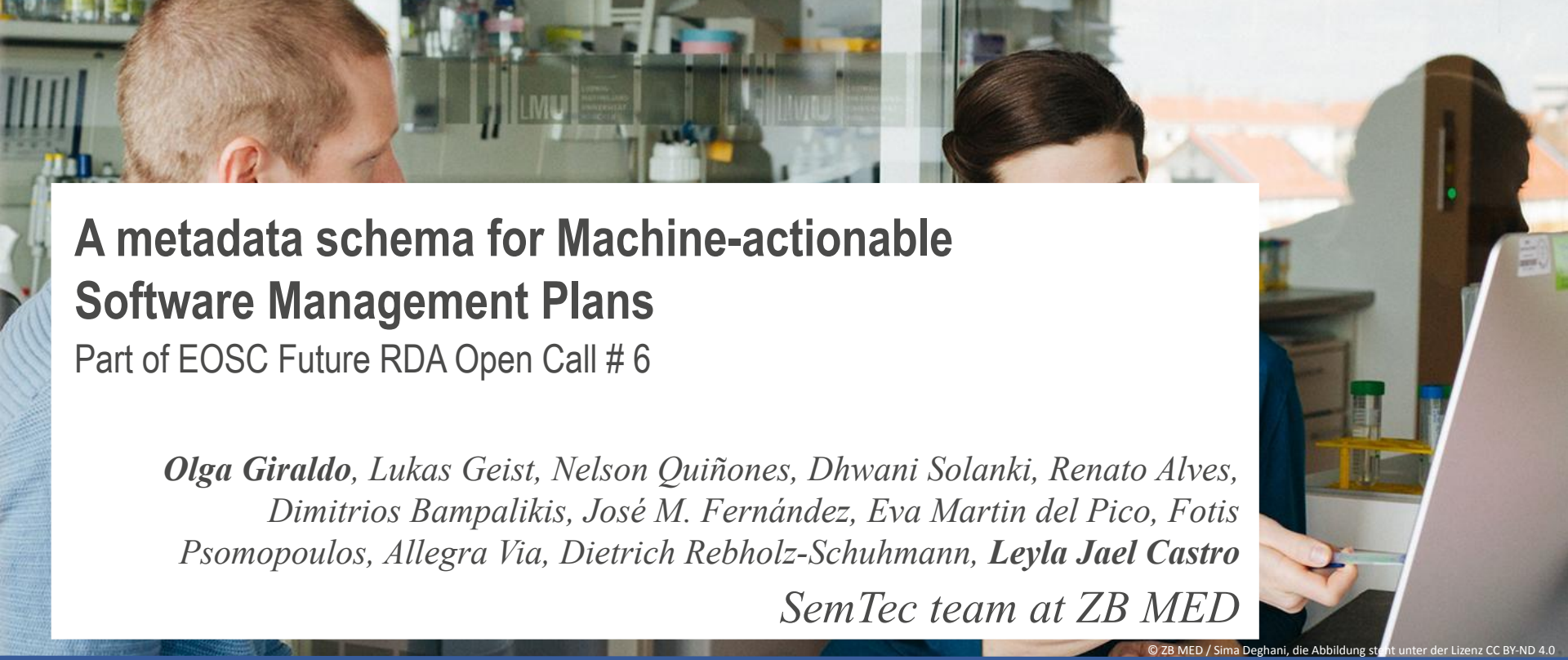


# A metadata schema for Machine-actionable Software Management Plans

Part of EOSC Future RDA Open Call # 6

***Olga Giraldo***, *Lukas Geist, Nelson Quiñones, Dhwani Solanki, Renato Alves, Dimitrios Bampalikis, José M. Fernández, Eva Martin del Pico, Fotis Psomopoulos, Allegra Via, Dietrich Rebholz-Schuhmann,* ***Leyla Jael Castro***

