An annotation scheme for references to research artefacts in scientific publications.

David Schindler
Institute of Communications Engineering
University of Rostock
Rostock, Germany
David.Schindler@uni-rostock.de

Kristina Yordanova
Institute of Computer Science
University of Rostock
Rostock, Germany
Kristina.Yordanova@uni-rostock.de

Frank Krüger
Institute of Communications Engineering
University of Rostock
Rostock, Germany
Frank.Krueger@uni-rostock.de

***Abstract*—The extraction of mentions of research artefacts from scientific papers is a necessary precursor for multiple applications ranging from simple search for literature based on particular research artefacts to semantic analyses of the investigations described in the literature. Techniques of natural language processing like named entity and relation extraction allow to establish detailed knowledge about such artefacts. The application of supervised classifiers relies on annotated datasets in order to provide a basis for training and evaluation. In this work, we present an annotation scheme for research artefacts in scientific literature which not only distinguishes between different types of artefacts like datasets, software and materials but also allows for the annotation of more detailed information such as amount or concentration of materials. Furthermore, we present first preliminary results in terms of inter-rater reliability.**

***Index Terms*—annotation, named entity, open science**

I. Introduction

A large number of research artefacts is typically employed and created during scientific investigations. This includes digital artefacts like software and datasets but also physical artefacts like materials or devices. Furthermore, methodological artefacts can be distinguished that include standardized approaches or procedures or methods for data analysis. Knowledge about such artefacts is of interest for several reasons:

- to provide credit to the creators of such artefacts [15],
- to provide a measurement of impact of such artefacts [24],
- to allow replication of investigations,
- to provide a quality control mechanism with respect to artefacts, e.g. systematically analyse studies for use of flawed research artefacts such as erroneous software [15].
- to provide a mapping of overall available artefacts [23].
- to facilitate selecting suited artefacts for new study designs [23].
- to facilitate the comparison of study results [2].

These artefacts, including specific information about their usage (e.g. parameters for software or devices), are typically mentioned in the scientific publications about the investigation. They are, however, often not formally specified, but "hidden" in the textual description. The manual extraction of such references and mentions is intractable, as it would require to manually scan a huge amount of publications. A more practical way of extracting mentions of research artefacts from publication is to employ automatic methods. Unfortunately, the mentions of such artefacts are very heterogeneous and there is no universally agreed standard. Two different approaches to automatic extraction of such references exist:

- *Unsupervised methods* (e.g. iterative bootstrapping), which start with a defined set of "seed" words, extract patterns about their occurrences and iteratively analyse the text corpus in order to extend the set of words and their usage pattern. Such approaches typically achieve high precision of the recognition [2, 28].
- *Supervised approaches* (e.g. conditional random fields) that employ an annotated dataset to train a machine learning classifier to extract occurrences of scientific artefacts from the publication. While both precision and recall of supervised approaches are typically high, the application strongly depends on the annotated dataset [11, 26].

While only supervised approaches require an annotated dataset to be trained on, both approaches rely on the existence of such a dataset when it comes to the evaluation of the recognition performance of the classifier.

In this work, we introduce an annotation scheme, which allows to create labels with respect to the references to research artefacts in scientific publications. The remainder of this paper is structured as follows. We first provide an overview of different research artefacts that are of interest in typical scientific investigations and discuss their role in the automatic information extraction in Section II. Afterwards, in Section III, we introduce our annotation scheme. Finally, we present first preliminary results and conclude.

II. Mentions of Research Artefacts

Research artefacts are all components that are used by researchers in order do conduct their investigations. Scientific publications often contain a very detailed description of the research artefacts that were employed during the investigation in order to get the described results. In the following, we describe the most common types and discuss their relevance to reference extraction.

The most prominent examples of artefacts with respect to the extraction of references are data and software. Both types

This work was funded by the German Research Foundation (Deutsche Forschungsgemeinschaft, DFG) within the Collaborative Research Centre 1270 ELAINE.

