

Management of Metadata Types in Basic Cardiological Research

Harald KUSCH ^{a,b,1}, Robert KOSSEN ^a, Markus SUHR ^a, Luca FRECKMANN ^a,
Linus WEBER ^a, Christian HENKE ^a, Christoph LEHMANN ^a,
Sophia RHEINLÄNDER ^a, Georg ASCHENBRANDT ^a, Lea K. KÜHLBORN ^a,
Bartłomiej MARZEC ^a, Julia MENZEL ^{a,c}, Blanche SCHWAPPACH ^{b,c},
Laura C. ZELARAYÁN ^{d,e}, Lukas CYGANEK ^{b,c,f}, Gregory ANTONIOS ^f,
Tobias KOHL ^f, Stephan E. LEHNART ^{e,f}, Marcel ZOREMBA ^d, Ulrich SAX ^a, and
Sara Y. NUSSBECK ^{a,g}

^a *Department of Medical Informatics, University Medical Center Göttingen (UMG), Göttingen, Germany*

^b *Cluster of Excellence “Multiscale Bioimaging: from Molecular Machines to Networks of Excitable Cells” (MBExC), University of Göttingen, Göttingen, Germany*

^c *Department of Molecular Biology, UMG, Göttingen, Germany*

^d *Institute of Pharmacology and Toxicology, UMG, Göttingen, Germany*

^e *German Center for Cardiovascular Research (DZHK), Partner Site Göttingen, Göttingen, Germany*

^f *Heart Research Center Göttingen, Clinic for Cardiology and Pneumology, UMG, Göttingen, Germany*

^g *Central Biobank UMG, UMG, Göttingen, Germany*

Abstract

Introduction: Ensuring scientific reproducibility and compliance with documentation guidelines of funding bodies and journals is a topic of greatly increasing importance in biomedical research. Failure to comply, or unawareness of documentation standards can have adverse effects on the translation of research into patient treatments, as well as economic implications. In the context of the German Research Foundation-funded collaborative research center (CRC) 1002, an IT-infrastructure sub-project was designed. Its goal has been to establish standardized metadata documentation and information exchange benefitting the participating research groups with minimal additional documentation efforts. **Methods:** Implementation of the self-developed menoci-based research data platform (RDP) was driven by close communication and collaboration with researchers as early adopters and experts. Requirements analysis and concept development involved in person observation of experimental procedures, interviews and collaboration with researchers and experts, as well as the investigation of available and applicable metadata standards and tools. The Drupal-based RDP features distinct modules for the different documented data and workflow types, and both the development and the types of collected metadata were continuously reviewed and evaluated with the early adopters. **Results:** The menoci-based RDP allows for standardized documentation, sharing and cross-referencing of different data types, workflows, and scientific publications. Different modules have been implemented for specific data types and workflows, allowing for the enrichment of entries with specific metadata and linking to further relevant entries in different modules. **Discussion:**

¹ Harald Kusch, Department of Medical Informatics, University Medical Center Göttingen, Robert-Koch-Str. 40, 37075 Göttingen, Germany; E-mail: harald.kusch@med.uni-goettingen.de.

Taking the workflows and datasets of the frequently involved experimental service projects as a starting point for (meta-)data types to overcome irreproducibility of research data, results in increased benefits for researchers with minimized efforts. While the menoci-based RDP with its data models and metadata schema was originally developed in a cardiological context, it has been implemented and extended to other consortia at Göttingen Campus and beyond in different life science research areas.

Keywords. Metadata schema, research data platform, data management

1. Introduction

Interventional clinical studies and trials involving human beings are subject to regulatory and legal framework conditions and require a highly standardized documentation. In contrast, biomedical basic research relying on model organisms is very outcome open and the documentation is usually not strictly standardized. The lack of standardized documentation strategies is in most cases not due to non-availability of standards. Many standards exist and the MIBBI project [1] is only one example of reporting guidelines for biological and biomedical investigations, which can be found at <https://fairsharing.org/collection/MIBBI>. However, very often researchers are not aware of existing standards in their area of research [2] and even if they are aware of standards, it might be too difficult to apply or use them. Utilities to enable or facilitate implementation of standards such as checklists or templates do not always exist.