*SemTec team at ZB MED*



DaMaLOS 2023

RDA RESEARCH DATA ALLIANCE

EOSC Future

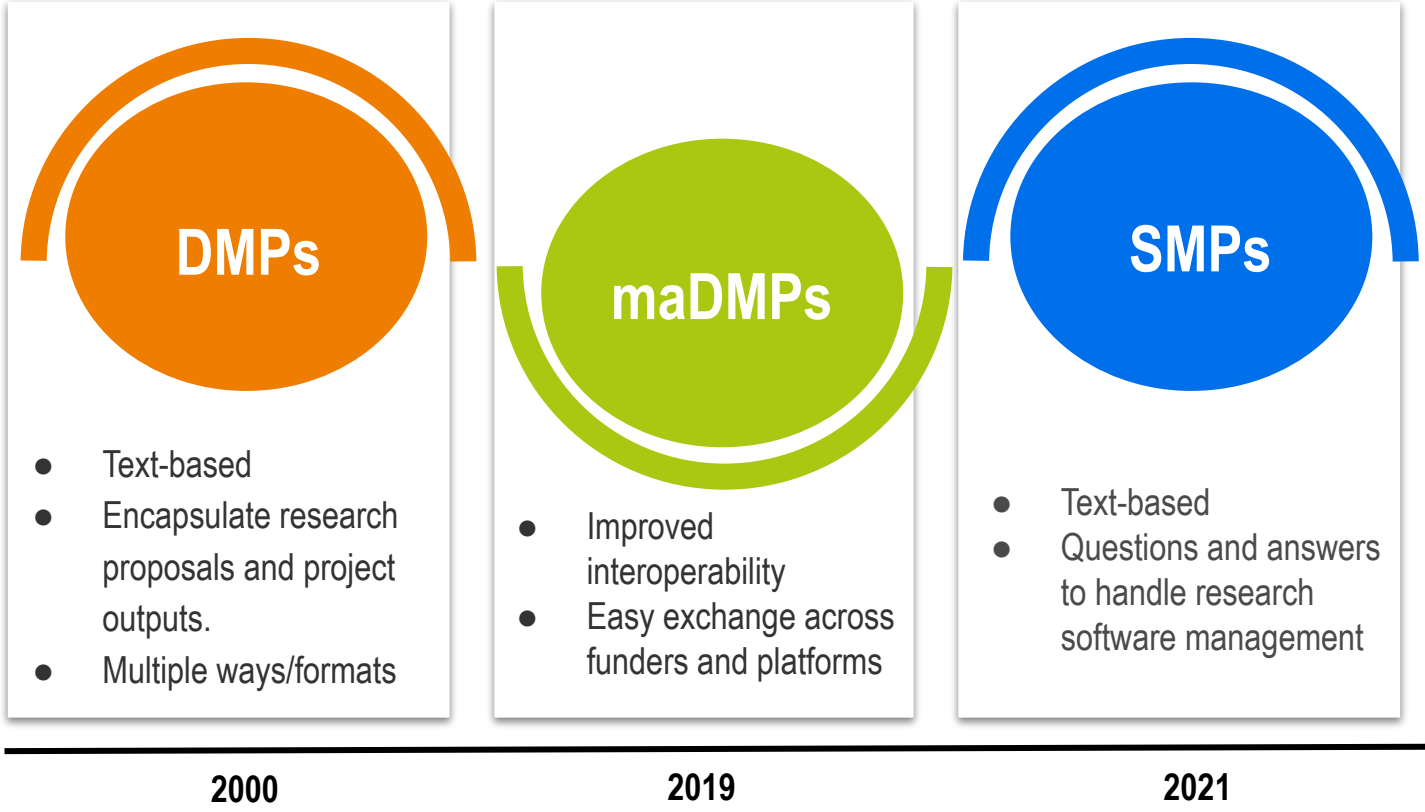