978-1-5386-9151-9/19/$31.00 ©2019 IEEE

play a central role in the modern data intensive sciences. Nowadays, many scientific investigations rely on **Data**. This includes but is not limited to sensory measurement results, simulation results or images. While data sharing allows the reproduction of scientific results, the publication of data in the sense of open science is not common in all scientific communities. Recently, the number of publishers that require the publication of all data increased [29]. The kind of data publication varies from making data available on request to publishing the dataset according to the FAIR data principles [37]. These principles include the deposition of the data in public accessible repositories, which allow the long term access by of persistent identifiers as for instance digital object identifiers. While citation standards have been proposed [7], they are rarely enforced by scientific journals [25]. With respect to extracting references to data from scientific texts, the main reasons are:

- to reproduce the results of the original investigation,
- to allow meta analysis of different studies,
- to reuse data for analyses, different from the original purpose, and
- to find datasets for one's own research [23].

Similar as for the citation of data, guidelines for the citation of **Software** exist [34] (e.g. by use of the digital object identifier), but are typically cited in heterogeneous unstructured ways [27]. This ranges from the plain mention of software names in the full text to the use of hyperlinks to complete specifications of the software including the version and the manufacturer. The literature describes the major reasons for the extraction of software mentions as

- providing a reward system for software as a form of research output [27] and
- establishing a mechanism of quality control and reproducibility [15].

Tracking the use of scientific software for the purpose of providing reward to the developers is for instance done by projects like Depsy [4]. On the other hand, Eklund et al. [12] illustrate potential problems with the use of flawed software packages by use of an example for the FRMI domain [20] and thus provide additional evidence for the need of automatic extraction of mentions of such.

Beside software and data, other key artefacts of today's research are **Models**. Models are typically built upon theories [13] and are thus based on observations and assumptions about the real world. This includes data-based models as well as mathematical models in terms of equations, geometrical models, but also models for the purpose of simulation. While there are no domain independent standards for the citation of models, individual domains established standard locations for the deposition of publicly available models. One example is the BioModels database [19], which contains simulation models from the life sciences. Beside the manual reproduction of scientific investigations, the purposes of automatic extraction of references to models are:

- to provide provenance information [35] and
- to automatically discover new knowledge from automatic model variation [21].

Scientific investigations are often built upon implicit and explicit **Standards**. The application of standard procedures not only builds a strong basis for scientific investigations but also allows the comparison and summarization of their respective results. Standards range from lab specific standard operating procedures that regulate how to use particular devices to domain specific or international recognized methods to ensure the quality of processes. Here, we consider standards to include all kinds of structural and methodological guidelines. This includes standardized methods for data exchange and presentation such as file formats, but also normed workflows, common assessments and questionnaires, and statistical tests. The extraction of standards is of interest for the following reasons:

- to assess the quality and compare different investigations
- to summarise the setup of clinical studies [9]
- to investigate the incorrect use of statistical tests [16]

One major source of data in scientific investigations are **Devices**, which thus play a central role in the quality of the data itself. Devices are typically used to measure physical effects in order to analyse the generating process. Typical examples of devices are sensors, such as accelerometers or thermometers, but also microscopes and MRT scanners. Another category of devices are actuators that do not sense but influence their environment. Actuators are often used together with sensors as the effect of the influence is measured. Examples for such actuators include devices to generate physical forces to specimen. Similar as for standards, the purposes of extracting information about devices include:

- to assess the data quality,
- to establish comparability of different investigations, and
- to provide formalized information about scientific investigations [8]

Apart from device use in scientific publications [1], the extraction of information about devices is also investigated in the context of electronic health records [31].

Materials refer to physical materials or substances that were particularly produced or collected, either by the researchers themselves or by external manufacturers. In the biomedical domain, materials include drugs, antibodies, cell lines and other biochemical substances. Other examples include the blend of ingredients in the material sciences or ground samples from the agricultural sciences. Materials are often not subject to standardized citation guidelines. They are typically mentioned by the their name, their general type and their manufacturer. However, recently publication and metadata standards were proposed to improve this [10]. The literature reports interest in information about materials from scientific publication for the following reasons:

- to improve the development of new drugs [36],
- to provide formalized knowledge about drug usages [6, 18] and
- to build systematic maps of drug-disease relationships [30]

In the following section we provide an overview of our annotation scheme.