These three effects, among others, can add up to irreproducible research results, which lead to a bottleneck in or a crisis of translational research of preclinical results towards the treatment of patients [3]. In addition, there is an economic effect to the irreproducibility, which according to Freedman et al was [4] estimated to 28.2 billion US\$ annually.

However, the problem is heterogeneous as the irreproducibility is caused by very different factors [4,5], which require diverse approaches to overcome them. For example, many journals have updated their policies to describe how they expect studies to be conducted [6]. In parallel, major funding agencies including the German Research Foundation (DFG) have developed and updated their guidelines for safeguarding of good scientific practice, which require the archiving of research data for at least ten years².

The lack of reproducibility is also affecting cardiovascular sciences [3]. To overcome some of the aspects mentioned above, the DFG-funded collaborative research center (CRC) 1002 established an IT-infrastructure (INF) sub-project³ to include the medical informatics expertise in data management into this cardiovascular basic science 12-year project, which started in 2012. So far, the expertise mostly originated from developing and operating IT-infrastructures for clinical non-interventional and interventional studies, such as studies and registries within the Competence Network for Congenital Heart Diseases [7] and the clinical research group KFO179⁴ [8]. Additionally, a close collaboration with the research data officer⁵ of the University of Göttingen as

² https://www.dfg.de/download/pdf/dfg_im_profil/reden_stellungnahmen/download/empfehlung_wiss_praxis_1310.pdf

³ <https://gepris.dfg.de/gepris/projekt/223309987>

⁴ <https://gepris.dfg.de/gepris/projekt/50725709>

⁵ <https://www.uni-goettingen.de/en/221506.html>

well as with the Göttingen Campus-wide eResearch Alliance⁶ provides options for sustainability of harmonized approaches. Since 2017, the CRC 1002 also includes a clinical observational intervention study [9] recruiting patients undergoing transcatheter aortic valve replacement.

With a prevalence of 2-7% in the European and North American population aged above 65, aortic valve stenosis has become the most frequent valvular heart disease [10]. Within the CRC 1002, researchers from 17 individual research groups, together with two service projects, investigate modulatory units in heart failure to identify targets for future therapeutic approaches. For heart failure, the field of basic research is very broad as reflected by the background of the different participating research groups: biophysics, imaging, small versus large animal models, proteomics, genomics, functional microdomains, etc. Thus, a huge variety of complex research data is produced.

From a medical informatics perspective, the basic approach to improve data management in order to avoid irreproducibility of research results, is to generate detailed metadata about the experiments themselves, to provide and enable long-term interpretations of research results and secondary use, which is especially relevant for interdisciplinary research collaborations. Due to intense campus-wide interactions in Göttingen, the data management approaches reported here are principally transferable and applicable to other initiatives and research areas following the FAIR (Findable, Accessible, Interoperable, Reproducible) guiding principles [11].

The overarching research question of this IT-infrastructure sub-project has been: how can the standardized documentation of CRC 1002 metadata and information exchange be of benefit regarding reproducibility of research results for the participating groups while introducing minimal additional documentation effort? The two derived, more precise sub-questions to be answered are: which IT tools can support this task and which metadata standards can be used for this purpose in a (partially) automated manner?

2. Methods

2.1. Requirements analysis, concept development, and implementation

Regarding the implementation of our self-developed Drupal-based menoci Research Data Platform (RDP) [12], essential core and driver of the requirements engineering according to Pohl [13] was the close collaboration with selected pilot experts and early adopters of a standardized documentation workflow. Most of the projects make use of the two service projects in the CRC 1002 (S01, “In Vivo and In Vitro Disease Models”, S02 “High resolution fluorescence microscopy and integrative data analysis”). This approach facilitated the contact to various “users” of the service projects as information providers and led to a prioritization of data types like advanced light microscopy and nanoscopy (ALMN), antibodies, mouse echocardiography, and human stem cells.

⁶ <https://www.eresearch.uni-goettingen.de/>



Figure 1. Schematic overview of the standardized analyses of workflows for new menoci modules

Investigation of the corresponding workflows and documentation processes was initiated as depicted in figure 1. The collection of requirements was assembled by four main aspects. The first steps included the direct observation of complex experimental procedures and their documentation as well as comprehensive stakeholder interviews. For the latter two steps search portals like biosharing.org or scholar.google were used. Search categories included official standards (e.g., ISO, Dublin Core, ISA tab), community/de facto standards, and local common or consented schemas [14] or standard operating procedures (SOPs). A list of investigated standardizing resources are available as supplemental table 1.

