ELITMa
Module 2 - RDM strategy

Best practices from the ELIXIR portfolio of data management tools and services. A quick guide to adopting ELIXIR resources and optimising an ELIXIR Node's data management strategy

Facilitators: Mijke Jetten & Celia van Gelder
Finland, 27-28-29 June, 2023

Day 1: Code of Conduct

In case you experience any issues, note them down or address them to Mijke or Celia

Day 1/2: Parking Lot

Let us collect topics that we should not forget and/or that you miss in the current programme. We will try to come back to them as much as possible during the module and will also make sure that these topics are not forgotten going forward.

- Celia
  - Terms4fairskills
  - Collect other/background information about other ELIXIR (Node) Services along the same line (to complete the slide collection :-)) - could be a good idea maybe. Plus we should also think about bringing the talks of yesterday to a larger audience
  - Slides from EOSC-Life/Health-RI workshop 13/14 June
  - Depletable vs non-depletable services
- Collected from the preparation phase:
  - Are there important topics missing?
    - DM Community building (EE, UK, LU, CH, EE, CZ)
    - Create engagement
    - Share experiences

- Role of emerging RDM community (o.a. financial model, relation to Platforms)
- Policies related to the data management strategies and best practices in different nodes (FI)
- Formal education & to hear from other nodes if there is stories to share to establish formal programs on RDM (NO)
- Monitoring routines for RDM (NO)
- Best practice and emerging patterns in core tool deployments (LU)
- DSW deployments and KM customisations in nodes (LU)
- RDMkit instances, plans, if any, for further instances (LU)
- Guidance documents for our application purposes (DE)
- Show-and-tell sessions to exchange experiences (DE)
- DMP and FAIRification as tightly bound issues (CZ, IT)
- Formal DS curriculum (ECTS awarded with a certificate or degree) (NO, CZ)
- Role and position Node data manager
- Training for Node data manager

Day 1: Introduction round

Add your Node, name, role, email address, F2F/online + a fun fact about yourself
- Portugal
  - Gil Poiares Oliveira - Bioinformatics training technician - data management, gil.oliveira@tecnico.ulisboa.pt, F2F, I've been trained on meditation
  - Ines Chaves, plant community researcher, data steward, ichaves@itqb.unl.pt, online, cat lover
  - Mariana de Freitas - Impact Officer - Fun fact: Climbed Mount Fuji :)
- France
  - Angela Sáenz - Node Coordinator angela.saenz@france-bioinformatique.fr F2F + fun fact: I love Bollywood dance
  - Paulette Lieby - RDM Community - paulette.lieby@france-bioinformatique.fr ; mostly involved with constructions of DMPs and working across communities ; online ; a fun fact: I adore(d) bushwalking in Central Australia

- ○ Jean-François Dufayard - RDM Community - jean-francois.dufayard@cirad.fr - mainly working in animating communities (national infrastructures) to build and gather FAIRisation standards for Recherche Data Gouv (french national research data initiative) ; wingman of Paulette for many tasks - fun fact: there is absolutely nothing fun about me, I've no humour at all, don't understand jokes ; I like cats too with a lot of black pepper
- Finland
  - ○ Riku Riski, Node coordinator - riku.riski@csc.fi - the desperate wooden boat custodian
  - ○ Minna Ahokas, Data coordinator - minna.ahokas@csc.fi F2F, I'm a historian
  - ○ Susanna Repo, deputy HoN - susanna.repo@csc.fi, F2F, Fun fact: I sing in a symphony choir as a soprano
- Norway
  - ○ Korbinian Bösl - Data Management Coordinator korbinian.bosl@uib.no (online) + fun fact: used to be a paramedic
  - ○ Espen Åberg - Services Coordinator espen.aberg@uit.no (online) Fun fact: I do gymnastics.
  - ○ Nazeefa Fatima - Data Steward, nazeefaf@uio.no (Online) - Fun Fact: I can survive very well without coffee and sugar.
  - ○ Matus Kalas - Ontologist (online)
  - ○ Ingeborg Winge- incoming Node coordinator Ingeborg.Winge@uib.no (online, on-off)
- Germany
  - ○ Helena Schnitzer (online), Training Coordinator, hschnitz@cebitec.uni-bielefeld.de, Fun Fact: you probably see my cats in the camera more often than my face
  - ○ Person 2 etc
- Sweden
  - ○ Jessica Lindberg (F2F), Node Coordinator, jessica.lindberg@nbis.se - Fun fact: Trained chef
  - ○ Niclas Jareborg (online), Data manager, niclas.jareborg@nbis.se - Fun fact: do improv theatre and play role playing games
- Netherlands
  - ○ Mijke Jetten (F2F), FAIR lead/Community Manager Data Stewardship, mijke.jetten@health-ri.nl. Fun fact: I won't have a meal without lots and lots and lots of black pepper on it :-)
  - ○ Lifang Liu,(online) Elixir-NL - NL node coordinator. Lifanfg.liu@health-right.nl. Fun part: following Training online today:)
  - ○ Celia van Gelder (F2F), celia.vangelder@health-ri.nl Training Programme Manager Health-RI, roles in ELIXIR: Training Coordinator ELIXIR-NL, (Node coordinator until July 1, 2023).   Fun fact: I played a drunken lady in a cafe in the performances of my theatre club a few weekends ago :-)

