Linkflows: Towards Genuine Semantic Publishing in Science

Bucur, Cristina-Iulia

2024

DOI (link to publisher)
10.5463/thesis.592

document version
Publisher's PDF, also known as Version of record

Scientific publishing is the main way in which novel research ideas, evidence, data, and scientific results are communicated, shared and assessed. Currently this process of scientific publishing lacks, most of the time, transparency and machine-interpretable representations. As such, scientific publication is still done in scientific articles, basically long coarse-grained text with complicated structures in natural language that are optimized for human readers and not for automated systems. Moreover, peer reviewing continues to be the main method of quality assessment in science, despite the many and serious flaws emphasized by the scientific community like the lack of transparency, accuracy and efficiency of such a practice. And, as science is rapidly changing and moving more towards a digital environment, with scientific contributions increasing in volume and complexity each day, scientific publishing that still follows this old paradigm of publishing seems to be at odds with the current scientific progress.

Altogether, our research shows that it is possible to bring scientific publishing closer to automated systems and make the elements of publications and their assessments, along with their corresponding processes more machine-interpretable, especially if semantics are captured from the start. The complementary models proposed can be used in the publishing practice and can be adequately used in tools that support the publishing activities and workflows. And, regardless of the challenges encountered along these processes, after reasonable guidance, users are able to use these tools and models with ease despite their novelty, demonstrating that overall our ideas work. This can be a first step in the direction of involving machines more in the research community, where a new way of publishing that considers the Semantic Web principles can help in creating a society where computers can assist the Open Science research community become more efficient and make scientific contributions FAIR: easier to find, more accessible, interoperable and reusable.

Cristina-Iulia Bucur has a background in Computer Science with a specialisation in High Performance Computing. Prior to this PhD project, she worked as a researcher in projects with Centrum Wiskunde & Informatica (CWI), the National Research Institute for Mathematics and Computer Science in the Netherlands, VU University Amsterdam, the Rijksmuseum, the British Museum, the Naturalis Biodiversity Centre and the Netherlands Institute for Sound and Vision.

9 789464 697933

Towards Genuine Semantic Publishing in Science

C.I.Bucur

Cristina-Iulia Bucur

Linkflows: Towards Genuine Semantic Publishing in Science

DOCTORAL THESIS

Linkflows: Towards Genuine Semantic Publishing in Science

Cristina-Iulia Bucur

SIKS Dissertation Series No. 2024 - 10

The research reported in this thesis has been carried out under the auspices of SIKS, the Dutch Research School for Information and Knowledge Systems.

The research reported in this thesis was partially funded by IOS Press.

The research reported in this thesis was also partially funded by The Netherlands Institute for Sound and Vision.

DOI 10.5463/thesis.592
ISBN 978-94-6469-793-3

Illustrations by Melisa Machuret || melisamachuret.com
Black and white sketches from www.shutterstock.com
Printed by ProefschriftMaken || www.proefschriftmaken.nl

VRIJE UNIVERSITEIT

LINKFLOWS: TOWARDS GENUINE SEMANTIC PUBLISHING IN SCIENCE

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad Doctor aan
de Vrije Universiteit Amsterdam,
op gezag van de rector magnificus
prof.dr. J.J.G. Geurts,
in het openbaar te verdedigen
ten overstaan van de promotiecommissie
van de Faculteit der Bètawetenschappen
op woensdag 6 maart 2024 om 11:45 uur
in een bijeenkomst van de universiteit,
De Boelelaan 1105

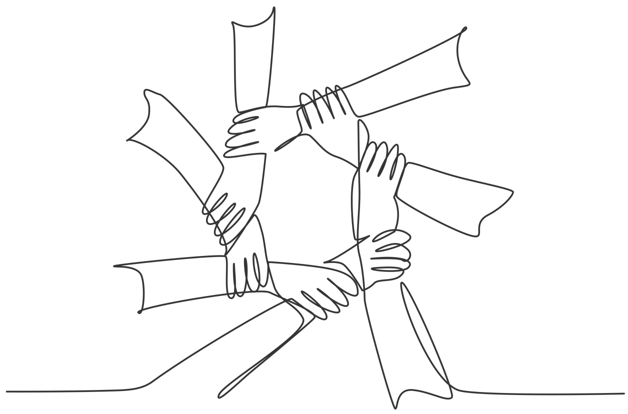
door

Cristina-Iulia Bucur

geboren te Boekarest, Roemenië

promotor:	prof.dr. J.R. van Ossenbruggen
copromotoren:	dr. T. Kuhn dr. D. Ceolin
promotiecommissie:	prof.dr. J. Domingue prof.dr. O. Corcho prof.dr. M. Dumontier prof.dr. A.C.M. ten Teije dr. I. Tiddi

To my family. Without your support throughout the years, this thesis would not have been possible.



ABSTRACT

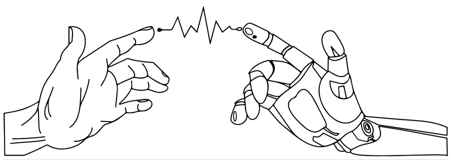
Scientific publishing is the main way in which novel research ideas, evidence, data, and scientific results are communicated, shared and assessed. Currently this process of scientific publishing lacks, most of the time, transparency and machine-interpretable representations. As such, scientific publication is still done in scientific articles, basically long coarse-grained text with complicated structures in natural language that are optimized for human readers and not for automated systems. Moreover, peer reviewing continues to be the main method of quality assessment in science, despite the many and serious flaws emphasized by the scientific community like the lack of transparency, accuracy and efficiency of such a practice. And, as science is rapidly changing and moving more towards a digital environment, with scientific contributions increasing in volume and complexity each day, scientific publishing that still follows this old paradigm of publishing seems to be at odds with the current scientific progress.

We address some of these issues of scientific publishing that make it hard for researchers in the Digital Era to disseminate, share and assess their work in an automated fashion and we focus on making the publication process more organized, structured and transparent. Mainly, we investigate new approaches in the digital environment of scientific publishing by focusing on making scientific articles machine-readable from the start by using existing technologies or newly developed approaches from the Semantic Web. Thus, we explore how we can further use machines to publish, aggregate, summarize, and automate more of the scientific publishing process.

In this thesis, we propose and evaluate models that tackle how scientific articles and their findings together with their reviews are written and published and how these models can be used in concrete applications that can assist the publishing process for humans and machines alike. First, we introduce a novel fine-grained semantically rich model for reviews in which reviews consist of review comments that contain formal links and semantics from the start and as such are able to make the reviewing process better organized and more accurate. Next, we use this fine-grained reviewing model in conjunction with semantic representations of all the elements of publications and their structure together with all the involved processes, actors and provenance to represent in a unified semantic model the scientific publishing process. Then, we focus on finding a semantic pattern that is able to formally represent the content of high-level scientific findings in a way that can be automatically interpreted. Afterwards, we test these models and applications on a multitude of scientific publication

roles such as reviewers, editors, authors and eventually readers in a field study where we prove that not only the scientific publication process (with submissions and final articles) can be represented with formal semantics from the start, but also the whole process in between, including reviews, responses and decisions.

Altogether, our research shows that it is possible to bring scientific publishing closer to automated systems and make the elements of publications and their assessments, along with their corresponding processes more machine-interpretable, especially if semantics are captured from the start. The complementary models proposed can be used in the publishing practice and can be adequately used in tools that support the publishing activities and workflows. And, regardless of the challenges encountered along these processes, after reasonable guidance, users are able to use these tools and models with ease despite their novelty, demonstrating that overall our ideas work. This can be a first step in the direction of involving machines more in the research community, where a new way of publishing that considers the Semantic Web principles can help in creating a society where computers can assist the Open Science research community become more efficient and make scientific contributions FAIR: easier to find, more accessible, interoperable and reusable.



PUBLICATIONS

The publications on which the chapters in this thesis are based on are:

[1] Cristina-Iulia Bucur. "Linkflows: Enabling a Web of Linked Semantic Publishing Workflows." In: *The Proceedings of the 15th Extended Semantic Web Conference Satellite Events*. ESWC'18. 2018, pp. 262–271. DOI: 10.1007/978-3-319-98192-5_45.

[2] Cristina-Iulia Bucur, Tobias Kuhn, and Davide Ceolin. "Peer Reviewing Revisited: Assessing Research with Interlinked Semantic Comments." In: *Proceedings of the 10th International Conference on Knowledge Capture*. K-CAP'19. 2019, pp. 179–187. DOI: 3360901.3364434.

[3] Cristina-Iulia Bucur, Tobias Kuhn, Davide Ceolin, and Jacco van Ossenbruggen. "A Unified Nanopublication Model for Effective and User-Friendly Access to the Elements of Scientific Publishing." In: *Proceedings of the 22nd International Conference on Knowledge Engineering and Knowledge Management*. EKAW'20. 2020, pp. 104–119. DOI: 10.1007/978-3-030-61244-3_7.

[4] Cristina-Iulia Bucur, Tobias Kuhn, Davide Ceolin, and Jacco van Ossenbruggen. "Expressing High-Level Scientific Claims with Formal Semantics." In: *Proceedings of the 11th International Conference on Knowledge Capture*. K-CAP'21. 2021, pp. 233–240. DOI: 10.1145/3460210.3493561.