From the collected catalogue of requirements, aspects were prioritized and selected to be further elaborated into concepts for implementation. For each selected data type, modules were designed and integrated into the Drupal-based menoci framework. The decision process and comparison of alternative frameworks as well as detailed technical aspects of the implementations are described by Suhr et al. [12]. In brief, the software development process followed the agile development paradigm integrating user (selected pilot experts and early adopters) feedback and continuous evaluation within the development cycles. Developments and entered metadata were evaluated and enhanced by an agile approach in data and software review meetings with the help of the early adopters from within the consortium.

3. Results

Aiming at the implementation of the FAIR guiding principles, the menoci-based Research Data Platform enables a unique data management approach of consortia- and publications-centric representations to cross-link research metadata and data collections of general as well as specific data types by applying available technological standards as well as metadata standards. Although started in the cardiological setting, the developments focused on data types that are relevant for many biomedical research consortia in order to facilitate a broad re-use potential of menoci functionality.

3.1. Published data registry

The central pillar of the menoci-based RDP is the “Published Data Registry” (see supplemental figure 1). It is a registry of publications generated by the researchers of the CRC 1002. Each publication entry provides information and metadata specific to the respective article, cross-links to relevant assets in other menoci modules, and directly links to different external resources (see figure 2).

Publication

Property	Value
Working Groups	AG Hasenfuß, AG T. Fischer, AG Toischer, AG Zelarayan
Subproject	A11, CD7, S01, S02, INF
Open Access	Yes
Publication Type	Journal Article
PMID	29394407
DOI	10.1093/marigly049
Publication Year	2018
Title	A context-specific cardiac β -catenin and GATA4 interaction influences TCF7L2 occupancy and remodels chromatin driving disease progression in the adult heart.
Journal	Nucleic acids research
Authors	Iyer LM, Nagarajan S, Woelfler M, Schoger E, Khadjeh S, Zafeiriou MP, Kari V, Herting J, Pang ST, Weber T, Rathjens FS, Fischer TH, Toischer K, Hasenfuß G, Noack C, Johnsen SA, Zelarayan LC

External Resources

- SRP103273 Mouse sequencing raw data (sequence read archive, SRA)
- SCR 003070 SciCrunch identifier (RRID:SCR_003070, ImageJ)
- SCR 015687 SciCrunch identifier (RRID:SCR_015687, DESeq2)
- SCR 015807 SciCrunch identifier (RRID:SCR_015807, GraphPad Prism)
- academic_081006 Supplemental material

Linked Antibodies

PID	AG	Type	Antigen symbol	Antibody Registry ID	Name
umg-afb11022-antibody-primary-234	11	primary	SHISA3	AB_2682593	Anti-SHISA3 antibody produced in rabbit
umg-afb1002-antibody-primary-235	12	primary	TCF7L2	AB_10734317	ChIPab+ TCF-4 antibody

Linked Echo-Studies

Study Name	Study Owner
C07_merCremer_TAC_MW4	Laura Zelarayan-Behrend

Linked Mouselines

Research Group	PID	Name	Created
11	umg-mic-mouseline-0016	B6	2020-01-29
12	umg-mic-mouseline-0018	D6.FVB(129)A1q[9]M6.0e[0]r1[1]g[0]k[1]	2020-02-08

Figure 2. Excerpt of a single publication from the Published Data Registry, illustrating the collected metadata and information. The entry for each registered publication provides specific data regarding the article, such as the publication type, DOI, journal, publication year. Furthermore, direct links are given to both external resources as well as entries in other modules of the menoci-based Research Data Platform (here: entries in the Antibody Catalogue, the Echocardiography and Mouseline Catalogue)

When registering a new publication, basic metadata can be automatically imported and subsequently manually corrected and enhanced by using the EuropePMC, Crossref and DataCite application programming interfaces and entering the article specific PubMed ID or digital object identifier (DOI). Moreover, the manual entry of metadata is also an option. For the basic metadata set in this registry see either figure 2 or the projects website <https://menoci.io>.