elixir

Information Centre for Life Sciences

ZB MED

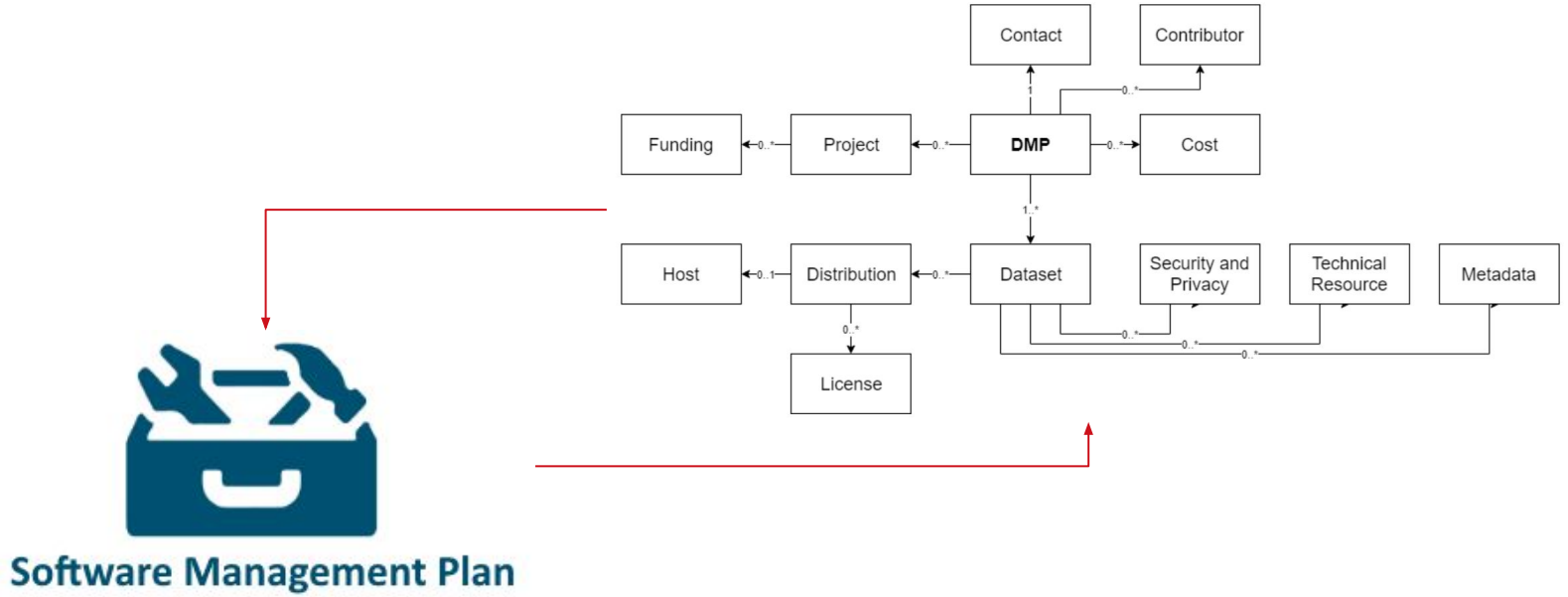
INFORMATION. KNOWLEDGE. LIFE.