[5] Cristina-Iulia Bucur, Tobias Kuhn, Davide Ceolin, and Jacco van Ossenbruggen. "Nanopublication-Based Semantic Publishing and Reviewing: A Field Study with Formalization Papers." In: *PeerJ Computer Science* 9 (2023). DOI: 10.7717/peerj-cs.1159.

The art of writing a beautiful fugue lies precisely in this ability,
to manufacture several different lines, each one of which gives
the illusion of having been written for its own beauty, and yet which
when taken together form a whole, which does not feel forced in any way.

— Douglas Hofstadter, *Gödel, Escher, Bach: an Eternal Golden Braid*

ACKNOWLEDGMENTS

The journey of completing a PhD is not possible without the support, effort and help of many people. My supervisors, PhD committee members, funding partners, group members, former collaborators, my friends and family have all played a role in this research endeavour.

Above all, I would like to express my sincere gratitude and thanks to my supervisors Jacco van Ossenbruggen, Davide Ceolin and Tobias Kuhn. Jacco's overall vision of the project, his attention to detail and his support both in terms of research and on a personal level were amazing. Davide, thank you for believing in me and helping me to overcome my low points. The PhD journey has indeed been an ultramarathon! Finally, I would like to express my deepest gratitude and many thanks to Tobias. You have been a constant help throughout the years we have worked together and an incredible role model. I am confident that I would not have reached this stage without your constant support and I am grateful to you for my development. I am still amazed by your efficiency not only in academic writing but also in your technical skills. In addition, your ability to plan and execute these plans is nothing short of amazing. Thank you for pushing me out of my comfort zone and ensuring that my perfectionism did not get in the way when it came to conducting or writing about our research.

I would also like to thank the members of my doctoral committee, John Domingue, Oscar Corcho, Michel Dumontier, Annette ten Teije and Ilaria Tiddi, for taking the time to read, evaluate and provide useful comments on the manuscript of this thesis. I am honoured to have the opportunity to defend my thesis before you.

In addition to VU Amsterdam, this research was funded by two other institutions: the publishing house IOS Press and the Netherlands Institute for Sound and Vision. I am therefore grateful for the constant support of IOS Press, represented by Stephanie Delbeque and Maarten Fröhlich. Stephanie, thank you for your enthusiasm and energy during our regular meetings, and Maarten for your calm and inquisitive manner. Both of you were always willing to try out new ideas and put research into practice. Erwin Verbruggen and Johan Oomen from the Netherlands Institute for Sound and Vision provided their insight and expertise throughout the project.

I am very grateful to all current and former colleagues in the User Centric Data Science group at VU Amsterdam: Xander, Chris, Oana,

Anca, Victor, Archana, Xueli, Roderick, Shuai, Dayana, Dena, Niels, Valentina, Ronald and to the newcomers Emma, Elena, Margherita, Gu, Sarah and Mirthe. And of course to Jacco, Tobias and Davide. Thank you for giving me the opportunity to be part of a research community where I could share ideas and have fruitful discussions. The feedback you gave me over the years helped shape my academic career and made my work feel like part of something bigger. Our secretary, Mojca, was always there for all of us, ready to help whenever anything was needed, always with a smile in her pocket. We are very lucky to have you!

Xander, I have a deep admiration for you as a colleague and friend, and I am deeply grateful for all the insightful discussions we have had. Your technical and research skills, coupled with your determination to get things done, make you an incredible asset to any research group. Thank you for all your words of encouragement and for believing in me. Chris, thanks to you I have started to think more seriously about a career in research. Some extended thanks to you below. Oana and Anca, thank you for the enlightening discussions about *CrowdTruth* and the tips on teaching. Whenever I wanted to feel at home or needed help with something, I knew I could always come to your office. Victor, I have always loved how clearly you can explain almost anything and how you can delve into the technical details of various research projects. Archana, thank you for making me smile more times than I can count. Xueli, thank you for the many times we had tea together, always accompanied by interesting discussions. Roderick, for a short time we had opposite desks in the new building and I wish we had more time to discuss about *IoT*. Shuai, it was great to read some of your stories and share some of mine with you. Dayana, how you managed to supervise so many students at the same time still amazes me. Dena, I have yet to hear you play the piano and I miss our discussions about literature. Niels, thank you for introducing me to *Data2Documents* and the discussions around it. Valentina, thank you for explaining serendipity to me and for taking the time to talk about mental health and its importance in the long run. Ronald, it was nice to hear your updates on the interesting projects you were working on and to talk with you about (your) cats.

Prior to this PhD, it was a real pleasure to work with Lora Aroyo, Chris Dijkshoorn and Leon Wessels on research projects with the Rijksmuseum, the British Museum, the Naturalis Biodiversity Centre and the Netherlands Institute for Sound and Vision. Our collaboration is one of the reasons why I ended up doing a PhD. Leon, thank you for our many (not only) historical discussions and the many good times we had talking about SPARQL over lunch or coffee and during our research trip to London. Chris, you lit the spark that convinced me to pursue an academic career in the form of a PhD. Your determination, hard work, professionalism and planning skills made every project we worked on a success. I fondly remember the days we worked in the "tower" of the Rijksmuseum, programming or just debugging together, while wandering around the museum at lunchtime like simple tourists. Lora is a force of nature and how she managed to be involved in so many projects at the same time is still a mystery to me.

To my friends Flori, Roxana and Răzvan, Nadi and Rosh, Vladimir and Denisa, Vlad, Coco, Georgi, Iri, Luci and Anca, and last but by no means least Adelina, Gianni, Jackie, Andreea and Gențiana. Thank you for making Amsterdam feel like home! In the end, I would like to give my heartfelt thanks to my extended and close family. I am deeply grateful for your encouragement and support over the years. *Gras, cu Ana formăm cea mai "tare" echipă din lume! Îți mulțumesc pentru tot!*

Thank You

CONTENTS

1 INTRODUCTION 1
1.1 Semantic Web technologies 3
1.2 Genuine Semantic Publishing 5
1.3 Nanopublications 6
1.4 The Linkflows project and this thesis 7
1.5 Research Questions and Contributions 8
1.6 Approach and Structure of this Thesis 11
2 PEER REVIEWING WITH INTERLINKED SEMANTIC COMMENTS 15
2.1 Introduction 17
2.2 Related work 19
2.3 Approach 21
2.3.1 General Approach 21
2.3.2 Linkflows Model of Reviewing 21
2.4 Evaluation Design 24
2.4.1 Dataset 24
2.4.2 Ground-Truth and Manual Annotation 26
2.4.3 Peer Questionnaire 27
2.4.4 Automated Sentiment Analysis 27
2.5 Results 28
2.5.1 Descriptive Analysis 28
2.5.2 Accuracy of Sentiment Analysis Methods 30
2.5.3 Comparative Analysis 31
2.6 Discussion and Conclusion 33
3 A UNIFIED NANOPUBLICATION MODEL FOR THE ELEMENTS OF SCIENTIFIC PUBLISHING 35
3.1 Introduction 37
3.2 Background 39
3.3 Approach 41
3.3.1 Semantic Model and Nanopublications 41
3.3.2 Use Case with Competency Questions 42
3.3.3 Dataset 44
3.3.4 Interface Prototype for Use Case 44
3.4 Evaluation 46
3.4.1 Evaluation Design 46
3.4.2 Evaluation Results 48
3.5 Discussion and Conclusion 51
4 EXPRESSING HIGH-LEVEL SCIENTIFIC CLAIMS WITH FORMAL SEMANTICS 53
4.1 Introduction 55
4.2 Background 57
4.3 Approach and Methods 59

4.3.1 Super-Pattern 59
4.3.2 Dataset A 63
4.3.3 Dataset B 64
4.4 Evaluation 65
4.4.1 Descriptive Analysis 65
4.4.2 Design of Formalization Study 67
4.4.3 Formalization Study Results 68
4.5 Discussion and conclusion 70

5 NANOPUBLICATION-BASED SEMANTIC PUBLISHING AND REVIEWING 73
5.1 Introduction 75
5.2 Background 79
5.2.1 Scientific knowledge representation 79
5.2.2 Semantic publishing and semantic papers . . . 80
5.2.3 Nanopublications 81
5.2.4 Genuine semantic publishing 81
5.3 Methodology 83
5.3.1 Approach 83
5.3.2 Formalization Papers 84
5.3.3 Tools 85
5.3.4 Field study design 85
5.3.5 User Feedback 90
5.4 Results 91
5.4.1 Analysis of Formalizations 92
5.4.2 Analysis of Nanopublications 93
5.4.3 User Feedback Analysis 96
5.5 Discussion and Conclusion 99