In addition to the automatically extracted bibliographic metadata, information on (sub-)projects, associated research groups or the open access status can be set. The section “external resources” of each publication entry allows to standardize the cross-linking to relevant resources (e.g., material, code, datasets), by providing direct URLs or other resolvable pairs of base URL and (persistent) identifier, such as DOIs, SciCrunch IDs, NCBI IDs (e.g. GEO) or ORCID iDs (for details see supplemental figure 2 and supplemental table 2).

Data types and materials that were regarded as particularly important to enhance reproducibility of CRC 1002 experiments such as antibodies, mouse lines, cell lines, ALMN and echo data can be cross-linked via the edit view of each publication and are explained below in more detail.

3.2. Catalogue modules

A set of different modules in the menoci-based RDP allows for the documentation, storing, and sharing of specific data types, such as model organisms, research objects or biological specimens. While each of these modules is customized to provide specific features and relevant information for each of the data types, they share common properties and purposes: Each module e.g. provides the ability for a standardized documentation of the respective data types, the collection of relevant metadata and the connection of entries to specific articles in the Published Data Registry (see 3.1). Accessibility of metadata in the catalogue modules can be differentiated into three levels to enable the internal management of sensitive information. In the context of the CRC 1002, catalogue modules have been implemented for induced pluripotent stem cell (iPSC) lines, transgenic mouse lines, and antibodies. These catalogue modules and the (meta-)data collected therein will be described in more detail below.

3.2.1. Model organism modules: Cell models and mouse lines

Various experimental approaches in CRC 1002 research make use of differentiated cells derived from human iPSC lines generated from patient biopsies or from CRISPR-engineered iPSC models. Therefore, documentation and tracking of relevant metadata describing the provenance of involved cell models is important for reproducibility. The metadata collection which integrates general data of each cell line with the clinical data of the respective donor is available via the Cell Model Catalogue and involves standardization according to the Human Pluripotent Stem Cell Registry (hPSC^{reg7}) and interfaces to hPSC^{reg} and the biospecimen management system STARLIMS⁸ used in the Central Biobank UMG (for metadata details see supplemental table 3). The degree of metadata standardization was enhanced by the implementation of dropdown menus or radio buttons with predefined but extensible value sets in the web forms for data entry. Mouse line-related metadata is documented according to the ARRIVE⁹ and MGI¹⁰ guidelines (e.g., breeding type, strain name, genetic mutations) and is organized in the Mouse Line Catalogue.

3.2.2. Material module: Antibodies

The Antibody Catalogue provides the ability to document and search both primary and secondary antibodies, enriched with a detailed set of metadata such as the targeted antigen, the manufacturer, storage instructions, and the clonality (see supplemental figure 3 and supplemental table 4 for details). As described for the Cell Model Catalogue, the web forms for antibody data entry contain recommendations or predefined values sets, where applicable, to enforce standardization of metadata. This local data collection is complemented by external links to databases such as the Antibody Registry and GeneCards. Furthermore, researchers can add quality information about an antibody in different applications (such as Western blot or immunohistochemistry imaging) in terms of a simple star rating system. Antibody data can be exported in different structured formats (such as json or csv) that enhance re-usability and interoperability.

⁷ <https://hpscreg.eu/>

⁸ <https://www.informatics.abbott/>

⁹ <https://arriveguidelines.org/>

¹⁰ <http://www.informatics.jax.org/mgihome/nomen/gene.shtml>

3.3. Workflow support modules

Complementing the catalogue modules described above, menoci workflow modules were designed on the one hand to enable the deep documentation and teaching of complex experimental approaches and research procedures, and on the other hand for planning, conducting and data storing of two different imaging techniques: ALMN and Mouse Echocardiography. A more detailed description of the technical workflow employed in these modules can be found in Suhr et al. [12]. In short, a request for an imaging experiment can be put forth by researchers which will then be performed or supported by a corresponding service project of the CRC 1002. During the process, relevant experimental and image specific data and metadata are collected and made available to the consortium. For details on the metadata collected see <https://menoci.io> and supplemental materials.

3.4. Evaluation of the modules

In addition to the agile development process including iterative fine granular evaluations, the content of the menoci-based RDP instances is regularly monitored and summarized to facilitate data review tasks. The current status (22.03.2021) of data content for different research consortia is summarized in table 1 and supplemental table 5.