# Previous work

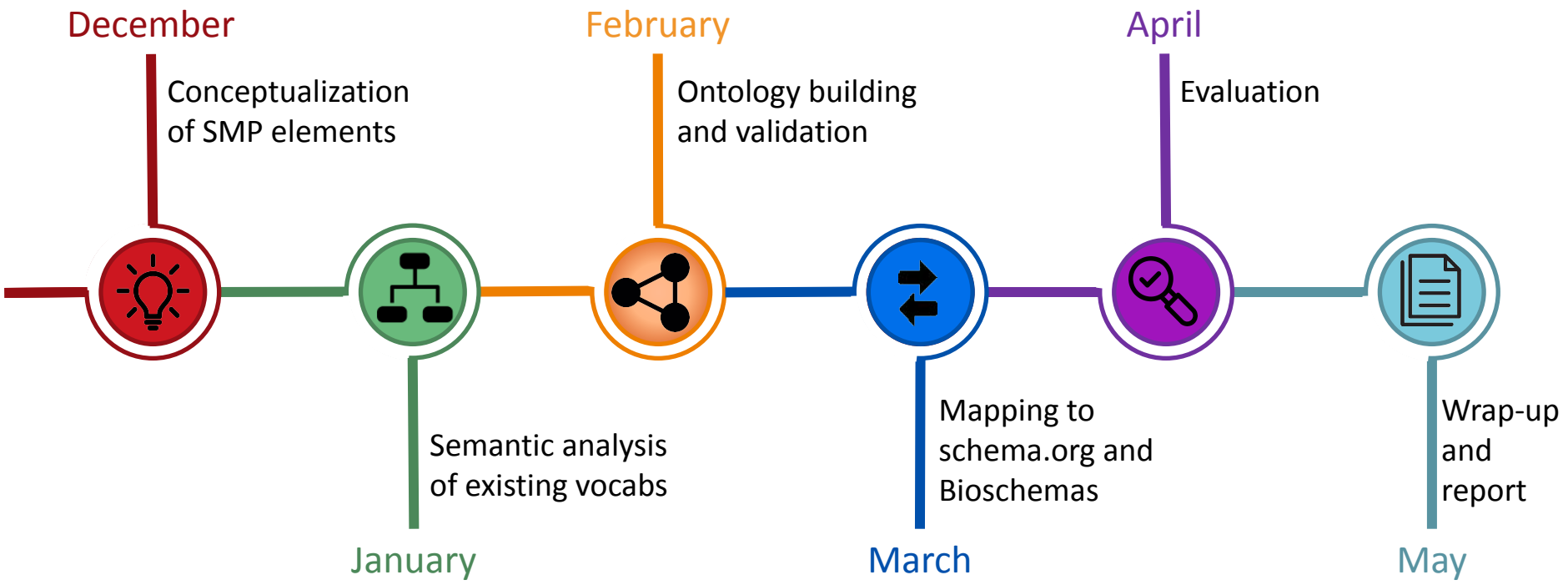


# Machine-actionable SMPs

*“Add semantics to the SMP requirements as a way to improve accessibility and usability of research software in life sciences."*



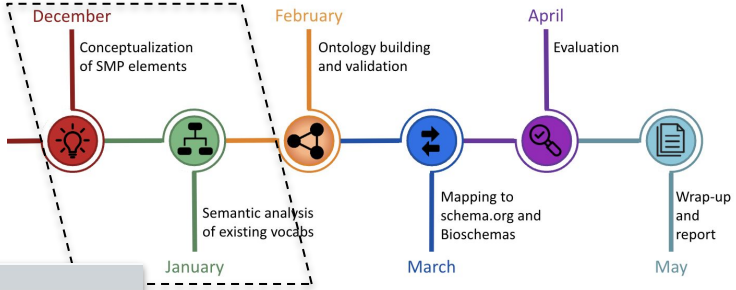
# Timeline


December
Conceptualization of SMP elements
January
Semantic analysis of existing vocabs
February
Ontology building and validation
March
Mapping to schema.org and Bioschemas
April
Evaluation
May
Wrap-up and report

# Conceptualization of SMP elements and semantic analysis

| Core requirement (Section 5.1) | Example SMP question(s) (Section 6.1) |
|---|---|
| Purpose | Please provide a brief description of your software, stating its purpose and intended audience. |
| Version control | How will you manage versioning of your software? |
| Repository | How will you make your software publicly available? If you do not plan to make it publicly available you should provide a justification. |
| User documentation | How will your software be documented for users? Please provide a link to the documentation if available.How will you document your software's contribution guidelines and governance structure? |
| Software licencing and compatibility | What licence will you give your software? How will you check that it respects the licences of libraries and dependencies it uses? |
| Deployment documentation | How will the installation requirements of your software be documented? Please provide a link to the installation documentation if available. |
| Citation | How will users of your software be able to cite your software? Please provide a link to your software citation file (CFF) if available. |
| Developer documentation | How will your software be documented for future developers? |
| Testing | How will your software be tested? Please provide a link to the (automated) testing results. |
| Software Engineering quality | Do you follow specific software quality guidelines? If yes, which ones? |
| Packaging | How will your software be packaged and distributed? Please provide a link to available packaging information (e.g. entry in a packaging registry, if available). |
| Maintenance | How do you plan to procure long term maintenance of your software? |