6 CONCLUSIONS 101
6.1 Research Questions Revisited 101
6.2 Discussion 103
6.3 Outlook 108

A APPENDIX 111

BIBLIOGRAPHY 115
SUMMARY 133
SAMENVATTING 137
REZUMAT 143
SIKS DISSERTATION SERIES 147

LIST OF FIGURES

Figure 1	Linkflows model for reviewing	22
Figure 2	Mock interface implementing the Linkflows model for reviewing	24
Figure 3	The part of the article that a review comment targets	28
Figure 4	The results of the model expert annotations	29
Figure 5	Accuracy of sentiment analysis methods	30
Figure 6	An example of a nanopublication-style communication interaction	41
Figure 7	Example nanopublication of a review comment	45
Figure 8	The reviewer-oriented view for the editor study	47
Figure 9	The section-oriented view for the editor study	48
Figure 10	Diagram with an example for the Super-pattern instantiation	60
Figure 11	Agreement between participants on the formalizations before and after discussion	70
Figure 12	Formalization paper template from Nanobench as used by the participants of our study	84
Figure 13	The Tapas interface listing submitted formalizations as results of SPARQL queries over the nanopublication service network	86
Figure 14	A nanopublication view of a formalization paper	89
Figure 15	The human-readable view of a formalization paper	91
Figure 16	Graphical representation of all the submissions to the Data Science special issue with formalization papers	95
Figure 17	Questionnaire Part 1: Average answers from participants on conceptual aspects of formalization papers	97
Figure 18	Questionnaire Part 2: Average answers from participants on technical aspects of formalization papers	97
Figure 19	Questionnaire Parts 3 and 4: Average answers from participants on general aspects about formalization papers and average rating of participants with respect to their background knowledge on various topics	98
Figure 20	Timeline publication for the special issue with formalization papers at the Data Science journal	113

LIST OF TABLES

Table 1 Overall approach of the thesis 12
Table 2 Percentages of questions where it was not possible to answer (Peers experiment) 29
Table 3 Disagreement scores for four of the Linkflows model dimensions 31
Table 4 Disagreement scores for the positivity/negativity dimension 31
Table 5 Model experts versus peers 33
Table 6 Descriptive statistics dataset 49
Table 7 Statistics nanopublications 49
Table 8 Results of the user study with editors 51
Table 9 Super-pattern qualifiers 61
Table 10 Usage of qualifiers in the dataset 65
Table 11 Usage of relations in the dataset 66
Table 12 Levels of agreement in the formalization study 68
Table 13 Instantiated super-patterns accepted for publication in formalization papers in the Data Science special issue 94
Table 14 Nanopublications created during the field study of the special issue with formalization papers at the Data Science journal 96

ACRONYMS

AI	Artificial Intelligence
BEL	Biological Expression Language
CISE	Compositional and Iterative Semantic Enhancement
CQ	Competency Question
daQ	Dataset Quality Information
DOI	Digital Object Identifier
DSKG	Data Set Knowledge Graph
FAIR	Findable, Accessible, Interoperable, Reusable
IPFS	InterPlanetary File System
JIF	Journal Impact Factor
LD	Linked Data
LDN	Linked Data Notifications
LLM	Large Language Model
LOD	Linked Open Data
LODC	Linked Open Data Cloud
LOV	Linked Open Vocabularies
OADM	Open Annotation Data Model
OBKMS	Open Biodiversity Knowledge Management System
ORCA	Ontology of Reasoning, Certainty and Attribution
ORCID	Open Researcher and Contributor IDentifier
ORKG	Open Research Knowledge Graph
OWL	Web Ontology Language
PRISM	Publishing Requirements for Industry Standard Metadata
PROV-O	Provenance Ontology
RASH	Research Articles in Simplified HTML
RDF	Resource Description Framework
RDFa	Resource Description Framework in Attributes
RO	Research Object
SKGO	Science Knowledge Graph Ontologies
SKOS	Simple Knowledge Organization System
SPAR	Semantic Publishing and Referencing Ontologies
SPARQL	SPARQL Protocol and RDF Query Language

URI Uniform Resource Identifier

URL Uniform Resource Locator

WADM Web Annotation Data Model

1 INTRODUCTION

Sometimes it seems as though each new step towards AI,
rather than producing something which everyone agrees is real
intelligence, merely reveals what real intelligence is not.
— Douglas Hofstadter, *Gödel, Escher, Bach: an Eternal Golden Braid*

In the world of science, academic publishing is at the core of science dissemination, providing a way to communicate, share and evaluate new research ideas and discoveries. Currently, the field of scientific publishing seems to face a major crisis [20, 41, 42, 53, 88, 127, 141]. While the paradigm of publishing has stayed basically the same for 300 years, now the increasing volume of articles published every day makes it very hard for scientists to stay up to date in their respective fields. In the recent years, with the pervasiveness of technology and the Internet, we changed not only the way we do science, but also how we perform and disseminate science [13, 122, 131, 137]. As such, scientific publishing became a more versatile and multifaceted process changing completely the initial paradigm of publishing towards a digital environment with new methods of electronic publication including scientific workflows, research protocols and standard operating procedures [10].

Especially in the natural sciences, moving towards a more digital environment and generating digital content seems to be more the rule and challenges the classic ways of publishing [34, 72, 128, 151]. As such, ontologies and controlled vocabularies have widely been used to aid the access to a great number of databases in the fields of biology, bioinformatics, physics, chemistry, medicine and other related areas [8, 9, 89, 160]. The heterogeneity and different semantics used in these various databases can be overcome by the use of structured texts and descriptions provided in a hierarchical concept definition together with domain knowledge. Fundamentally, in life sciences, ontologies used in databases can facilitate data integration, data access and analysis, supporting integrative research and data-driven science, thus providing better ways to access, process, and exploit data, information, and knowledge in a transparent and structured way. We believe that the use of ontologies and controlled vocabularies, together with the idea of scientific assertions that can be published independently in a machine-readable, interoperable, and citable manner [56, 85, 123] can be essential for generating a paradigm shift in the scientific publishing world.

In the digital publishing context, Linked Data (LD) is a key technology for this paradigm shift in the scientific publishing world, especially as it supports scientific publications by enabling the exchange, reuse and linking of data on the Web [16]. While the Linked Data set of best practices to connect and publish structured data on the Web is not sufficient to enable the entire scientific publication process, it is an important layer that facilitates it. The Linked Data principles encourage using dereferenceable HTTP Uniform Resource Identifier (URI)s for things like datasets, services, tools, etc. and including links to other URIs. Linked Data also supports provenance (meta-)information about the resources that are linked, thus giving a way to locate various versions of data and access information like ownership and copyright. In turn, this can support the shift brought forth by digital publishing, changing the sharing of scientific knowledge from distribution to discovery.

Important issues that have been identified with the current publishing practice remain and these include the lack of accuracy, efficiency, and reproducibilty, as well as the "Knowledge Burying" problem. We will discuss in more details these issues in the following paragraphs.

On the one hand, many have pointed out serious flaws of current scientific publishing practices [41, 90, 106, 129], including with respect to the accuracy and efficiency of the reviewing process [2, 44, 134], especially as scientific publishing has evolved from the form of a classical article to electronic publishing of scholarly journals. On the other hand, with the recent advances in Artificial Intelligence (AI), especially tools that use existing scientific databases like Semantic Scholar and Large Language Model (LLM)s, the sharing and discovery of research work and results can be made faster within the research communities, accelerating the scientific process and connecting methods across the various scientific disciplines [158].

All these advances in AI, together with the criticism brought forth by the reviewing process, might need us to reconsider the scientific publication process in the near future. Especially if it is still pertinent to package not only the scientific work in scientific article structures, but also the peer reviews in big bulks of natural language texts that are mostly suited for human cognition. While the advances with regard to LLMs have recently been significant, especially with models like ChatGPT, in this thesis our focus will not be on using such models[1], but on AI techniques used in the field of knowledge representation and reasoning, especially with respect to the scientific publishing process.

Reproducibility plays a crucial role in scientific research because it allows others to test, check and verify the validity of one's claims and methods [101] and it permits further collaboration and reuse of scientific discoveries. Unfortunately, according to a study published in Nature [6], over 70% of the 1500 interrogated scientists admitted

1 However, we used DeepL for the automated translation of this thesis summary from English to Dutch and Romanian.

to have failed to reproduce the work of other researchers at some point in time. The Findable, Accessible, Interoperable, Reusable (FAIR) principles for scientific information [165] can be key factors in guiding towards reproducible research. According to these, data should be (i) findable both for humans and machines; (ii) accessible on the long term; (iii) interoperable by the use of shared vocabularies, for example; and (iv) reusable for both humans and machines. And, following these guidelines should, in turn, support reproducibility. While we do not directly address the issue of scientific reproducibility in this thesis, we touch upon some aspects of scientific publishing that can enable it, like the initial publication in a structured, machine-interpretable way of scientific knowledge that considers the FAIR principles as guiding factors.

As Mons [103] notices, another important issue of traditional scientific articles and another hurdle in the way of accessibility, discovery and reproducibility of research is the process of "Knowledge Burying". This term entails that all information is written and published in one bulk of text - the scientific article - that contains the scientific hypotheses, arguments, methods and results that is adapted solely to human cognition, using compact natural language text that cannot be read by automated systems. So, in order to make the scientific information available to machines, we need to extract knowledge from the original published scientific article and restructure information in a structured form that is machine-readable. For this, additional methods like text mining need to be applied, methods that result in a loss of knowledge as information is adapted and put into various structured forms. In this thesis we propose a few approaches that consider the structured publication of data from the start, methods meant to avoid "Knowledge Burying" altogether.

Some of the topics discussed in this Introduction are based on the following publication:

– *Linkflows: Enabling a Web of Linked Semantic Publishing Workflows* in the Proceedings of the European Semantic Web Conference Satellite Events, PhD Symposium Track, 2018 [21].