Table 1. Count of annotated entries in the released instances of menoci-based research data platform.

instance	started in	Publications	External resource ¹	Antibodies	Mouse lines	Cell models	Echo studies	Fly stocks
CRC 1002	2015	366	2597	1043*	34*	510	491	-
CRC 1190	2016	139	1608	1043*	34*	-	-	-
CRC 1286	2018	111	1289	209	-	-	-	-
EXC 2067	2020	237	2238	184	-	-	-	-
FOR 2705	2021	33	155	466	-	-	-	10
TRR 274	2021	16	235	24	1	-	-	-

¹ The different external resource types are shown in detail in the supplemental material.

* Antibody and mouse line databases are shared among CRC 1002 and 1190.

4. Discussion

Since 2012 three different types of modules have been developed within the menoci framework serving different purposes. Our experience shows that the largest benefit regarding reproducibility and effort is reached for all participating researchers, when data generated by service projects is chosen for standardization. The IT-tools developed for this purpose in close collaboration with the personnel of the two service projects, were the two workflow support modules (echocardiography as well as ALMN). These collaborations led to the development and use of three additional intertwined catalogue modules (antibodies, stem cells, mouse lines). Measures to minimize the additional documentation effort include the development and implementation of a data importer for already documented antibodies in the different labs from Excel sheets and for publication information from PubMed. Moreover, a semi-automated data export from the Cell Model Catalogue to the European Stem Cell Registry hPSC^{CRG} was realized. Metadata standardizing approaches such as Dublin Core or Antibodyregistry were applied as far as applicable (see Supplemental table 1). Altogether, these developed IT tools as well as the semi-automated use of metadata standards benefitted the participating research

groups with regard to reproducibility of research results while at the same time the additional documentation effort was kept low. At the moment, the most numerous type of annotated entries across all menoci instances are antibodies, as their exact description is key to reproducibility [15]. In addition, INF student assistants help in digitalizing and linking data out of the participating research groups and their conducted experiments into the menoci platform. At the beginning, a lab notebook registry was developed to link the paper-based notebooks containing descriptions of experiments via a sticker to the respective published data and manuscripts. Meanwhile, many participating research groups use or transition to an electronic laboratory notebook documentation [16], making the use of the lab notebook registry less important. As requirements of the researchers are subject to constant change, a next generation sequencing (NGS) module is currently under development.

A variety of research data management (RDM) frameworks serving different application purposes are available. We selected the generic Drupal framework with a great potential to combine readily available web portal functionality with the possibility to add specific modules for special requirements. A comparison of functional categories of menoci and other RDM frameworks (such as OpenBis or Seek) has been described previously [12].

The described data models have been developed for the purpose of managing experimental data of basic heart research within CRC 1002. Due to the good network fostered by the eResearch Alliance at Göttingen Campus, these developments and the menoci platform have been successfully reused by other research consortia with different backgrounds e.g., neurosciences, cellular biology (see table 1). Moreover, it is planned to use the platform in other settings, such as the CURE-PlaN network. The antibody database of CRC 1002 was subsequently developed as a shared resource with the biochemistry CRC 1190, as several investigators participate in both CRCs. Thus, not only data models but also metadata, results, and infrastructure is deeply reused at this CRC IT interface. menoci core modules (e.g., published data registry, antibody catalogue) have also been implemented in the neuroscience consortia FOR 2705 and TRR 274. All menoci developments are oriented along the FAIR guiding principles. Findability of menoci data is enhanced by e.g., the application of persistent identifiers and metadata enrichment where applicable. Accessibility is regulated by the role and rights model of menoci. menoci datasets can be enriched with structured and standardized metadata as far as standardizing resources are available for the respective data types. Altogether, these measures increase the re-usability potential of menoci datasets. However, complete FAIR compliance remains to be challenging: e.g. complete standardized machine readability is a complex task and yet only partially implemented in menoci. Technical and organizational aspects of data review and quality improvement are summarized in supplemental table 6.

Challenges that can occur during any software development project, such as feature overload, insufficient stakeholder engagement or insufficient communication during the initial phases of the project, were strongly mitigated by the involvement of pilot users/early adopters, the requirements analysis and close observation/investigation of workflows, as described in the methods section. Furthermore, in-house development of the platform allowed us to minimize dependencies on external/licensed software and libraries, enabling a targeted approach to the researchers requirements without relying on external developers/updates, while the use of the Drupal backbone helped us to reduce the time spent on basic development.