- Janet Vos (online),janet.vos@health-ri.nl, , ELIXIR-NL node coordinator,
- Italy
  - Ale Cestaro (F2F) - Node Coordinator, CNR, a.cestaro@ibiom.cnr.it, FunFact: I smoke pipe everywhere …
  - Ivan Mičetić (F2F), Data Management Coordinator, ivan.micetic@unipd.it. Fun fact, been busy searching for auroras but couldn't find the night either…
  - Flavio Licciulli (Online), RDM Community, Interoperability Platf member; flavio.licciulli@ba.itb.cnr.it. Fun fact: Until a couple years ago, I was a race car driver
- Ireland
  - Siobhán Cleary (online), Research Data Management Coordinator, siobhan.cleary@ucd.ie, Fun fact: I have an identical twin sister!
  - Maria Doyle (F2F), Training Coordinator, maria.doyle@ul.ie, Fun fact: Have just taken up Reformer Pilates
- Luxembourg
  - Pinar Alper (online), Data Management Coordinator, pinar.alper@elixir-luxembourg.org Fun fact: I currently am working in a container

- Czech Republic
  - Jiří Vondrášek (F2F), Head of Node, jiri.vondrasek@uochb.cas.cz Fun fact: during winter season my head looks like a doughnut
  - Anna Strachotová (F2F), Node Coordinator, anna.strachotova@uochb.cas.cz
- Hub
  - Fabio Liberante (online) - Programme Manager for Data & Interoperability Platforms, fabio.liberante@elixir-europe.org - my first job was as a pizzaiolo 🍕 - pizza chef
  - Clare Garrard (online), Scientific Officer - Interoperability & Communities, clare.garrard@elixir-europe.org, Fun fact: you will probably hear my cat snoring in my Zoom at some point…
  - Ivana Versic (F2F) - Head of Ops, ivana.versic@elixir-europe.org, Fun fact: Ivan Micetic and I are from the same town in Croatia

- Estonia
  - Kairi Koort (online), Deputy Head of Node, kairi.koort@tlu.ee, Fun fact: I spend my weekends working from an island in the Baltic sea called Hiiumaa/Dagö. I love languages and art.
  - Diana Pilvar (online), Data Steward, diana.pilvar@ut.ee , I'm interested in books that hurt my soul, desserts that give me diabetes and i am a night owl

ELIXIR Training Programme in Management

Day 1: Questions & Answers for speakers/presentations

RDMkit

Pinar
- Ivan: what do you intend by assembly development (tools). Could you name some of these?
    - We have assembled the following tools for the End2End Assembly:
        - **Keycloak**, Authentication/Authorisation (Open Source)
        - **REMS**, Tool to manage access requests and orchestrate the DAC decision process (ELIXIR FInland)
        - **Data Catalog**: A lightweight catalog based on the DATS model. Allows discovering datasets and kickstarting access requests (tool from the FAIRPlus project)
        - **DAISY**: A project and data registry targeted for data stewardship teams. We use it to record GDPR-required documentation on datasets. It is also our access control master, Keycloak is just an enforcer. (From ELIXIR LU)
- Paulette: I am quite impressed by your E2E assembly! Superb!
    - Thanks Paulette! We're in the process of writing it up as a paper.