III. Annotation of Research Artefact

A. Types of mentions

With respect to the mention of research artefacts in scientific publications, we distinguish four types: *Allusion, Usage, Creation* and *Deposition*. The semantic meaning of the mention type depends on the context only.

The mention type *Allusion* alone allows to determine which artefacts are generally present in the text. It could be used for overall mappings of research artefacts. The more restricted type *Usage* allows to answer which artefacts actually contributed to an investigation. This corresponds to a re-use of a given artefacts and could be used to establish a credit system or impact measure. The type *Creation* makes the distinction between newly created and existing artefacts. This creates the possibility of mapping new creations and tracing original publications of artefacts. The last type *Deposition* is quite similar to *Creation* with the added constraint of the artefacts public availability.

Allusions refer to research artefacts that are described in the text but not necessarily used in the investigations. In general, it is not possible to conclude whether the artefact was used or not. Allusions are typically used in reference to other studies to describe the current state of the art.

The following example illustrates how the material *anti-TNF-*α is mentioned with respect to a general effect. As can be seen it is not possible to determine whether the material was actually used in the published investigation.

Example 1: There is some evidence that depressive symptoms may be reduced in CD patients after infusion of anti-TNF-α, [...]. (from Guloksuz et al. [14])

In contrast to Allusions, artefact **Usage** refers to the actual application in the investigation being described. This typically means that the research artefact makes a necessary contribution to the investigation.

In the following, an example for the usage of a material is given. Here, the administration of anti-TNF-α is described, which clearly signals a usage in the context of the study.

Example 2: Administration of anti-TNF-α significantly increased scores on all dimensions of quality of life and reduced depression scores measured by the HDRS, BDI and SCL-90 (Table 3). (from Guloksuz et al. [14])

The distinction between usage and allusion can be captured by the question whether the authors of the study were able to conduct the study without research artefact.

The publication of research artefacts that were created in the context of the study is accompanied with statements about the actual location of the artefact. Such **Deposition** statements allow access to these artefacts. Such statements often refer to research artefacts that are particularly developed with the purpose of the study.

In the following example, a software tool was developed but also made available by providing an URL.

Example 3: Here we develop a simulator, TABASCO http://openwetware.org/wiki/TABASCO, which enables [...] (from Kosuri et al. [17])

Creation refers to research artefacts that were created for the purpose of the investigation presented in the publication. In contrast to deposition statements, creation statements do not directly allow access to the respective artefact.

The software, which is mentioned in the following example, was developed for the purpose of the study but it can not directly be concluded from the sentence whether it is published.

Example 4: We designed and developed Tabasco to revisit approximations encoded within previously available gene expression simulation algorithms [...] (from Kosuri et al. [17])

B. Annotation of artefacts

To capture additional information about research artefacts, the annotation scheme includes additional categories. All additional information are optional, as typically not all information is available from the textual description. A description including examples is provided in the following.

With respect to **data**, it is of interest, whether only parts of dataset or even entire collections of datasets were used in the study. This aspect is covered by the annotation tag **Granularity** which allows the specification of *Database* for collections of datasets, *dataset* for individual datasets and *subset* for parts of the respective dataset. The **Source URL** allows to annotate the location of a dataset, which might be an URL or a deposition identifier. The **Type** of the location is further specified by *Repositories*, *Database*, *Institutional Archives*, or *Personal Servers* and **Reference** provides information about how the data is referred to, as for instance a description in *data articles*. **Creator** is used to annotate the author of the data and **Version** for the version of the dataset.

Example 5: The Opportunity dataset [59] contains sensor-based data [...] To evaluate [...] we used the body-worn sensors only and ignored other sensors in our experiments." (from Liu et al. [22])
The term "Opportunity" in the first sentence is an allusion of a dataset, which provides details about the dataset itself. A reference to the original publication is provided. The further specification of "body-worn sensors" is a usage statement with refers to a subset of the Opportunity dataset. This association has to be labelled by use of a relation.

Like for data, additional information which further specify the reference to the **software** are included in the annotation scheme. In fact, software uses the same types of additional information as data, except the field **Type**. Also for the label **Granularity** different options make up the set of possible values: *Basic Software*, *Package/Extension* and *Specific Function*.