However, more specific challenges in the context of basic biomedical research IT infrastructures include the often complex and very specific types of metadata and processes, as well as securing knowledge about these in the long term. To address these potential problems, we not only employed documentation and tracking tools (GitLab based issue management, Wiki entries), but also performed detailed research, description and discussion of the metadata, data types and processes with the biomedical experts of the research consortia.

The menoci-RDP as central point of information for CRC 1002 research data builds a starting point for an extended consortium-related knowledge graph by cross-linking metadata via unambiguous and/or persistent identifiers (supplemental figure 03). Especially linking out to Wikidata, which was initially enabled through the RDP, has the potential to connect CRC 1002 data with a cross-disciplinary and -border knowledge base.

5. Data availability

All additional data and materials are available in the research data repository of Göttingen Campus (<https://doi.org/10.25625/TISZMJ>) and via the project website: <https://menoci.io>

6. Conflict of interest/contributions of authors and funding

All authors declare that no conflict of interest exists. HK, RK, MS, BSP, LCZ, SEL, US, SYN elaborated the concepts and wrote the manuscript. MS, LF, LW, CH, CL, SR, GAs, LKK, BM, JM, LC, GAn, TK, MZ elaborated the concepts and developed the software. The human resources for conceptualization, development, and continuous improvement of the menoci platform were mainly funded by the DFG through project funding of the CRC 1002 infrastructure (INF) project, A09, S01, and S02 as well as the Z projects of CRC 1190 and CRC 1286, and Germany's Excellence Strategy - EXC 2067/1-390729940. SEL is a principal investigator of DZHK, partner site Göttingen; and the Transatlantic Excellence Network CURE-PLaN funded by Foundation Leducq Paris.

References

- [1] C.F. Taylor, D. Field, S.-A. Sansone, J. Aerts, R. Apweiler, M. Ashburner, C.A. Ball, P.-A. Binz, M. Bogue, T. Booth, A. Brazma, R.R. Brinkman, A. Michael Clark, E.W. Deutsch, O. Fiehn, J. Fostel, P. Ghazal, F. Gibson, T. Gray, G. Grimes, J.M. Hancock, N.W. Hardy, H. Hermjakob, R.K. Julian, M. Kane, C. Kettner, C. Kinsinger, E. Kolker, M. Kuiper, N. Le Novère, J. Leebens-Mack, S.E. Lewis, P. Lord, A.-M. Mallon, N. Marthandan, H. Masuya, R. McNally, A. Mehrle, N. Morrison, S. Orchard, J. Quackenbush, J.M. Reecy, D.G. Robertson, P. Rocca-Serra, H. Rodriguez, H. Rosenfelder, J. Santoyo-Lopez, R.H. Scheuermann, D. Schober, B. Smith, J. Snape, C.J. Stoeckert, K. Tipton, P. Sterk, A. Untergasser, J. Vandesompele, And S. Wiemann, Promoting Coherent Minimum Reporting Guidelines For Biological And Biomedical Investigations: The Mibbi Project. *Nat Biotechnol* **26** (2008), 889–896.
- [2] S.Y. Nussbeck, M. Rabone, E.E. Benson, G. Droege, J. Mackenzie-Dodds, And R.T. Lawlor, "Life In Data"-Outcome Of A Multi-Disciplinary, Interactive Biobanking Conference Session On Sample Data. *Biopreservation And Biobanking* **14** (2016), 56–64.