1.1 SEMANTIC WEB TECHNOLOGIES

Semantic publishing is an umbrella term that covers a range of activities and technologies around the general idea of making scientific publishing machine-interpretable. This is not a new concept, as its roots are tightly coupled to the notion of the Web, with Tim Berners-Lee mentioning that the semantic web "will likely profoundly change the very nature of how scientific knowledge is produced and shared, in ways that we can now barely imagine" [13]. Despite the fact that semantic publishing is not directly linked to the Web, its progress was highly influenced by the rise of the "semantic web." As such, in the beginning, it referred to mainly publishing information on the Web in the form of documents that additionally contain structured annotations, so extra information that is parsable by machines in the form of semantic markup (with markup languages like Resource Description Framework in Attributes (RDFa), for example). This allowed published information on the Web to be machine-interpretable, to the limited extent to which the markup languages allowed. A next step was to use semantic web languages like Resource Description Framework (RDF) and Web Ontology Language (OWL) to publish information in the form of data objects, together with a specific detailed representation called ontology that is able to represent the domain of the data in a formal (thus machine-interpretable) way. The information published in such structured ways provides not only a "semantic" context through the metadata that describes the information, but also a way for machines to understand the structure and even the meaning of the published information [5, 7, 31, 71].

In this way, semantic publishing would allow for the "automated discovery, enables its linking to semantically related articles, provides access to data within the article in actionable form, or facilitates integration of data between papers" [166]. However, despite all the advancements in the semantic web technologies in the past years, semantic publishing is not "genuine" [82] in the sense that the current scientific publishing paradigm has not changed much as we are still using long articles written in natural language that do not contain formal semantics from the start that machines can process and interpret in an automated manner. So, with scientific publishing often stuck to formats optimized for print such as PDF, we are not using the advances that are available to us with technologies around the semantic web and linked data.

Some researchers consider that semantic publishing is inevitable and that it will happen in incremental steps [140], as it is already possible to publish data as RDF statements in the Linked Open Data Cloud (LODC) [69]. Semantic Web technologies have launched a revolution in the field of scientific publishing and the idea is to create and facilitate an open access ecosystem where both content and metadata of scientific articles is accessible, together with formalized internal structures of the documents and components, enriched and with semantic connections to other related or similar documents.

In the view of the prevalence of the Semantic Web, considerable research was done in enriching the meaning of a traditional article in the digital publishing environment, facilitating its automatic discovery, having access in a semantic way to and within the article and also being able to link to other related articles or other related parts of articles. Especially notable in this sense are the Semantic Publishing and Referencing Ontologies (SPAR) [114], the ontologies central to the task of semantic publishing. All these techniques, methods and approaches can facilitate the scientific publishing domain and our research.

As datasets, documents and, in general, knowledge is spread in the web of the Internet, where everything can be shared and reused and linked, decentralization is a key concept. Decentralization implies that there is no control of a central authority anymore, e.g. a publishing house, over the open content that exists on the Web. In the past, techniques to ensure the functioning of a secure and decentralized global file system over the Internet to entice collaborations have been described [97]. Then, the BitTorrent communication peer-to-peer file sharing protocol over the Internet to distribute and access data in the digital publishing environment was studied [35], while peer-to peer networks for RDF data were developed [49] and a decentralized architecture to support nanopublications, scientific RDF snippets, was built [79]. There is a lot of research in this area of computer science, but we will focus especially on technologies related to the field of digital scientific publishing.

In terms of assessing the quality of scientific publications, the most widely used indicator is the Journal Impact Factor (JIF) [54], but this metric has been the subject of multiple debates in the past as it was shown that it can be favourably manipulated [74]. For example, the JIF can be biased towards journals that publish a high number of non-research items (e.g. research notes, comments) and have higher publishing numbers [42]. So, new ways of rating the quality of scientific publications is needed. Semantic Web technologies with ontologies like the Dataset Quality Information (daQ) [38] can support better and unbiased measures of quality, while new dimensions of quality that consider these technologies need to be taken into account.

However, new initiatives that try to change the old paradigm of publishing are proposed, together with new publishing workflows. This paradigm shift entails to move from the textual representation of information to a more data-centric one, similar to the proposal for the next-generation Web [13], where instead of documents, the interest shifts from the syntactic (e.g. HTML) to the semantic level (e.g. RDF, OWL). On the semantic level, we would be able to express the content and not just the structure of what is now in narrative documents. Moreover, formats that are based on HTML, like Research Articles in Simplified HTML (RASH) [117], have been proposed, where scientific articles that include semantic annotations can be represented.

1.2 GENUINE SEMANTIC PUBLISHING

The majority of current approaches attempting to make scientific texts machine-readable share a commonality: they assume the current paradigm of scientific articles as a given and use it as a starting point to extract information. Hence, in most of the present methods, semantics are solely addressed subsequent to the issuance of scientific articles, while semantic annotations, semantic interlinking, and semantic integration are implemented to further amplify and retrieve information from research that is presently being published in natural language texts. While it is important to try to process the vast amount of existing scientific literature that has the form of long English texts (and sometimes long texts in other languages), we should also think about how we can improve the way how we publish scientific insights in the first place.

An important aspect of this is the vision of semantic publishing, which we mean here in the sense of *genuine semantic publishing* [82], where the machine-interpretable formal semantics cover the main scientific claims the work is making. As such, in the *genuine semantic publishing* vision, semantics are considered prior to, during, and after publication, and are an integral aspect of the publication itself. Additionally, genuine semantic publishing involves formally representing the structure and content of research work from the very beginning while authentic, fine-grained representations are published by authors as primary components of a published entity, without the need for a separate narrative article.

Still, with the exception of a few fields such as biology (particularly biodiversity data with markup-languages such as TaxPub [109] and systems like the Open Biodiversity Knowledge Management System (OBKMS) [108]) and knowledge representation (of scientific knowledge in projects like the Science Knowledge Graph Ontologies (SKGO) [47]), genuine semantic publishing remains a vision that lacks practical evidence on its functionality. As a result, there is still a significant gap in terms of making scientific knowledge machine interpretable, despite the useful attempts and approaches towards this goal. As such, genuine semantic publication requires consideration not only of the semantic representation of a fine-grained and authentic primary component of a publication entity from the very beginning, but also of all aspects related to its publication. However, we have observed in the aforementioned projects that the combination of all these requirements is currently lacking in research.

Nanopublications [58], which are small RDF-based semantic packages, have emerged as a powerful concept and technology for enabling such genuine semantic publishing.

1.3 NANOPUBLICATIONS

Nanopublications [58] are a specific concept and technology based on LD to publish scientific results and their metadata in small publication units. Each nanopublication has an assertion that contains the main content (such as a scientific finding), and comes with provenance about that assertion (e.g. what study was conducted to derive at the assertion; or which documents it was extracted from) and with publication information about the nanopublication as a whole (e.g. by whom and when it was created). All these three parts are represented in RDF and are thereby machine-interpretable.

It has been shown how nanopublications can also be used for other kinds of assertions, including meta-statements about other nanopublications [78], and in order to make nanopublications verifiable and immutable, *trusty URIs* [80, 81] can be used as identifiers, which include cryptographic hash values that are calculated on the nanopublication's content. A decentralized server network has been established based on this, through which anybody can reliably publish and retrieve nanopublications [79]. In order to group nanopublications into larger collections and versions thereof, index nanopublications have been introduced [84]. With these technologies, small interconnected LD snippets can be published in a reliable, decentralized, provenance-aware manner.

In this research, we utilize the concept of nanopublications as a wrapper for the integration of various forms of scientific content. They serve as a container for embedding a wide range of elements, including the primary narrative of scientific articles, the peer reviews, the entirety of the publication process, and the publication of scientific claims or findings. Their usage throughout this thesis illustrates the adaptability and flexibility of nanopublications as a technological framework well-suited for the paradigm described as "genuine semantic publishing" in Section 1.2.

1.4 THE LINKFLOWS PROJECT AND THIS THESIS

The project to which the research of this PhD thesis belongs to is called Linkflows. This name comes from the original research proposal under which funding was granted. The main objective of the Linkflows project was "to make scientific contributions on the Web, e.g. articles, reviews, etc. better valorized and efficiently assessed in a way that allows for their automatic interlinking, quality evaluation and inclusion in scientific workflows". Additionally, an approach mentioned in the research proposal to integrate and assess scientific contributions was to use provenance-aware semantic modeling and publishing.

One of the suggested use cases of the Linkflows project was using semantic publishing in the field of scientific publication as a possible way to "unleash the value of scientific contributions in the Web age". Furthermore, new ways to organize contributions around scientific journals was one of the aims, in a way that these new approaches could be tested in practice with the goal to "experiment with new innovative approaches to scientific publishing that might shape the scientific publishing landscape in the future". Hence, in this thesis titled "Linkflows: Towards Genuine Semantic Publishing in Science", we experiment with using Semantic Web technologies in the field of scientific publication, hoping to get closer to the idea of genuine semantic publishing (as described in Section 1.2) in science.

1.5 RESEARCH QUESTIONS AND CONTRIBUTIONS

The research described in this thesis tries to address some of the problems of scientific publishing presented above by using Semantic Web principles and nanopublications. Therefore, our research is guided by the following main research question:

How can we apply the Semantic Web principles and nanopublications to make genuine semantic publishing possible to scientific findings as well as their assessments?

In this thesis, we address this main research question by answering four sub-research questions, each focusing on a different aspect of the publication process.