The example below, illustrates the usage of the software "R" with is further specified by use of the version "3.3.1.".

Example 6: All analyses were performed with R version 3.3.1. (from Salatzki et al. [33])

It has to be noted that sole mention of programming languages, like Java or Python are excluded from the annotation of software. However, frameworks or packages that are associated with the programming language (as in the above example) are included.

Similar to software and data, references to **Models** are often further specified in publications. Apart from the additional information known from software, the details of a model include the **Specific Type** which provides information about the actual implementation of the model. Possible values include modelling frameworks or modelling formalisms.

Example 7: To investigate the effects of PV length and diameter on arrhythmia inducibility and arrhythmia dynamics, bi-atrial bilayer meshes were constructed from MRI data for twelve patients. (from Roney et al. [32])
The term "bi-atrial bilayer" is considered as model as allows to map the MRI-data with respect to computational investigations. The term "meshes" is considered as the specific type of the model.

As for the previous artefacts, the annotation for **Standards** also includes the detail labels: **Reference**, **Creator**, and **Version**. In the following example, statistical tests like ANOVA and t-tests are annotated as standards.

Example 8: To compare differences between groups of mice phenotypes, we performed two-way ANOVA (Bonferroni post test), [...], and data was analysed with GraphPad Prism Software. (from Salatzki et al. [33])

Annotation for **Devices** comprises the fields **Reference**, **Creator (Manufacturer)**, and **Type**. Here **Type** refers to the general type of device, e.g. *microscope* or *camera*, while the base annotation covers the specific name, if provided.

Example 9: Samples were analysed by direct infusion on a QExactive mass spectrometer (Thermo Scientific) equipped with a TriVersa NanoMate ion source (Advion Biosciences). (from Salatzki et al. [33])
The terms "QExactive" and "TriVersa NanoMate" are considered as devices, with the respective types "mass spectrometer" and "ion source". The manufacturers are given for both.

The detail annotation tags of **Materials** follow the same principle as devices with the extension of the **Concentration** label. This covers all forms of material concentrations, e.g. *dilution factors*.

Example 10: Briefly, cells were starved for 24h and afterwards stimulated for 6 h with a mix of C16:0, C18:1 and C18:2, dissolved in 10% FFAs-free BSA. (from Salatzki et al. [33])
From the example, the terms "C16:0", "C18:1", "C18:2" and "FFAs-free BSA" are considered materials. For the last the concentration "10%" is provided.

IV. Preliminary Results

In order to assess the potential quality of the proposed annotation scheme, two annotators were asked to provide annotations of research artefacts for scientific articles. The annotators were selected from a pool of annotators according to their achieved annotation quality with respect to software on a sample article from the corpus of [11]. The agreement was determined by comparison with the original annotation provided in the corpus.

TABLE I
PROGRESSION OF COHEN'S κ FOR RESEARCH ARTEFACTS AND MENTION TYPE FOR EACH ITERATION. THE NUMBER N IN BRACKETS PROVIDES THE ACTUAL NUMBER OF ANNOTATIONS.

Iteration	κ Artefact	(n)	κ Mention Type	(n)
init	-0.01	(179)	0.00	(16)
1.	0.19	(436)	0.19	(98)
2.	0.62	(419)	0.15	(280)
3.	0.76	(426)	0.21	(336)
4.	0.96	(415)	0.80	(373)

The annotation corpus was randomly sampled from all papers available from PLoS ONE from January 2013 to June 2018, equally distributed over the years. Sampling and downloading of the full-texts was done using the *rplos* package [3]. The annotation was done with the anafora annotation tool [5], as illustrated in Figure 1.

The preliminary results presented here show the process of training the annotators on the first randomly selected paper of the corpus. The annotators were trained in iterations, where after each annotation round problems and annotation discrepancies were discussed in meetings and the annotators revised their annotations afterwards.

We calculated the inter-rater agreement using Cohen's Kappa. The agreement was evaluated on the type of research artefact as well as the mentioning type. The results are given in Table I. Starting with an initial explanation of the annotation task (*init*), the agreement was determined iteratively. The agreement for the type of mention type was only calculated on overlapping base annotation.