Korbinian
- Pinar**:** How do you use RDMkit in training? Can you elaborate a bit.
    - We use RDMkit as a reference in workshops and courses
    - Also we have been using RDMkit in flipped classroom approaches so that the learners had to read up on a topic in advance that was discussed later in the workshop
- Paulette: excellent talk too! You gave me ideas about expanding and consolidating the work we do using DSW as our tool
- Jean-François: Agree with Paulette, we develop machine actionable structure DMP templates, on DSW, for many many lifesciences subdomains. We hope to be able to find synergies with other needs and initiatives in Europe

DSW

Ana
- Celia: Are there links between TeSS and DSW? I know this was the plan but not sure if this is already implemented

- Celia: If you have all these separate instances, e.g. for institutes, what does this mean for the KM? WIll these also diverge? Do you keep track of additions to the KM?
- Celia: Is there a place in ELIXIR where Nodes implementing the DSW come together/discuss (and not only the technical challenges). I know there is a separate DSW Slack Workspace too.

Korbinian
- Paulette:
    - you use the cloud solution, and you still have full control? As opposed to a self-hosted instance?
        - We have full admin control in the instance (not on the deployment, DB, …)
    - Also, how to you use project templates to provide guided/specific recommendations
        - We have templates from other research infrastructures for specific technologies/domains - this works OK but the users often would wish to combine these
        - P: I see
    - For feeding data into SEEK etc… I guess you have developed the appropriate output templates?
        - We are using maDMP to stay interoperable (instead of DSW json), maDMP is not as much information as we would want
        - You mean RDA DMP DCS - yes – ok, got it; but you could write/expand the output templates?
            - https://registry.ds-wizard.org/document-templates/dsw:rda-madmp:1.14.0
    - How do you link to 'examplar DMPs'?
        - dsw-instance.org/projects/create/from-template?selected=templateUUID
            - Ok, thanks
    - Issues with text export: could you say more? I didn't quite catch this
        - Institutional support is frustrated with the ScienceEurope+Horizon Europe/2020 export as they are hard to backtrack on wrong choices in the questionnaire.
    - I have the feeling have an extended chat with you would be quite enlightening
- Mijke: how "recent" is the comparison of similar tools (slide 57)
    - https://ds-wizard.org/comparison says April 2023 (but it's a bit coloured)

- we have one made by the Norwegian universities here:
  https://doi.org/10.5281/zenodo.7428542 (Norwegian)
- Niclas, can you share a similar comparison you made in Sweden/referred to at the Prague session from Yvonne and a researcher I think
- Celia: Is there a place in ELIXIR where Nodes implementing the DSW come together/discuss (and not only the technical challenges). I know there is a separate DSW Slack Workspace too. now discord: https://discord.gg/R9kmh9Vd

Federated EGA

Riku

- Ivan: I would really appreciate some of the common building blocks, tools, hints… In the name of "economics of scale "
  - Re-use of GA4GH standards - here is re-use with GDI starter kit https://github.com/GenomicDataInfrastructure namely REMS/LS AAI is also in FEGA
  - Link to FEGA session materials at the last AHM, perhaps worth to take a look at https://drive.google.com/drive/folders/1J3_VficsN_425zHKnrLKHG1UpOqF_xk-
  - Service description https://research.csc.fi/-/fega

Korbinian

- Paulette:
  - interesting that NO has developed its own TRE! Great!
    - It's not a ELIXIR-NO achievement, but we collaborate extensively
    - In fact we actually have 3 TREs in Norway, which we work with:
      - Services for sensitive data (TSD) - University of Oslo
      - HUNT Cloud - NTNU
      - SAFE | Employee Pages | UiB

  - Vulnerable to staff turnover: indeed! We all know about this ;)

Molgenis

Joeri

Korbinian:
  ○ Is there a use case/example connecting Molgenis with ELN(s)?
Niclas:

- Pointers to real-world examples of uses of Molgenis, to better understand how it can be used and what problems it can solve

EDAM

Herve
- RDM terms extension of EDAM
  https://edamontology.github.io/edam-browser/#topic_3071&version=latest
- Latest extension on RDM topics in EDAM (not yet in the browser)
  https://webprotege.stanford.edu/#projects/69591619-4eda-4f03-9e7f-65b213038fe1/edit/Classes?selection=Class(%3Chttp://edamontology.org/topic_3071%3E) - registration required

FAIRsharing

Allyson
Clare: Do you have much data on whether/how much people are using FAIRsharing to improve citability?
Allyson: we have standard Google analytics, but have not yet gathered individualised record-level citation statistics. I think that would be something interesting to do!

