Plan (DMP)



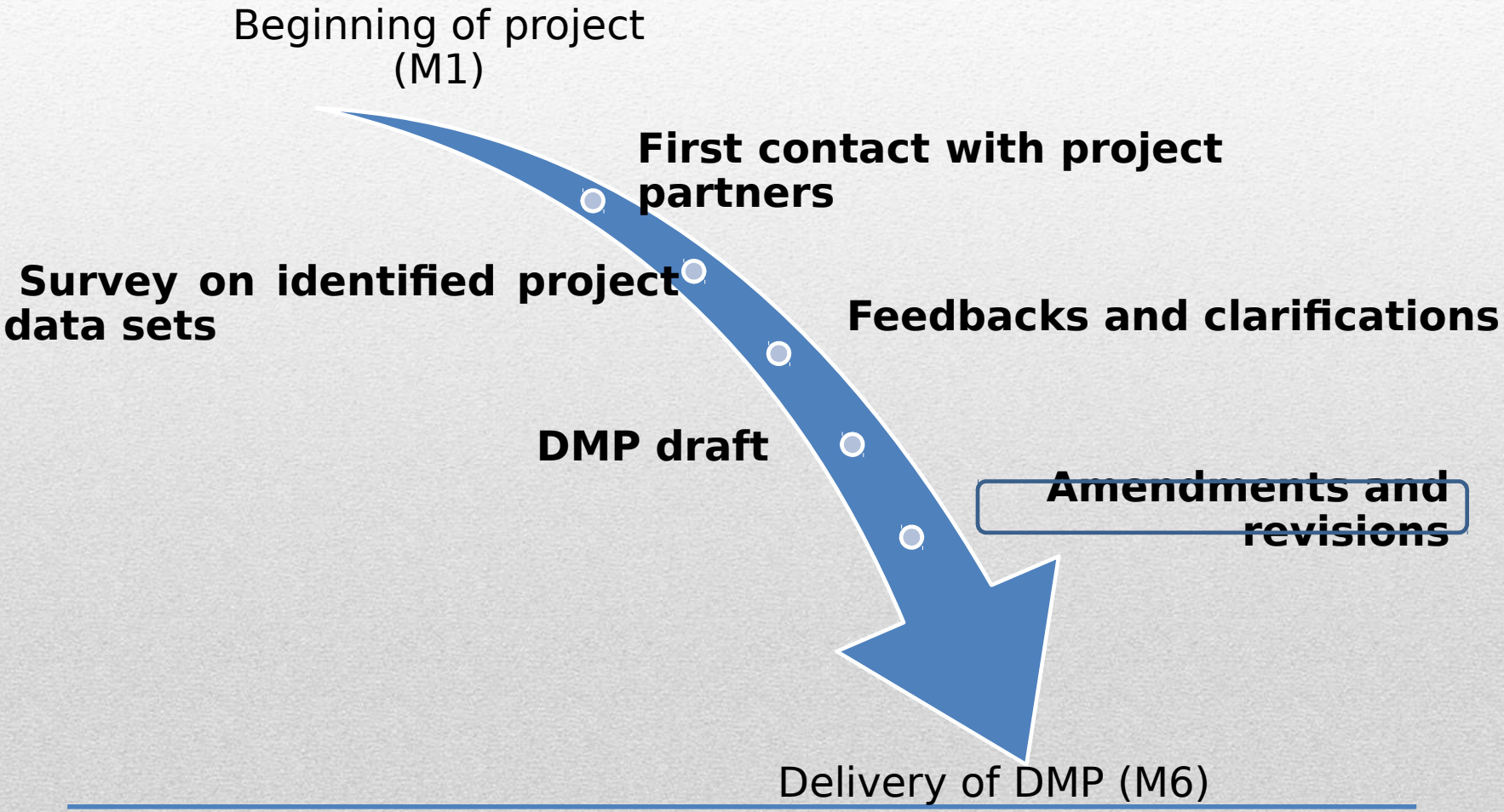
DMP draft

- Integrate all information collected through the questionnaire
- Define the DMP structure
  - document history, authors and contacts
  - executive summary
  - general data management policies regarding:
    - data sets identification and description (data sets and file naming…)
    - standards and metadata
    - sharing: list of repositories
    - storage and management
    - preservation and archiving
  - creators and contributors (roles and contacts)
  - description of the Data sets produced within the different WPs

DMP draft

- Introduction with a general description of WP aims, the data produced, their management, sharing and archiving procedures
- Tabular description of each data set
  - data set name, ID and description
  - creators, contributors, contact persons, rights holders
  - status
  - scale
  - repository

Roadmap to Data Management Plan (DMP)