- [3] K.R. Sipido, Irreproducible Results In Preclinical Cardiovascular Research: Opportunities In Times Of Need. *Cardiovasc Res* **115** (2019), E34-E36.
- [4] L.P. Freedman, I.M. Cockburn, And T.S. Simcoe, The Economics Of Reproducibility In Preclinical Research. *Plos Biol* **13** (2015), E1002165.
- [5] M.L. Grieneisen And M. Zhang, A Comprehensive Survey Of Retracted Articles From The Scholarly Literature. *Plos One* **7** (2012), E44118.
- [6] K.R. Sipido, B. Casadei, P. Holvoet, S. Janssens, A. Luttun, And M. Sampaolesi, Bedside To Bench: A Look At Experimental Research With A Clinical Trial Checklist. *Cardiovasc Res* **101** (2014), 1–3.
- [7] K. Helbing, S.Y. Demiroglu, F. Rakebrandt, K. Pommerening, O. Rienhoff, And U. Sax, A Data Protection Scheme For Medical Research Networks. Review After Five Years Of Operation. *Methods Inf Med* **49** (2010), 601–607.
- [8] A. Dangl, S.Y. Demiroglu, J. Gaedcke, K. Helbing, P. Jo, F. Rakebrandt, O. Rienhoff, And U. Sax, The It-Infrastructure Of A Biobank For An Academic Medical Center. *Stud Health Technol Inform* **160** (2010), 1334–1338.
- [9] B.E.C. Beuthner, R. Topci, M. Derks, T. Franke, S. Seelke, M. Puls, A. Schuster, K. Toischer, M. Valentova, L. Cyganek, E.M. Zeisberg, C. Jacobshagen, G. Hasenfuss, And S.Y. Nussbeck, Interdisciplinary Research On Aortic Valve Stenosis: A Longitudinal Collection Of Biospecimens And Clinical Data Of Patients Undergoing Transcatheter Aortic Valve Replacement. *Open Journal Of Biosources* **7** (2020).
- [10] V.T. Nkomo, J.M. Gardin, T.N. Skelton, J.S. Gottdiener, C.G. Scott, And M. Enriquez-Sarano, Burden Of Valvular Heart Diseases: A Population-Based Study. *The Lancet* **368** (2006), 1005–1011.
- [11] M.D. Wilkinson, M. Dumontier, I.J.J. Aalbersberg, G. Appleton, M. Axton, A. Baak, N. Blomberg, J.-W. Boiten, L.B. Da Silva Santos, P.E. Bourne, J. Bouwman, A.J. Brookes, T. Clark, M. Crosas, I. Dillo, O. Dumon, S. Edmunds, C.T. Evelo, R. Finkers, A. Gonzalez-Beltran, A.J.G. Gray, P. Groth, C. Goble, J.S. Grethe, J. Heringa, P.A.C. 'T Hoen, R. Hooft, T. Kuhn, R. Kok, J. Kok, S.J. Lusher, M.E. Martone, A. Mons, A.L. Packer, B. Persson, P. Rocca-Serra, M. Roos, R. Van Schaik, S.-A. Sansone, E. Schultes, T. Sengstag, T. Slater, G. Strawn, M.A. Swertz, M. Thompson, J. Van Der Lei, E. Van Mulligen, J. Velterop, A. Waagmeester, P. Wittenburg, K. Wolstencroft, J. Zhao, And B. Mons, The Fair Guiding Principles For Scientific Data Management And Stewardship. *Sci Data* **3** (2016), 160018.
- [12] M. Suhr, C. Lehmann, C.R. Bauer, T. Bender, C. Knopp, L. Freckmann, B. Öst Hansen, C. Henke, G. Aschenbrandt, L.K. Kühlborn, S. Rheinländer, L. Weber, B. Marzec, M. Hellkamp, P. Wieder, U. Sax, H. Kusch, And S.Y. Nussbeck, Menoci: Lightweight Extensible Web Portal Enhancing Data Management For Biomedical Research Projects. *Bmc Bioinformatics* **21** (2020), 582.
- [13] K. Pohl, *Requirements Engineering: Fundamentals, Principles, And Techniques*, Springer, Berlin, 2010.
- [14] J. Menzel, P. Weil, And S.Y. Nussbeck, Metadata Capture In An Electronic Notebook: How To Make It As Simple As Possible? *Gms Medizinische Informatik, Biometrie Und Epidemiologie* (2015), 1–10.
- [15] K.J. Rhodes And J.S. Trimmer, Antibodies As Valuable Neuroscience Research Tools Versus Reagents Of Mass Distraction. *J Neurosci* **26** (2006), 8017–8020.
- [16] S.Y. Nussbeck, P. Weil, J. Menzel, B. Marzec, K. Lorberg, And B. Schwappach, The Laboratory Notebook In The 21st Century: The Electronic Laboratory Notebook Would Enhance Good Scientific Practice And Increase Research Productivity. *Embo Rep.* **15** (2014), 631–634.