> *1. Can an approach for scientific publishing based on a fine-grained semantic model help to make reviewing better structured and more accurate?*

We address this research question in Chapter 2. Research has pointed out serious flaws of current scientific publishing practices, including the lack of accuracy and efficiency of the reviewing process. To address some of these problems, we apply the general principles of the Web and the Semantic Web to scientific publishing, focusing on the reviewing process. We demonstrate that a fine-grained model of the scientific publishing workflow can help us make the reviewing processes better organized and more accurate, by ensuring that review comments are created with formal links and semantics from the start. Our contributions include a novel reviewing model called Linkflows that allows for such detailed and semantically rich representations of reviews and the reviewing processes. We perform the evaluation of this model on a dataset that was manually created and curated from recent open peer reviewed Computer Science journals, creating ground-truth data with the help of the original reviewers. In general, our analysis shows that our model is well understood and easy to apply, and it revealed the semantic properties of such review comments. This chapter is based on the following publication:

– *Peer Reviewing Revisited: Assessing Research with Interlinked Semantic Comments* in the Proceedings of the 10th International Conference on Knowledge Capture, 2019, and was co-authored by Tobias Kuhn and Davide Ceolin [22].

2. Can we use nanopublications as a unifying data model to represent the structure and links of scientific articles and their assessments in a precise, transparent, and provenance-aware manner?

We address this research question in Chapter 3. In this chapter we mainly address the issue of the lack of transparency and machine-interpretable representations in scientific publishing. As scientific articles are published in long coarse-grained text with complicated structures, they are optimized for human readers and not for automated means of organization and access. Moreover, peer reviewing is their main method of quality assessment, but these peer reviews are nowadays rarely published and their own complicated structure and linking to the respective articles are not accessible. In order to address these problems and to better align scientific publishing with the principles of the Web and Linked Data, we propose here an approach to use nanopublications as a unifying model to represent in a semantic way the elements of publications, their assessments, as well as the involved processes, actors, and provenance in general. Our contributions include: 1) a set of seven competency questions that focus on the specific scenario of editors performing a meta-review and we show how they can be executed as SPARQL Protocol and RDF Query Language (SPARQL) queries; 2) a prototype of a user interface for editors in their task of performing a meta-review that is able to answer the specific competency questions proposed. We evaluated our approach on a nanopublications dataset that represented an interlinked network of article elements and review comments, while using this dataset to create interfaces for editors in the task of meta-reviewing. Overall, we demonstrate that a unified and semantic publication model based on nanopublications can make scientific communication more effective and user-friendly. This chapter is based on the following publication:

- *A Unified Nanopublication Model for Effective and User-friendly Access to the Elements of Scientific Publishing* in the Proceedings of the 22nd International Conference on Knowledge Engineering and Knowledge Management, 2020, and was co-authored by Tobias Kuhn, Davide Ceolin and Jacco van Ossenbruggen [23].

3. To what extent can a semantic template called the super-pattern be used to formalize the main claims of scientific articles from different disciplines?

We address this research question in Chapter 4. With the use of semantic technologies gaining significant traction in science communication with a wide array of applications in disciplines including the life sciences, computer science, and the social sciences, languages like RDF, OWL, and other formalisms based on formal logic are applied to make scientific knowledge accessible not only to human readers but also to automated systems. However, these approaches have mostly focused on the structure of scientific publications themselves, on the used scientific methods and equipment, or on the structure of the used datasets, while the core claims or hypotheses of scientific work have only been covered in a shallow manner, such as by linking mentioned entities to established identifiers. In this research, by making use of existing semantic formalisms, we are able to fully express the content of high-level scientific claims using formal semantics in a systematic way. Our contributions include the proposal of a clear semantic pattern which we call the "super-pattern". Our evaluation shows that the instantiation of the five slots of this super-pattern leads to a strictly defined statement in higher-order logic that can be successfully applied to an enlarged sample of scientific claims with a high degree of consistency and convergence given the complexity of the task and the subject. This can allow, on the longer run, for researchers to express their high-level scientific findings in a manner that can be automatically interpreted and this in turn will allow for automated consistency checking, question answering, aggregation, and much more. This chapter is based on the following publication:

– *Expressing High-Level Scientific Claims with Formal Semantics* in the Proceedings of the 11th International Conference on Knowledge Capture, 2021, and was co-authored by Tobias Kuhn, Davide Ceolin and Jacco van Ossenbruggen [24].

4. Are nanopublications and the super-pattern appropriate concepts to enable a new paradigm of scientific communication where authors publish their scientific findings with formal semantics?

We address this research question in Chapter 5. As we have already mentioned, with the amount of scientific literature that continues to increase rapidly, it has become more difficult for researchers in different disciplines to keep up-to-date with the recent findings in their field of study. A possible solution for this problem is the automated processing of scientific articles, but the accuracy of such an approach remains very poor for extraction tasks beyond the most basic ones (like locating and identifying entities and simple classification based on predefined categories). By answering this research question, we want to show that despite the few approaches that have tried to change how we publish scientific results, it is possible to make articles machine-interpretable by expressing them with formal semantics from the start. In this research we demonstrate that we can formally publish high-level scientific claims in formal logic in a real journal setting, hence our contributions include the publishing of scientific claims in a special issue of an existing journal. We use the concept and technology of nanopublications for this endeavor, and represent not just the submissions and final papers in this RDF-based format, but also the whole process in between, including reviews, responses, and decisions. We do this by performing a field study with what we call formalization papers, which contribute a novel formalization of a previously published claim. Our evaluation shows the technical and practical feasibility of our approach as the participating authors mostly showed high levels of interest and confidence, and mostly experienced the process as not very difficult, despite the technical nature of the current user interfaces. This chapter is based on the following publication:

– *Nanopublication-based Semantic Publishing and Reviewing: A Field study with Formalization Papers* in the PeerJ in Computer Science Journal, 2023, and was co-authored by Tobias Kuhn, Davide Ceolin and Jacco van Ossenbruggen [25].

1.6 APPROACH AND STRUCTURE OF THIS THESIS

The purpose of this thesis is to present and provide novel approaches for the scientific publishing field that incorporate Semantic Web principles such that the publication process can be more transparent and machine interpretable. Moreover, we want to investigate how we can make the scientific publishing process more "genuine" in the sense explained in Section 1.2. Therefore, this thesis looks at the different scientific contributions of the publication process like articles, reviews and scientific findings and it makes both theoretical (model) contributions as well as proposals of concrete applications. The two types of contributions (theoretical through the proposal of models and practical through applications of these models) are needed as they complement each other. In addition, as scientific publishing is a process that requires the involvement of different kinds of actors for its functioning, our evaluations focus specifically on the different perspectives of these actors (authors, readers, reviewers, editors) with models and applications tailored for these various roles, in order to give a comprehensive coverage of the scientific publishing process.

It is important to consider the different perspectives and roles of the actors involved in the publishing process of scientific works, as each possesses a significant role during the various stages of the process. At times, individuals can occupy multiple roles, such as authors of articles serving as reviewers for others or even editors for journals. The publishing process commences with an authentic and meaningful contribution, comprising of a submitted manuscript to a journal or a conference by one or more authors. The assessment of this contribution's worth to the scientific community in its respective field is carried out using peer reviews. If the reviewers judge the scientific contribution to be substantial and innovative, editors can assist authors to incorporate feedback and publish their work. This ensures that research results are available to the wider scientific community. All of these roles embrace different aspects of the publication process,

	Chapter 2 **reviewing** MODEL	*Chapter 3* **editing** APPLICATION	*Chapter 4* **reading** MODEL	*Chapter 5* **authoring** APPLICATION
TOPIC	reviewing with semantic comments	unified nanopublication based publishing	scientific claims with formal semantics	field study formalizations
CONTRIBUTIONS	*Section 2.3.2* review ontology	*Section 3.3.1* scheme for interlinked publication elements *Section 3.3.4* editor user interfaces	*Section 4.3.1* SuperPattern ontology *Section 4.3.2 & 4.3.3* datasets with formalized claims	*Section 5.3.4* special issue with "formalization papers"
EVALUATION	*Section 2.4.2 & 2.4.3* user study with reviewers and model experts *Section 2.4.4* automated techniques	*Section 3.4.2.2* competency questions *Section 3.4.2.3* user study with editors	*Section 4.4.2 & 4.4.3* formalization study	*Section 5.4.1 & 5.4.2* field study with "formalization papers" *Section 5.4.3* user feedback questionnaire

Table 1: Overall approach of the thesis. Each chapter focuses on one of the main activities involved in scientific publishing, activities mainly performed by reviewers, editors, readers (both human and machines) and authors. The contributions of the thesis are categorized into models and applications, depending on whether the proposed contributions are theoretical or practical in nature with respect to the topic addressed by each chapter. The evaluations are also listed accordingly per chapter.

each with its specific tasks to fulfil. A comprehensive perspective of scientific publishing can be attained by acknowledging the differing roles and responsibilities of the various actors involved.

Table 1 shows the overall structure of the thesis. As we previously mentioned, we focus both on theoretical contributions (proposing models) and practical contributions (providing applications of the models proposed) focusing not only on the content of the articles and the reviews but also on the findings expressed in these articles. Additionally, the different actors of the publishing process (reviewers, editors, readers and authors) are taken into account in the evaluation of both the theoretical and practical contributions, in order to cover as much of the publication process as possible.