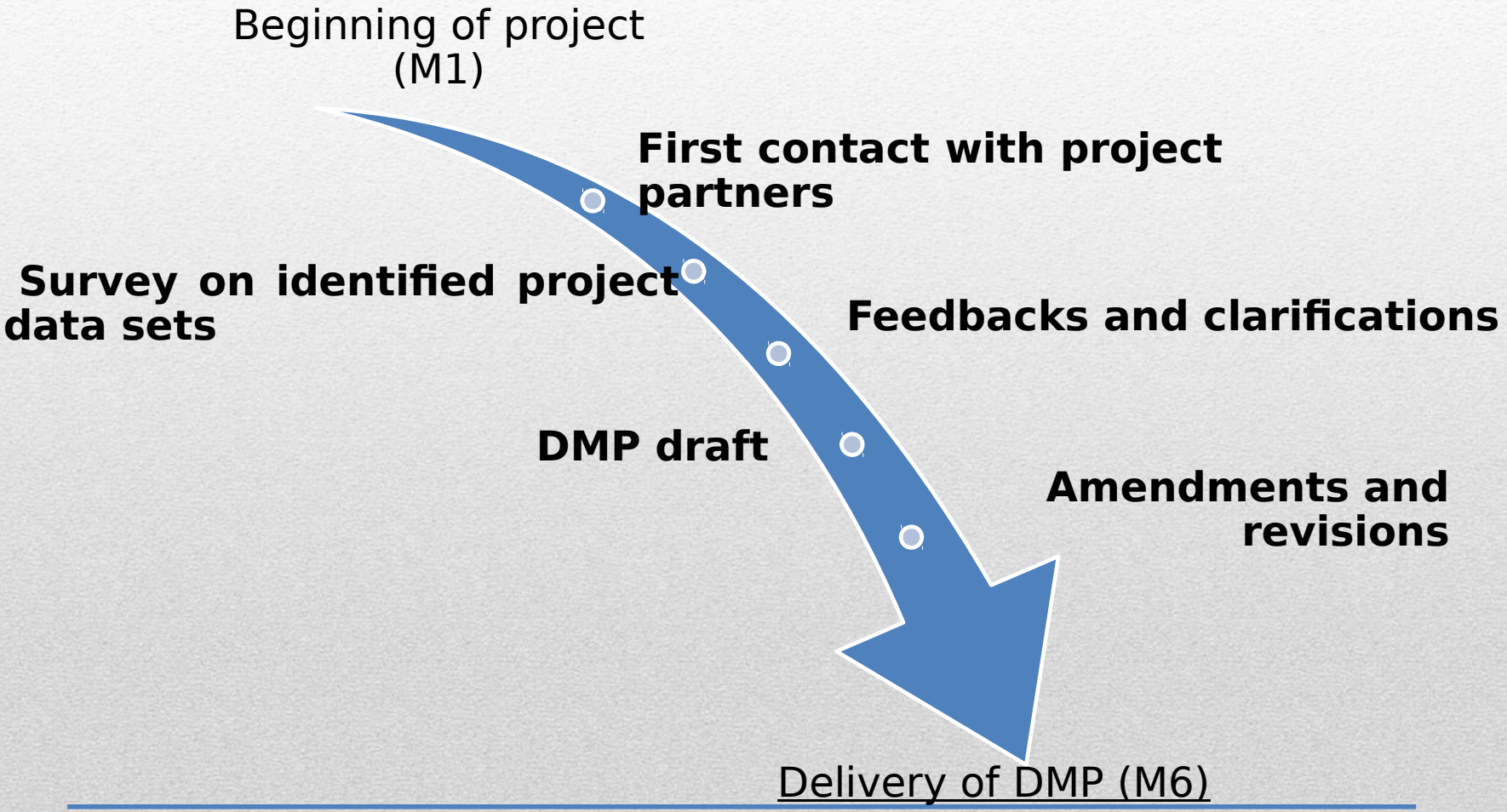


Amendments and revisions

- Two-rounds draft revision process:
  1. revision by project coordinator
  2. inclusion of coordinator’s amendments
  3. revision by all project partners
  4. inclusion of partners’ amendments
- Final approval by project coordinator

finally…

Roadmap to Data Management Plan (DMP)



Monitoring and reporting publication of data sets

Data sets publication is mainly linked to publication of project results. Establish an efficient monitoring procedure together with project coordinator and include it in the **Consortium publication policy.**

At the publication of results, check if the underlying data have been described and disseminated as in the DMP and check whether peer-reviewed articles are

[Option mandatory for the projects taking part in the 'Open research data pilot' and optional for the other projects

8. Open Research Data

More information on Data Management Plans (DMPs) in the Online Manual.

Digital Object Identifier, DOI (if available)	Title/Identifier (if no DOI available)	Is this dataset Openly accessible[5]?	Is this dataset re-usable[6]	If the dataset is linked to a publication, specify the DOI of the publication
[insert DOI reference]	[insert title or identifier]	*[YES] [NO]*	*[YES] [NO]*	[insert DOI reference of the publication]

- **Scientific publications and underlying data must be reported in PP.**
- **OpenAIRE provides metadata into PP automatically.**
- **The DMP must be updated when changes in the data sets occur.**

Our experience…so far

The reasons behind our choices

- We have been inspired by the DMPonline tool but we have decided to prepare **a different questionnaire** for collecting data set details, even though we own so much to the DCC tool!!! DMPonline is not really designed to meet H2020 ORD requirements and you have to start a new project for each data sets produced within each H2020 research project. The customized template can be more easily copied and modified by team leaders.