While these results have only preliminary character and cannot be generalised, they suggest that almost perfect agreement for the type of artefact and substantial agreement on the type mention can be reached. The results show a steady increase in agreement on artefact annotation. The stagnation for the agreement in mention type is most likely due to the focus on artefact annotation in the first discussions.

V. Conclusion

In this work, we presented a novel annotation scheme to provide a labelled dataset for the purpose of named entity recognition of research artefacts from scientific publication. We distinguished six classes of artefacts and showed that a general interest in the extraction of these types exist in the literature. Additionally, we discussed potential purposes of extracting knowledge about these research artefacts and provided evidence for this interest from the literature. We discussed how additional information about the research artefacts can be provided by use of additional sub labels for the respective artefact classes.

Finally, we presented first results that suggest that the annotation scheme allows to achieve a level of inter-rater reliability, which is of high importance when it comes to automated methods for artefact identification. In the future, it is planned to employ the annotation scheme on the entire

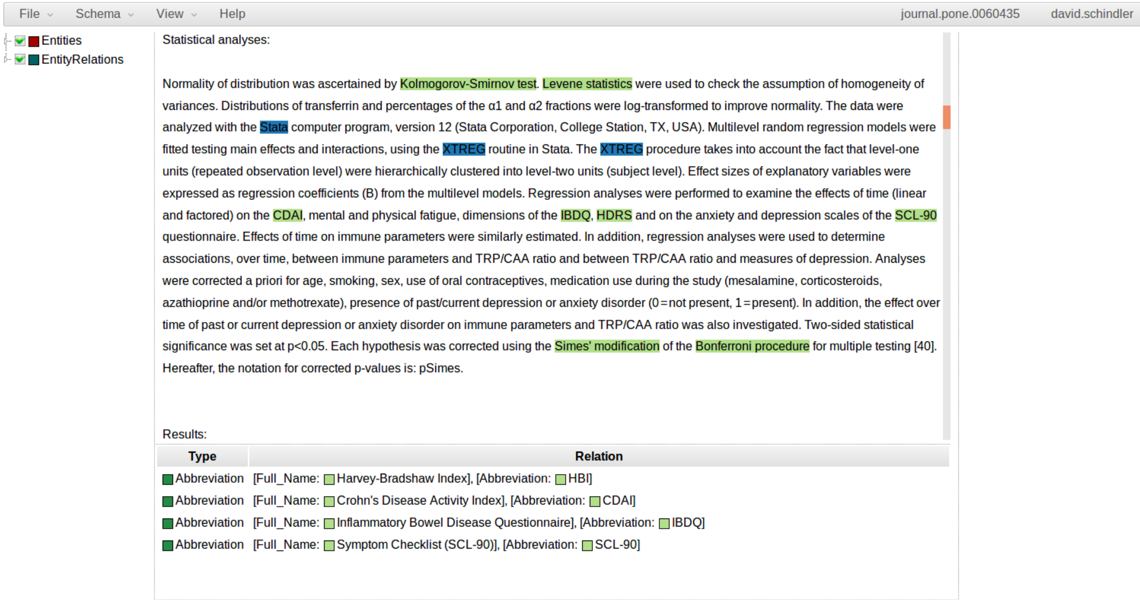

Fig. 1. Screenshot from the Anafora annotation environment.

set of articles in order to provide a dataset to train, evaluate and compare methods of named entity extraction with respect to research artefacts.

Acknowledgement

We thank the students that helped us with the annotation.

References

[1] I. Augenstein, M. Das, S. Riedel, L. Vikraman, and A. McCallum. Semeval 2017 task 10: Scienceie-extracting keyphrases and relations from scientific publications. *arXiv preprint arXiv:1704.02853*, 2017.

[2] K. Boland, D. Ritze, K. Eckert, and B. Mathiak. Identifying references to datasets in publications. In *International Conference on Theory and Practice of Digital Libraries*, pages 150–161. Springer, 2012.

[3] S. Chamberlain, C. Boettiger, and K. Ram. *rplos: Interface to the Search 'API' for 'PLoS' Journals*, 2017. URL `https://CRAN.R-project.org/package=rplos`. R package version 0.8.0.

[4] D. S. Chawla. The unsung heroes of scientific software. *Nature*, 529(7584):115–116, jan 2016. doi: 10.1038/529115a.