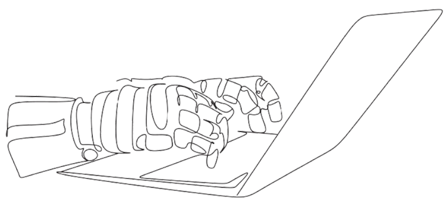
As we can see in Table 1, this thesis contains five chapters. In Chapter 2 we propose a novel fine-grained model for a detailed and semantic representation of peer reviews (that allows for reviews to be represented in small distributed knowledge snippets) and we evaluate this model with the help of reviewers. In Chapter 3 we further test this reviewing model in a concrete application with an evaluation done by editors of scientific journals. In Chapter 4 we propose another model, one that focuses on expressing with formal semantics the findings of scientific claims, using a semantic pattern we called "the super-pattern" which would make these findings accessible to both human readers and automated systems. And lastly, in Chapter 5 we combine all the models and applications proposed earlier and show that it is possible for authors to publish their formally-represented findings in a special issue of a real-life journal where the whole publication workflow, that includes reviews, responses and decisions can also be formally represented.

Throughout this thesis we make full use of the concept of nanopublications, using them as a container for embedding different types of scientific content from the actual narrative of scientific articles, to the reviews and the entire publication process and the publication of scientific claims or findings. This proves that nanopublications are a versatile and flexible technology that can be used in the "new generation of publishing", or "genuine semantic publishing" by all actors involved in the publication process, as we also show in our field study in Chapter 5. Furthermore, we summarize and represent semantically the main scientific claim we address in every chapter using the semantic template called "the super-pattern" that we propose in Chapter 4 and publish them as nanopublications. Each such formalization is mentioned at the beginning of every chapter.

We summarize the main high-level claim that we address in this thesis, together with its published nanopublication as follows:

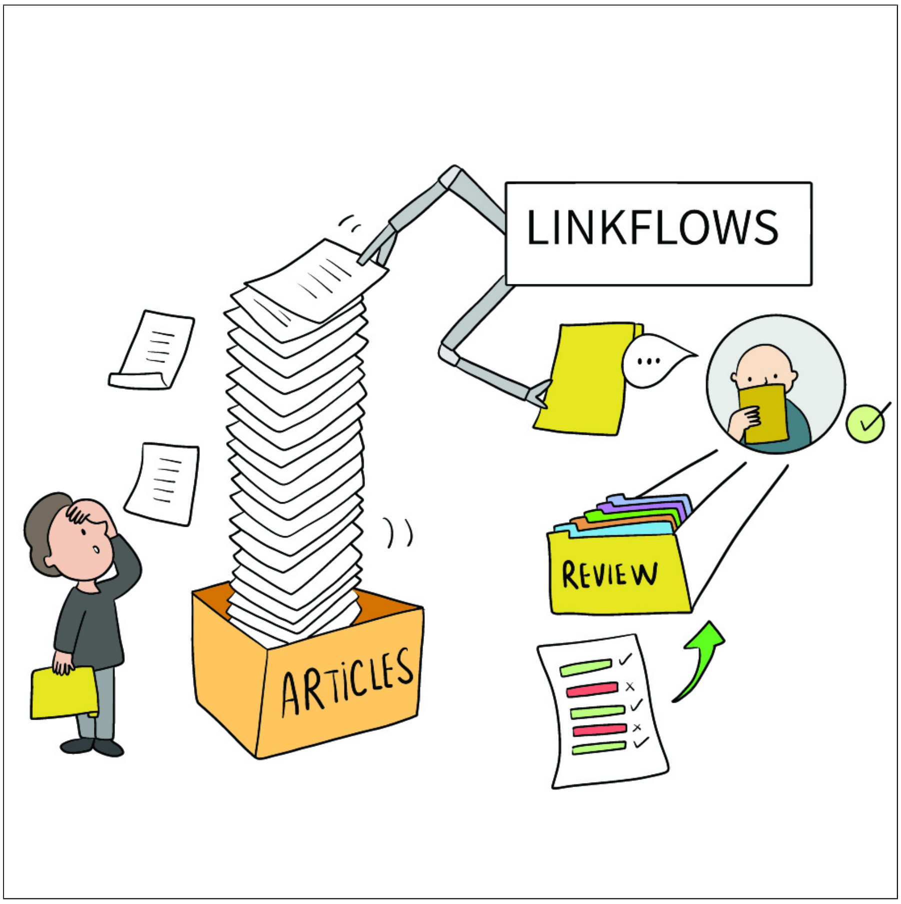
MAIN CLAIM FORMALIZATION: In the context of all things of type *scholarly communication*, things of type *application of Semantic Web principle and nanopublication can generally* have a relation of type *increases* to things of type *efficiency* that are in the same context (i.e. *scholarly communication*).

FORMALIZATION NANOPUBLICATION: (link to nanopublication)[2]



2 https://w3id.org/np/RA63RYoc27kVYbLSnoZAY69_vhm8aMTNmQaxl10GVNKY8

2 PEER REVIEWING WITH INTERLINKED SEMANTIC COMMENTS



MAIN CLAIM FORMALIZATION: In the context of all things of type *peer review*, things of type *application of Linkflows reviewing model generally* have a relation of type *enables* to things of type *fine-grained representation of review comment* that are in the same context (i.e. *peer review*).

FORMALIZATION NANOPUBLICATION: (link to nanopublication)[1]

PUBLICATION: This chapter was previously published as *Peer Reviewing Revisited: Assessing Research with Interlinked Semantic Comments* in the Proceedings of the 10th International Conference on Knowledge Capture, 2019, and was co-authored by Tobias Kuhn and Davide Ceolin [22].

1 https://w3id.org/np/RAwsTvk3vufcj27olpknfjHzlTPlUpCgS4zKPyTXkMAFs

ABSTRACT

Scientific publishing seems to be at a turning point. Its paradigm has stayed basically the same for 300 years but is now challenged by the increasing volume of articles that makes it very hard for scientists to stay up to date in their respective fields. In fact, many have pointed out serious flaws of current scientific publishing practices, including the lack of accuracy and efficiency of the reviewing process. To address some of these problems, we apply here the general principles of the Web and the Semantic Web to scientific publishing, focusing on the reviewing process. We want to determine if a fine-grained model of the scientific publishing workflow can help us make the reviewing processes better organized and more accurate, by ensuring that review comments are created with formal links and semantics from the start. Our contributions include a novel model called Linkflows that allows for such detailed and semantically rich representations of reviews and the reviewing processes. We evaluate our approach on a manually curated dataset from several recent Computer Science journals and conferences that come with open peer reviews. We gathered ground-truth data by contacting the original reviewers and asking them to categorize their own review comments according to our model. Comparing this ground truth to answers provided by model experts, peers, and automated techniques confirms that our approach of formally capturing the reviewers' intentions from the start prevents substantial discrepancies compared to when this information is later extracted from the plain-text comments. In general, our analysis shows that our model is well understood and easy to apply, and it revealed the semantic properties of such review comments.

2.1 INTRODUCTION

Scientific articles and the peer reviewing process they undergo are at the core of how scientific research and discoveries are communicated and assessed. For the last 300 years, such scientific articles have played a crucial role in facilitating scientific progress, and peer reviewing has been crucial to ensure the quality and integrity of these scientific contributions. While communication has changed dramatically in the digital age in almost all fields of life, the paradigm of scientific publishing has remained remarkably stable. Journals, articles, and peer reviews are now mostly produced and consumed in digital form, but their structure and appearance has mostly stayed the same.

Although alternative ways of scientific publishing have been explored, most of the suggested approaches still depend on large bulks of text in natural language, digitized and maybe semantically enriched, but essentially following the same old publishing paradigm. While this might not be seen as a problem by itself, the general statistics of scientific publishing point to a serious crisis of information overload [86]. An increasing number of articles is published every day, and the number of publishing scientists worldwide increases with approximately 4–5% per year [163]. As a direct consequence, scientists need more time to stay up to date with the recent developments in their respective fields. Already in 2004, for example, epidemiologists in primary care needed more than 20 hours per week if they wanted to read all the new articles in their field [3].

Moreover, we are also facing challenges with respect to the quality of research. Peer reviewing has been seen as a key pillar of the scientific publishing system and the main method of quality assessment. However, despite its established and central role, it is a hotly debated topic in the scientific community due to issues like taking unreasonable amounts of time, lack of transparency, inefficiency, lack of software integration, lack of attribution for reviewers, unsystematic quality assessment, and even poor scientific rigor [17]. Actually, the agreement level between peer reviewers has been found to be only slightly better than what one would expect by chance alone [94].

We can tackle some of these problems by applying the ideas and technologies of the Web and the Semantic Web, allowing for a transparent and accessible medium where contributions have a precise structure, can be reliably linked, and can be correctly attributed. In the research presented here, we propose to apply the general principles of the Web and Semantic Web to the reviewing process at a finer-grained scale. These principles entail the use of dereferenceable HTTP URIs to identify, lookup and access resources on the Web, and to interlink them, thereby forming a large network of interconnected resources (or "things"). Or, in Tim Berners-Lee's words [16], forming "a web of data that can be processed directly and indirectly by machines" and humans alike. In such a Web-based system, which does not depend on a central authority, everybody can say anything about anything. To put all these statements into context, we can moreover capture provenance information that will allow us to judge the quality and reliability of the provided information. Therefore, following these principles, we can represent information in the form of immutable nodes in a Web-wide distributed network and accurately track the history and provenance of each snippet of information.