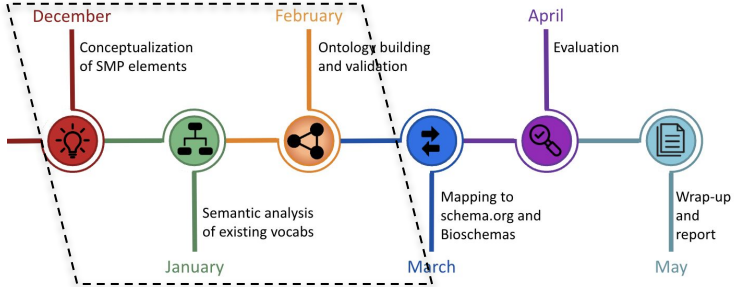
**Source:** https://doi.org/10.5281/zenodo.7248877

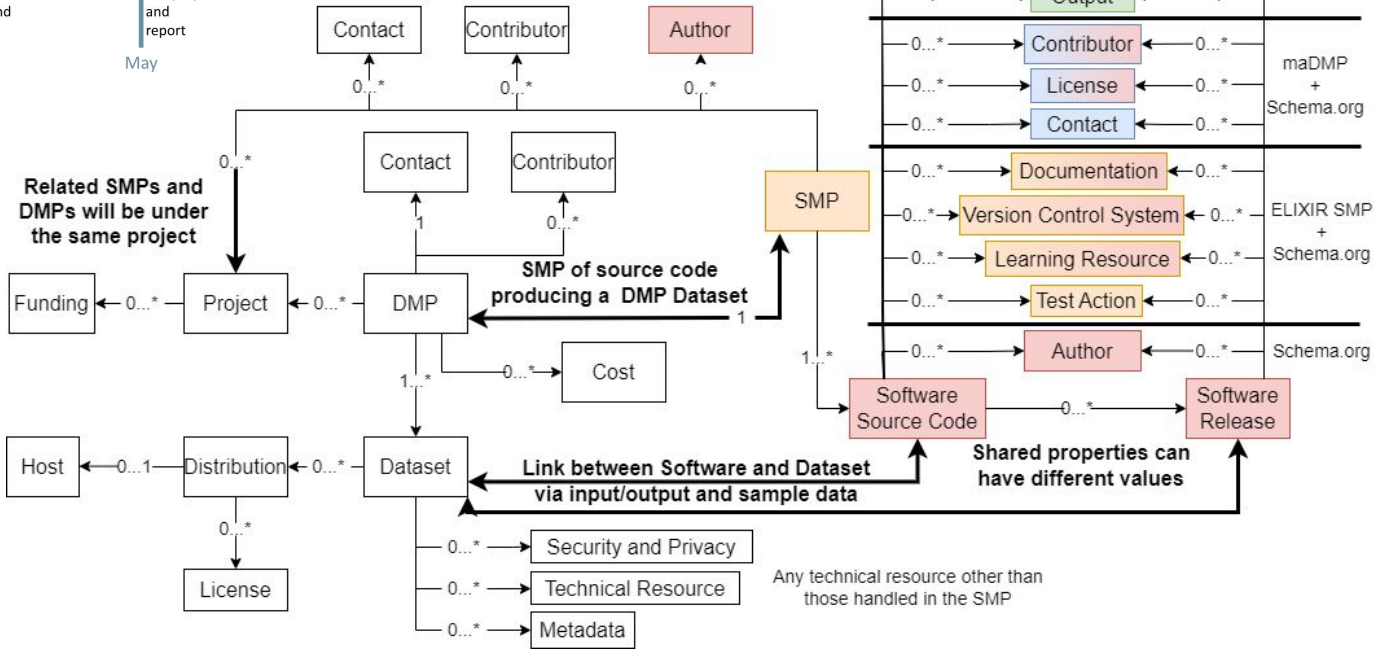



- Software Management Plan
  - Accessibility & License
    - What is the name of the software?
    - How can the software be accessed by third parties?
    - Does your software have a license?
  - Documentation
    - What type of documentation is available, provided with the s
    - Is the purpose of the software stated in the documentation?
    - Does the documentation describe how to
  - Testing
    - What type of testing do you use?
    - Are sample data and/or parameters that can be used to test
  - Interoperability
    - Do you use well-established standard input/output formats?
    - What programming languages are you using in your project?
  - Versioning
    - Do you use a version control system?
    - Do you use Semantic Versioning?

**Source:** https://smw.ds-wizard.org/

# Metadata model for a maSMP



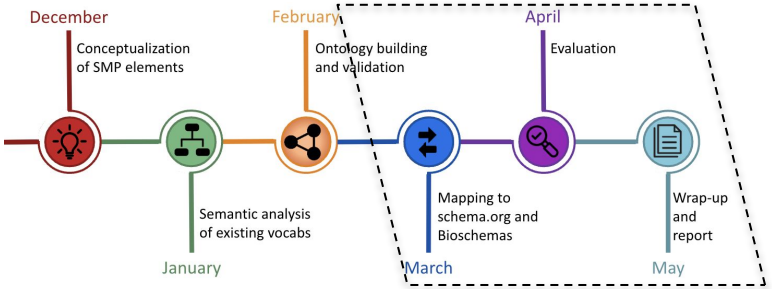




- An overview of concepts used in the metadata model for maSMPs is available at: https://github.com/zbmed-semtec/maSPMs
- The first version of maSMP ontology is available at: 10.5281/zenodo.7806638