[5] W.-T. Chen and W. Styler. Anafora: A web-based general purpose annotation tool. In *Proceedings of the 2013 NAACL HLT Demonstration Session*, pages 14–19. Association for Computational Linguistics, 2013. URL `http://www.aclweb.org/anthology/N13-3004`.

[6] S. Chiesa, M. García-Remesal, G. de la Calle, D. de la Iglesia, V. Bankauskaite, and V. Maojo. Building an index of nanomedical resources: An automatic approach based on text mining. In *International Conference on Knowledge-Based and Intelligent Information and Engineering Systems*, pages 50–57. Springer, 2008.

[7] Data Citation Synthesis Group. Joint declaration of data citation principles. 2014. doi: 10.25490/a97f-egyk.

[8] B. De Bruijn, S. Carini, S. Kiritchenko, J. Martin, and I. Sim. Automated information extraction of key trial design elements from clinical trial publications. In *AMIA Annual Symposium Proceedings*, volume 2008, page 141. American Medical Informatics Association, 2008.

[9] D. Demner-Fushman and J. G. Mork. Extracting characteristics of the study subjects from full-text articles. In *AMIA Annual Symposium Proceedings*, volume 2015, page 484. American Medical Informatics Association, 2015.

[10] A. Devaraju, J. Klump, S. J. Cox, and P. Golodoniuc. Representing and publishing physical sample descriptions. *Computers & Geosciences*, 96:1–10, nov 2016. doi: 10.1016/j.cageo.2016.07.018.

[11] G. Duck, A. Kovacevic, D. L. Robertson, R. Stevens, and G. Nenadic. Ambiguity and variability of database and software names in bioinformatics. *Journal of biomedical semantics*, 6(1):29, 2015.

[12] A. Eklund, T. E. Nichols, and H. Knutsson. Cluster failure: why fmri inferences for spatial extent have inflated false-positive rates. *Proceedings of the National Academy of Sciences*, page 201602413, 2016.

[13] R. Frigg and S. Hartmann. Models in science. In E. N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, summer 2018 edition, 2018.

[14] S. Guloksuz, M. Wichers, G. Kenis, M. G. Russel, A. Wauters, R. Verkerk, B. Arts, and J. van Os. Depressive symptoms in crohn's disease: relationship with immune activation and tryptophan availability. *PLoS One*, 8(3):e60435, 2013.

[15] J. Howison and J. Bullard. Software in the scientific literature: Problems with seeing, finding, and using software mentioned in the biology literature. *Journal of the Association for Information Science and Technology*, 67 (9):2137–2155, 2016.

[16] W. Hsu, W. Speier, and R. K. Taira. Automated extraction of reported statistical analyses: towards a logical representation of clinical trial literature. In *AMIA Annual Symposium Proceedings*, volume 2012, page 350. American Medical Informatics Association, 2012.

[17] S. Kosuri, J. R. Kelly, and D. Endy. TABASCO: A single molecule, base-pair resolved gene expression simulator. *BMC Bioinformatics*, 8(1):480, 2007. doi: 10.1186/1471-2105-8-480.

[18] M. Krallinger, F. Leitner, O. Rabal, M. Vazquez, J. Oyarzabal, and A. Valencia. Chemdner: The drugs and chemical names extraction challenge. *Journal of cheminformatics*, 7(1):S1, 2015.

[19] C. Li, M. Donizelli, N. Rodriguez, H. Dharuri, L. Endler, V. Chelliah, L. Li, E. He, A. Henry, M. I. Stefan, J. L. Snoep, M. Hucka, N. L. Novère, and C. Laibe. BioModels database: An enhanced, curated and annotated resource for published quantitative kinetic models. *BMC Systems Biology*, 4(1):92, 2010. doi: 10.1186/1752-0509-4-92.

[20] K. Li, E. Yan, and Y. Feng. How is r cited in research outputs? structure, impacts, and citation standard. *Journal of Informetrics*, 11(4):989–1002, 2017.