This approach of semantically modeled nodes in a network of information snippets can then also use reasoning techniques to aggregate information in dynamic views, instead of the static representations of classical journal articles. Peer reviews can then be represented on a finer-grained scale, with explicit links to the specific parts (e.g. paragraphs) of the scientific contribution they are about. This in turn enables more detailed and more precise quality assessments of both, the scientific contributions and the reviews themselves.

This chapter is structured as follows. In Section 2.2 we describe the current state of the art in the field of scientific publishing and the reviewing process in particular. In Section 2.3 we describe our approach with regard to performing the reviewing process on a fine-grained scale and introduce a novel model, Linkflows, that combines already existing ontologies in order to provide a more detailed, semantically rich view on the reviewing process. In Section 2.4 we describe in detail how we performed the evaluation of the Linkflows model in the reviewing context, while we report and discuss the results of this evaluation in Section 2.5. Future work and conclusion of the present research are outlined in Section 2.6.

2.2 RELATED WORK

In the context of the Semantic Web, considerable research has been done in formally enriching the meaning of traditional articles. This can facilitate the discoverability of the article and its linking to other related articles or concepts. Important outcomes of these efforts include SPAR[2], which are a set of core ontologies around semantic publishing, Publishing Requirements for Industry Standard Metadata (PRISM)[3], which uses metadata terms to describe published works, and Simple Knowledge Organization System (SKOS)[4], which provides a formal backbone model for knowledge organization systems.

In order to address the information overload crisis, solutions to automatically extract structured information from scientific texts have been proposed. The most common of these include text mining techniques like association rule mining and Inductive Logic Programming powered by large knowledge bases [71]. These approaches, however, do not attempt to transcend the current publishing paradigm and are therefore limited in their potential to address the core of the problem, which is that knowledge is "burried" in narrative texts in the first place, from which it has to be "mined" again in a costly and error-prone process [103]. To address this problem at its source, alternative approaches have been proposed to make scientific communication more structured and machine-understandable from the beginning. Efforts in this direction include the concept and technology of nanopublications, which allow for representing scientific claims in small linked data packages based on RDF and formal provenance [58], with extensions that allow for informal and partly formal representations to extend their application range [78]. The micropublication model is a related effort that emphasizes the importance of the argumentation structure when applying formal knowledge representation methods to scientific communication [34]. Approaches that build upon these

2 http://www.sparontologies.net/
3 https://www.idealliance.org/prism-metadata
4 https://www.w3.org/2004/02/skos/

concepts include work that combines the micropublication ontology with Open Annotation Data Model (OADM) to model knowledge about drug-drug interactions and to link them to their scientific evidence in articles [132]. All these approaches are aligned with the general ideas of "genuine semantic publishing" [82] and Linked Science [72].

With respect to quality assessment of scientific publications, one of the most widely used indicators in the last 40 years is the JIF [54, 88], but this metric has been the subject of extensive debates, as it was shown that it can be manipulated [74], and has problems like skewness of citations, false precision, absence of confidence intervals, and asymmetry in its calculation [36]. Also, the JIF can be biased towards journals that publish a larger number of non-research items (e.g. research notes or comments) or publish more articles overall [42]. To address these issues, approaches like the daQ [38] use Semantic Web technology to enable more accurate and more flexible measures of quality, while the dimensions of quality have been investigated in their own right in the specific context of Linked Data and the Semantic Web [169].

A deciding factor for the publication of a scientific article is the evaluation by the peers in the field: the peer reviews. While this system for evaluating the quality of research has been used for a long time, researchers have also outlined the flaws of this process [147], advocating for a change. Some point out that there is no conclusive evidence in favor of peer reviews [91], that the existing evidence actually corroborates flaws in this system [148], and that biases and preconceptions in judgments affect the overall evaluation of quality that the peer reviewers make [12]. Proposals that have been put forward to increase the quality of peer reviews, like better review structure [90, 152] or paying peer reviewers for their work [41], have seen little uptake so far. Interesting developments have emerged recently around making reviews more fine-grained [128] and representing the structure of reviews with a dedicated ontology[5].

As datasets, documents and knowledge in general are spread on the Web, where everything can be shared, reused, and linked, decentralization is a key concept. Decentralization implies that no central authority — like a publishing house — has global control over the content or participants of the system. A large amount of research has investigated the prerequisites and consequences of decentralization in general [138], and to a lesser extent specifically in the context of scientific publishing. The former includes recent initiatives like Solid[6], a completely decentralized LD framework. The latter include approaches to ensure the functioning of a secure and decentralized global file system [97], the application of the BitTorrent peer-to-peer file sharing protocol to distribute and access scientific data [96], decentralized data

5 http://fairreviews.linkeddata.es/def/core/index.html
6 https://solid.inrupt.com/

publishing approaches for RDF data [49], nanopublications [79], RASH[7] and dokieli [27].

2.3 APPROACH

In this work, we aim to answer the following research question: *Can an approach for scientific publishing based on a fine-grained semantic model help to make reviewing better structured and more accurate?*

2.3.1 *General Approach*

In our approach, we apply the general principles of the Web and the Semantic Web to the field of scientific publishing. While the Web consists of a distributed network of documents that link to each other, the Semantic Web adds to that a more fine-grained network at the level of data and knowledge, where the nodes are concepts, i.e. domain entities, not documents. Following the principles of the Web and the Semantic Web, we argue for fine-grained publication of scientific knowledge in the form of small distributed knowledge snippets that form the nodes in a network, to replace the long prose texts we currently find in articles and reviews. These knowledge snippets are represented with formal semantics, and identified and located with the help of dereferenceable URIs. In this way, scientific knowledge can be shared and accessed as a continuously growing decentralized network of machine-readable snippets of scientific contributions, instead of a static and machine-unfriendly collection of long texts.

2.3.2 *Linkflows Model of Reviewing*

At the core of our approach, we propose an ontology for granular and semantic reviewing. Figure 1 shows the main classes and properties of this ontology that we call the Linkflows model of reviewing. The Linkflows model visualization[8] was made with the online WebVOWL tool and the formal ontology specification[9] of the model can be found online.

The main class of the ontology is the *Comment* class, which includes review comments (subclass *ReviewComment*), on which we focus here, but the class also includes general text annotations or any kind of comment about a text snippet that comes with a dereferenceable URI. The general properties of the model are *refersTo*, which connects a comment to the entity the comment is about, *isResponseTo*, which declares that a comment is a response to another comment, *isUpdateOf*,

---

7 https://github.com/essepuntato/rash

8 https://service.tib.eu/webvowl/#iri=https://raw.githubusercontent.com/LaraHack/linkflows_model/master/Linkflows.ttl