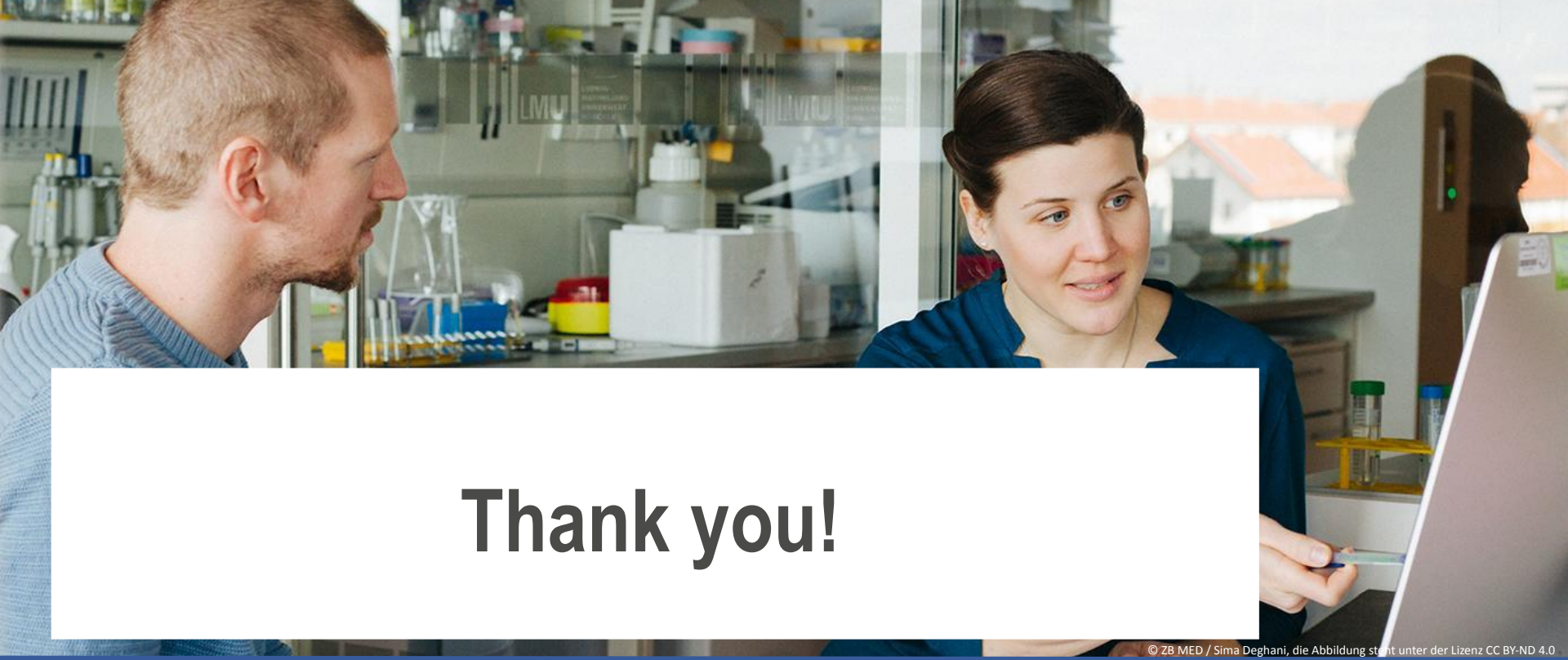
# Metadata elements for a maSMP

| | Software Source Code (aka SoftwareSourceCode in schema.org) | | Software Release (aka SoftwareApplication in schema.org) | |
|---|---|---|---|---|
| | **Property name** | **Possible values (range)** | **Property name** | **Possible values (range)** |
| From schema.org | identifier | PropertyValue, Text, URL | identifier | PropertyValue, Text, URL |
| | name | Text | name | Text |
| | description | Text | description | Text |
| | license | Text, URL | license | Text, URL |
| | author | Organization or Person | author | Organization or Person |
| | contributor | Organization or Person | contributor | Organization or Person |
| | citation | CreativeWork, Text, URL | citation | CreativeWork, Text, URL |
| | conditionsOfAccess | Text | conditionsOfAccess | Text |
| | isAccessibleForFree | Boolean | isAccessibleForFree | Boolean |
| | codeRepository | URL | releaseNotes | Text, URL |
| | programmingLanguage | ComputerLanguage, Text | memoryRequirements | Text |
| | targetProduct (aka Software Release) | SoftareApplication | operatingSystem | Text |
| | archivedAt | URL | processorRequirements | Text |
| | discussionURL | URL | storageRequirements | Text |
| | usageInfo | CreativeWork, URL | supportingData | Dataset |