[21] K.-W. Liang, Q. Wang, C. Telmer, D. Ravichandran, P. Spirtes, and N. Miskov-Zivanov. Methods to expand cell signaling models using automated reading and model checking. In *Computational Methods in Systems Biology*, pages 145–159. Springer International Publishing, 2017. doi: 10.1007/978-3-319-67471-1_9.

[22] Y. Liu, L. Nie, L. Liu, and D. S. Rosenblum. From action to activity: sensor-based activity recognition. *Neurocomputing*, 181:108–115, 2016.

[23] M. Lu, S. Bangalore, G. Cormode, M. Hadjieleftheriou, and D. Srivastava. A dataset search engine for the research document corpus. In *Data Engineering (ICDE), 2012 IEEE 28th International Conference on*, pages 1237–1240. IEEE, 2012.

[24] M. S. Mayernik, D. L. Hart, K. E. Maull, and N. M. Weber. Assessing and tracing the outcomes and impact of research infrastructures. *Journal of the Association for Information Science and Technology*, 68(6):1341–1359, 2017.

[25] H. Mooney and M. P. Newton. The anatomy of a data citation: Discovery, reuse, and credit. *Journal of Librarianship & Scholarly Communication*, 1(1), 2012.

[26] A. Névéol, W. J. Wilbur, and Z. Lu. Extraction of data deposition statements from the literature: a method for automatically tracking research results. *Bioinformatics*, 27(23):3306–3312, 2011.

[27] X. Pan, E. Yan, Q. Wang, and W. Hua. Assessing the impact of software on science: A bootstrapped learning of software entities in full-text papers. *Journal of Informetrics*, 9(4):860–871, 2015.

[28] X. Pan, E. Yan, and W. Hua. Disciplinary differences of software use and impact in scientific literature. *Scientometrics*, 109(3):1593–1610, 2016.

[29] H. Park and D. Wolfram. An examination of research data sharing and re-use: implications for data citation practice. *Scientometrics*, 111(1):443–461, 2017.

[30] T. C. Rindflesch, L. Tanabe, J. N. Weinstein, and L. Hunter. Edgar: extraction of drugs, genes and relations from the biomedical literature. In *Biocomputing 2000*, pages 517–528. World Scientific, 1999.

[31] A. Roberts, R. Gaizauskas, M. Hepple, G. Demetriou, Y. Guo, A. Setzer, and I. Roberts. Semantic annotation of clinical text: The clef corpus. In *Proceedings of the LREC 2008 workshop on building and evaluating resources for biomedical text mining*, pages 19–26, 2008.

[32] C. H. Roney, J. D. Bayer, H. Cochet, M. Meo, R. Dubois, P. Jaïs, and E. J. Vigmond. Variability in pulmonary vein electrophysiology and fibrosis determines arrhythmia susceptibility and dynamics. *PLOS Computational Biology*, 14(5):e1006166, 2018.

[33] J. Salatzki, A. Foryst-Ludwig, K. Bentele, A. Blumrich, E. Smeir, Z. Ban, S. Brix, J. Grune, N. Beyhoff, R. Klopfleisch, et al. Adipose tissue atgl modifies the cardiac lipidome in pressure-overload-induced left ventricular failure. *PLoS genetics*, 14(1):e1007171, 2018.

[34] A. M. Smith, D. S. Katz, and K. E. N. and. Software citation principles. *PeerJ Computer Science*, 2:e86, sep 2016. doi: 10.7717/peerj-cs.86.

[35] J. Valdez, M. Kim, M. Rueschman, S. Redline, and S. S. Sahoo. Classification of provenance triples for scientific reproducibility: A comparative evaluation of deep learning models in the ProvCaRe project. In *Lecture Notes in Computer Science*, pages 30–41. Springer International Publishing, 2018. doi: 10.1007/978-3-319-98379-0_3.

[36] M. Vazquez, M. Krallinger, F. Leitner, and A. Valencia. Text mining for drugs and chemical compounds: methods, tools and applications. *Molecular Informatics*, 30 (6-7):506–519, 2011.

[37] M. D. Wilkinson, M. Dumontier, I. 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. Gray, P. Groth, C. Goble, J. S. Grethe, J. Heringa, P. A. '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. *Scientific Data*, 3:160018, mar 2016. doi: 10.1038/sdata.2016.18.