FAIR Cookbook

Danielle
Question was raised what is ELIXIR-LU's commitment towards FAIR Cookbook now that it was included into the Node's service delivery plan. This led to a discussion of what is expected by ELIXIR of services included into the service delivery plans. It would be great to have guidelines available but also more clear understanding of the expectations.

→ some Nodes expressed reluctance to list services on SDP because of unclear expectations

ELIXIR Training Programme in Management

Day 1: Notes from wrap up session

3 minutes per Node, for example on:

- Identified challenge in/for your Node in relation to roles and competences
- Take-home messages for the day, eye openers, plans for the next steps for your Node DM strategy
- Parking Lot topics

CZ

- Challenge - to convince researchers
- 2021 started talking with funders
- A funder is member of the ELIXIR CZ Board
- Proper storage - national repository will be in place
- Ontologies - Parking Lot
- Want to establish Data steward CV - educational program

EE

- There is no university policy
- THere is policy on subtopics like GDPR
- Create DMP template for my university
- Compliance is not covered - no staff at the moment
- Parking Lot: DSW translations

IT

- Request for RDM only from the Commission
- No FAIR data requirements
- No push from above to do proper RDM
- People are doing this in smaller groups, on universities; they are not part of the ELIXIR Node
- Bottom-up approach
- NC knows about it
- Connect these local groups and bring to national level, work in progress, started to build local RDM community
- It would be push if the Commission also explicitly requires FAIR Data

FR

- Could fill nearly all the slots

- Doing almost everything
- We need to coordinate the work much better
- Push of French ministries helps
- Lot of projects coming in the next year, exercise was useful for that!

IE

- Challenging exercise, new people, but useful to get together
- Take the table to the Node

SE

- Challenge: the Node does not have effective a policy around RDM
- Government policy & university policy is there
- GDPR pretty well covered in the DM team
- Training for the non-DM staff in the Node - gap
- Node has legal responsibilities in data processing so users need to know
- Knowledge about RDM in the own staff needs to be upskilled
- Next step: Will look at some of the services presented today to see if we want to take some of them on in the Node.

NO

- Main stakeholders use different terminology, especially on institutional level
- Confusions
- Tick most of the 8 boxes.
- Gaps on influencing policy on national & institutional level. oOing ad-hoc now but will be a WP in the next phase.
- Gap in formal legal expertise
- Post-graduate degree for data stewards
- Gap: internal knowledge management; internal information, platform

UK

- Federated Node
- Status of RDM/data stewardship differs
- Funders are engaged and are making guidance documents
- Next steps: consultation back to the HoN and management team. Look to align with objects of the ELIXIR DM community
- BioFAIR

FI

- ○ Compliance: individuals working on this but lack of coordination
- ○ Infrastructure: we'll need to map the national RDM landscape and build a strategy for Node's RDM services
- ○ August workshop coming already - about RDM Services
- ○ ELIXIR-FI does not have a mandate to arrange RDM Services for life sciences in Finland (yet) and the question is also if we want/need this mandate

PT

- ○ Key limitation: Reference documentation/ knowledge base e.g. related to Services where & how to do it.

Day 2: Notes from plenary discussion on SE example

PT asked for advice on how to have a minimum viable solution for data stewardship?
Head of Node in SE recognises the importance of prioritising funding; Reflect on what our offer can be considering our competences
Qs: - What are the competences of the staff you are able to fund?
- - What is a good payoff at your local level?
- - Expectation management

IT Node is in a similar position as the PT Node: how to find the first financing module/jumpstart

FI Node has been following the SE Node with interest.
What is the drive of the growth?
How can we justify more funding? How do we map the needs of the research community?
Maybe a frame would be helpful

SE: Management also agrees it is really important and we have to do it.