| | version (i.e., semantic version) | Text | version (i.e., semantic version) | Text |
| From maDMP | hasContact | Organization or Person | hasContact | Organization or Person |
| From Bioschemas | input | FormalParameter, Dataset | input | FormalParameter, Dataset |
| | output | FormalParameter, Dataset | output | FormalParameter, Dataset |
| From maSMP (New elements) | hasAPIDocumentation | Documentation | hasAPIDocumentation | Documentation |
| | hasDeveloperDocumentation | Documentation | hasDeveloperDocumentation | Documentation |
| | hasUserDocumentation | Documentation | hasUserDocumentation | Documentation |
| | hasLearningResource | LearningReource | hasLearningResource | LearningReource |
| | hasVersionControlSystem | SoftwareApplication | hasVersionControlSystem | SoftwareApplication |
| | hasReadme | URL | hasReadme | URL |
| | testedWith | TestAction | testedWith | TestAction |

# Next step




- Carry out  a workshop with experts in Software Management Plans and machine-actionable Data Management Plans
  - To validate our maSMP model with them
  - To align different efforts
- Promote adoption
- Align to metadata related to Data Sciences and Artificial Intelligence on the NFDI4DataScience German project



# Thank you!



DaMaLOS 2023

RDA RESEARCH DATA ALLIANCE

EOSC Future

elixir

ZB MED Information Centre for Life Sciences

INFORMATION. KNOWLEDGE. LIFE.