- The questionnarie includes all the details requested by the final DMP template. We have not distinguished between a first more general template - as required at M6 - and more detailed later versions. We thought that it was better to give the partners the whole picture since the beginning of the project in order to let critical issues emerge right from the start.

Critical issues and a few suggestions

From the point of view of the support service:

- role recognition within the institution
- lack of experience
- scalability and staffing
- organizational workflow (timing, monitoring, info sharing, collecting feedbacks from partners) very burdensome...
- what we expect from EU:
  - sharing of best practices and success cases (data organization and selection, personal data treatment, IP, licences, ...)
  - evaluation of work done (?)
  - stressing the importance of DM Planning at

Critical issues

From the point of view of researchers:

- in general, lack of knowledge and awareness about Open Access:
  - repository and archiving procedures for both publications (Gold & Green roads) and data
  - open licences and copyright issues
  - service providers (Open AIRE)
- Low perception of the advantages of Open Access
- DMP perceived both as burdensome and time-consuming
- Difficulties in identifying and select the data sets and the responsabilities for data management: they want ready-made answers but it is a matter of reflecting

After all a DMP support service mission is to provide guidance to make researchers understand

The
Value &
Impact
of their research data

Thank you for your attention!

Further information and bibliography

1. EC's Guidelines on Data Management in Horizon 2020: http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf

2. EC's Guide on Open Access to Scientific Publications and Research Data in Horizon 2020: http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-pilot-guide_en.pdf

3. OpenAIRE (https://www.openaire.eu/), about the Open Research Data Pilot: https://www.openaire.eu/opendatapilot
   - OpenAIRE Research Data Management Briefing paper: https://www.openaire.eu/briefpaper-rdm-infonoads
   - Open Research Data Pilot factsheet: https://www.openaire.eu/or-data-pilot-factsheet

4. Digital Curation Center (DCC) http://www.dcc.ac.uk/
   - How-to Guides & Checklists: http://www.dcc.ac.uk/resources/how-guides
     - «Checklist for a data management plan»: http://www.dcc.ac.uk/resources/data-management-plans/checklist
     - «Five steps to decide what data to keep»: http://www.dcc.ac.uk/resources/how-guides/five-steps-decide-what-data-keep
     - «Where to keep research data»: http://www.dcc.ac.uk/resources/how-guides-checklists/where-keep-research-data
     - «Appraise and select research data»: http://www.dcc.ac.uk/resources/how-guides/appraise-select-data
     - «Develop a data management and sharing plan»: http://www.dcc.ac.uk/resources/how-guides/develop-data-plan
     - «License research data»: http://www.dcc.ac.uk/resources/how-guides/license-research-data
   - Zenodo community «Research data management (RDM) training materials» : https://

Further information and bibliography

5. European Universtity Institute (EUI) Research Data Guide: http://www.eui.eu/Documents/Research/Library/ResearchGuides/Economics/PDFs/EUIResearchDataServicesGuide3.0July2015.pdf

6. Deakin University Library "DATA MANAGEMENT TOOLKIT FOR RESEARCHERS“: http://www.deakin.edu.au/__data/assets/pdf_file/0018/267111/data-management-toolkit.pdf

7. Examples of H2020 Initial DMPs:
   - http://edsa-project.eu/edsa-data/uploads/2015/02/EDSA-2015-P-D55-FINAL.pdf
   - www.umobile-project.eu/index.php/download/category/17-wp6-dissemination-exploitation-and-standardisation?download=11:d-6-10-data-management-plan
   - www.benefit4transport.eu/index.php/component/phocadownload/category/12-reports?download=42:benefit_d6_1_dmp_df
   - http://www.waterinneu.org/deliverables/641821_Draft%20Data%20Management%20Plan.pdf
   - http://www.pathway2health.eu/wp-content/uploads/2016/02/PATHway-Deliverable-D1.3-Report-on-Data-Management.pdf
   - https://www.point-h2020.eu/wp-content/uploads/2015/07/POINT_D6.5-Data_Management_Plan.pdf
   - http://www.freme-project.eu/resources/FREME_Deliverable_D7-4.pdf
   - http://heron-project.eu/images/Deliverables/HERON_D.2.6_DataManagementPlan_final.pdf
   - http://www.woscap.eu/documents/131298403/131299900/Data+Management+Plan_D1.3.pdf/05262d10-8721-434e-8aee-1a88ecc0156d
   - https://verc.enes.org/esiwace/results/deliverables/Data_Management_Plan_Annex_final.pdf

Further information and bibliography

7. Open licences info:
   - http://opendefinition.org/licenses/
   - https://creativecommons.org/licenses/

8. Metadata:
   - http://www.ddialliance.org/
   - https://www.datacite.org/

10. Useful and interesting tools:
    - DCC's «DMPonline»: https://dmponline.dcc.ac.uk/
    - B2 Service Suite: https://www.eudat.eu/

11. Repositories:
    - Registries of repositories:
      - https://www.openaire.eu/search/data-providers
      - http://www.re3data.org/
      - http://www.opendoar.org/
      - http://roar.eprints.org/
    - Zenodo: https://zenodo.org/
    - Data seal of approval http://datasealofapproval.org/en/