On the support side, SE has a helpdesk function - Researchers have data management issues and want support for that (consultations)
A lot of it is depositing data in data repositories.
Data management planning (consultations) and providing a nationalised data management module. SE does not have tech solutions in their portfolio (exception: FEGA)

FI Node has 1 FTE (funder's perspective)

Perhaps not every node needs to have the same types of functions. - SE Node

Day 2: Notes from template discussion

Maturity Model may not be the best name -> Profile
It is not competitive
For self-assessment
Defining where you are and where you want to go

Use this as a tool to define your mission/vision , for yourself and others
Plus the other way around: Use your mission / vision to fill the template

Day 2: Notes from discuss next steps

Italy: Ivan: Mandatory first step for Italy is to make a connection with the government/national funders and only after to draft a strategy. Without this commitment from above, nothing will progress. We are currently working in this direction.
France: Paulette: support Robert's and Niclas in saying that 1) looking out for other nodes once ready to initiate some action (writing, …) – Robert can correct me here, and 2) emphasing the bottom-up aspect
UK: Robert -> plan follow-up meeting(s)

Day 2: Notes from RDM community and wrap up

STEERS project - how do we see the relation?
More information needs to be distributed first because the majority of the participants are not aware. Can be taken further in follow-up meetings

ELIXIR Training Programme in Management

Portfolio template

Please copy and paste yourself to your Node portfolio

Section 1 - Node description & Roles
- Short description of the Node and the roles in your Node management
  For inspiration, look at the [slide deck of the 2021 Node triangle workshop](#).

Section 2 - Current services

What services do you already offer?
- What are the current services you have in your Services Portfolio / Services Delivery Plan?

Section 3 - Nodes' good practices

Based on the presented examples, reflect:
- What services are relevant to your Node?
- What do you need (from others) to implement it?

Section 4 - Funder requirements for FAIR data

Based on the ZonMw presentation and examples:
- Reflect on how funder requirements to FAIR data may affect your Node DM strategy and identify potential implications

Considerations:

1. Funders evolve from RDM requirements for single projects towards organising research communities, and domain specific FAIR implementations
   - What examples do you know about?
   - What is your opinion on this role?
   - What do you expect from a funder?

2. Collaboration for FAIRification
   - What opportunities for collaboration do you know? Were they fruitful?
   - Who are / were collaborating?
   - What opportunities do you see for FAIRification within your Node?

Section 5 - Data stewardship roles & competences

ELIXIR Training Programme in Management

Analyse the competences of your current (and future) Node team, based on the identified competence areas for data stewardship. 8 competence areas (https://competency.ebi.ac.uk/framework/datasteward/1.0)

- **Policy and Strategy**: Development, implementation and monitoring of the research data management policy and strategy of the institute
- **Compliance**: Compliance with relevant codes of conduct, legislation and field specific standards
- **FAIR data**: Alignment to the FAIR data principles and the principles of Open Science
- **Services**: Availability of adequate support on research data management, in staff or services
- **Infrastructure**: Availability of adequate infrastructure for research data management
- **Knowledge management**: Adequate level of knowledge and skills on research data management in the institute
- **Network**: Obtaining and maintaining a network of aligned expertise areas and relevant organisations by the institute
- **Data archiving:** Adequate support and infrastructure for FAIR and long-term archiving of the data of the institute

Competence area	- How is this competence area tackled in your Node? - How does your Node want to tackle this going forward?	By whom? Name and/or role (NC, DM, other,...)
Policy and Strategy		
Compliance		
FAIR Data		
Services		
Infrastructure		
Knowledge Management		
Network		
Data archiving		

Section 6 - Learn from & share with

Based on today's presentation, examples and discussions:

- What are your take-home messages for the day?
- Lessons learned, eye openers, plans for the next steps for your Node DM strategy?

Section 7 - Nodes services maturity model

Model aspects	- How do you cater for these aspects / are you offering services / if not, do you plan to?	If out of scope explain why
Strategy		
Policies		
Training		
Guidelines		
Metadata standards		
Data storage		
Data sharing		
DMP		