9 https://github.com/LaraHack/linkflows_model/blob/master/Linkflows.ttl

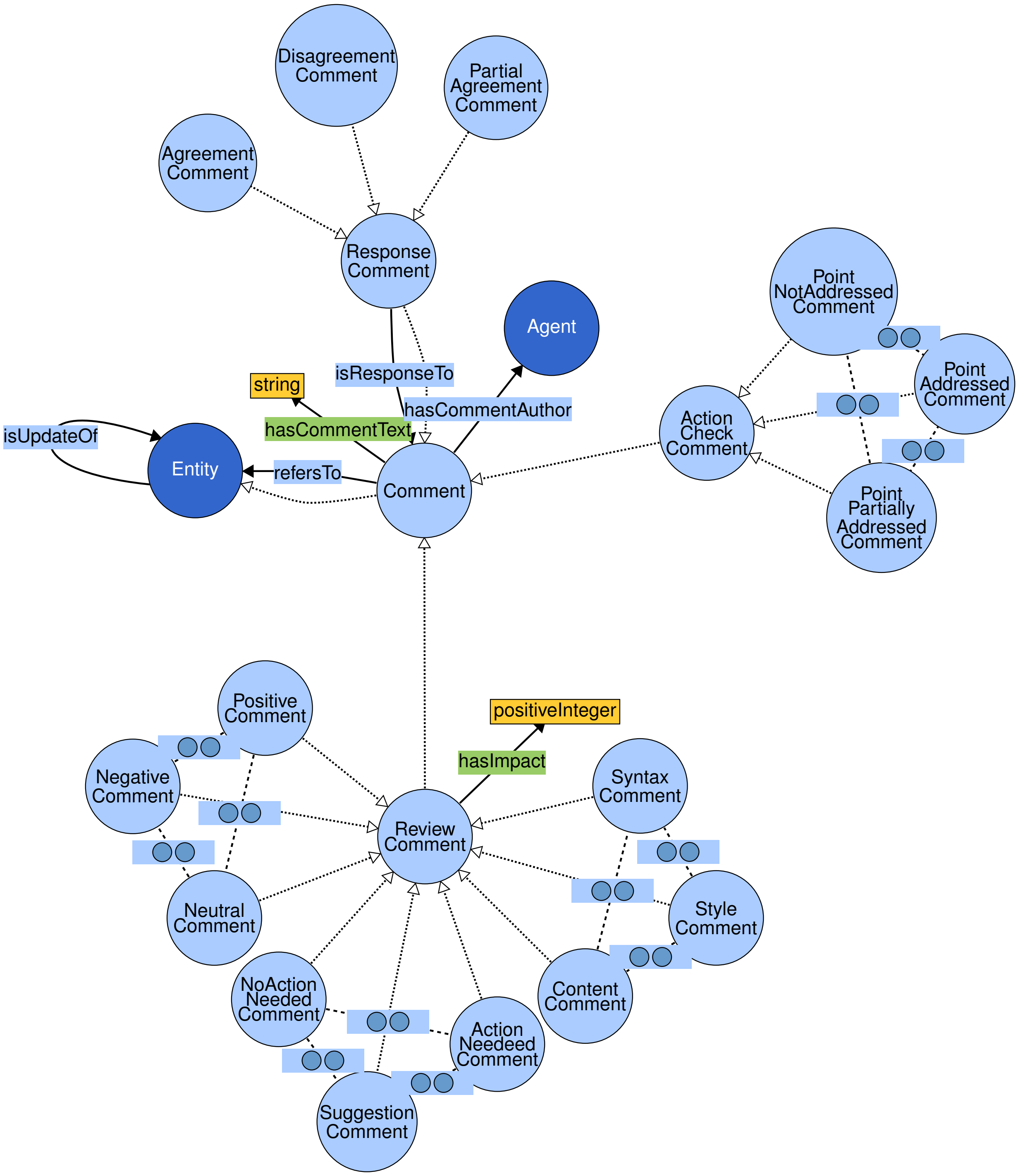


Figure 1: Linkflows model for reviewing.

which connects an entity such as a comment to its previous version, *hasCommentText*, which links to the text content of a comment, and *hasCommentAuthor*, which declares the person who wrote a comment.

Our model defines three dimensions for review comments with different categories, each defined as subclasses of the class *ReviewComment*, which form the core of our semantic representation of reviews. The first dimension is about whether the point raised in the review comment is about individual spelling or grammar issues (*SyntaxComment*), the general style of the text including text structure, text flow, text ordering, and consistent wording (*StyleComment*), or the content of the text, e.g. about missing literature, the presented arguments, or the validity of the findings (*ContentComment*). The second dimension is the positivity/negativity of the review comment: *PositiveComment* for review comments that mainly raise positive points, *NeutralComment* for neutral or balanced points raised, and *NegativeComment* for the cases with mainly negative points. The third dimension captures whether an action is needed in response to the review comment (according to the reviewer): *ActionNeededComment* means that the reviewer thinks his or her comment necessarily needs to be addressed by the author; *SuggestionComment* stands for comments that may or may not be addressed; and *NoActionNeededComment* represents the comments that do not need to be addressed, such as plain observations. On top of that, we define a datatype property *hasImpact* that takes an integer from 1 to 5 to represent the extent of the impact of the point raised in the review comment on the overall quality of the article according to the reviewer. For negative comments this score indicates what would be the positive impact if the point is fully addressed, while for positive points it indicates what would be the negative impact if this point were not true. To further clarify these dimensions and how they could be implemented in an actual reviewing system, we show a mockup of such an interface in Figure 2.

For representing the interaction between reviewers and authors and follow-up actions that are necessary or requested by the reviewer, the *ResponseComment* and *ActionCheckComment* were created as subclasses of *Comment*. The author of the text snippet that was commented upon can agree, disagree, or partially agree with the review comment of the reviewer and they can indicate this classifying their response comment accordingly as *AgreementComment*, *PartialAgreementComment*, or *DisagreementComment*.

Finally, reviewers or editors can indicate in another follow-up comment whether they think that the author indeed addressed the point raised by the reviewer (to which they might have agreed, disagreed, or partially agreed). This can be expressed by the subclasses *PointAddressedComment*, *PointPartiallyAddressedComment*, and *PointNotAddressedComment*.

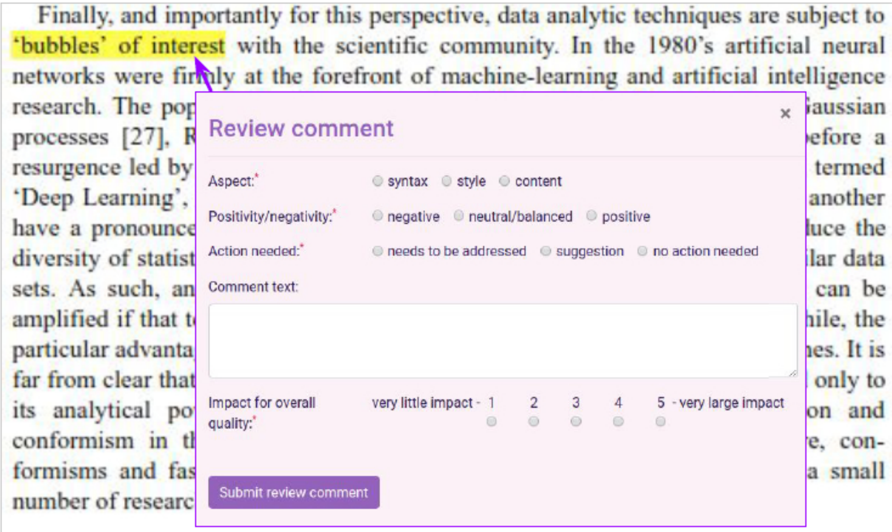


Figure 2: Mock interface implementing the Linkflows model for reviewing. *The entries reported in the interface correspond to the review dimensions identified in the model.*

2.4 EVALUATION DESIGN

We evaluate our approach based on recently published articles that come with open peer reviews. We use these data to simulate how the Linkflows model for reviewing would have worked, had this fine-grained semantic model been applied to these manuscripts from the start. Specifically, we performed the following steps: First, we created a dataset consisting of 35 articles with open reviews and rebuttal letters (where available). Second, the review snippets were manually annotated by model experts by applying our model. Third, we asked reviewers of the selected articles to rate one of their own review comments with regard to our model, thereby giving us ground-truth data. Fourth, we asked experts via a questionnaire to apply our model to the review snippets for which we have ground-truth answers. Fifth and lastly, we applied different automated sentiment analysis methods on the review comments to check if these automated methods would be able to correctly identify the positivity or negativity in these review comments. These five steps are explained in more detail below.

2.4.1 *Dataset*

In order to study how well our model can capture the reviewing process, we needed a dataset of manuscripts and their reviews. For that, we selected journals and conferences in Computer Science that make their reviews openly available in a non-anonymous way. Because this data preparation involves a lot of manual work, we had to restrict ourselves to a subset of all available articles, and a subset of the reviewers they had.

Specifically, we started by considering all the 38 articles published in 2018 in the Semantic Web Journal (SemWeb)[10], the 13 articles published in the first edition of the journal Data Science (DS)[11], and the 25 articles published in 2018 in the PeerJ in Computer Science (PJCS)[12] journal. Additionally, we collected data from two conferences where article versions and reviews are openly accessible via the openreview.net platform [150]. The only conferences in openreview.net that had a complete number of submissions (all articles and reviews) uploaded on the platform at the time we created our dataset were two workshops in 2018: Decentralizing the Semantic Web from the International Semantic Web Conference (ISWC-DeSemWeb)[13] and the International Workshop on Reading Music Systems from the International Society for Music Retrieval Conference (ISMIR-WoRMS)[14]. ISWC-DeSemWeb had 10 submissions and ISMIR-WoRMS had a total of 12 submissions.

From this set of articles, we first filtered out the ones that were eventually rejected (we had to focus on accepted papers, because the reviews are typically not openly available for rejected ones), had only anonymous reviewers (because of the ground truth step below), or where one of the authors of this chapter was editor or reviewer (for objectivity reasons). After this first filtering step, we selected seven articles from each of the five data sources (the three journals and the two workshops), resulting in a dataset of 35 articles. This selection was done randomly by applying the SHA-256 hash function on their title strings and picking the seven articles with the smallest hash values. We then applied the same procedure on the reviewer names to select a random non-anonymous peer review for each of the selected papers. We always picked the first-round reviews for the papers that went over multiple rounds of reviews.

Finally, we took the resulting reviews and split them into smaller, finer-grained snippets of review comments according to our model. Each of these snippets covers a single point raised by the reviewer, and they often correspond to an individual paragraph or an individual list item of the original review. We did an initial annotation of the review comments corresponding to the 35 selected articles and reviews with regard to the target of the review comments. This target of a review comment can be (a) the entire article, (b) a section or (c) a paragraph (or a similar type of structure in terms of the granularity of the ideas expressed like figure, table, diagram, sentence, footnote, listing, reference, etc.). The level of granularity of the review comments can later indicate, for example, at which level the text received most comments,

10 http://www.semantic-web-journal.net/issues#2018
11 https://content.iospress.com/journals/data-science/1/1-2
12 https://peerj.com/computer-science
13 https://openreview.net/group?id=swsa.semanticweb.org/ISWC/2018/DeSemWeb
14 https://openreview.net/group?id=ISMIR.net

or to what extent the authors chose to address the issues raised on the different levels. All this is aligned with the Linkflows view of scientific communication as a network of inter-connected distributed information snippets. A review comment and the paragraph it addresses, for example, are two connected nodes in this network, and so are all the paragraphs and sections of an article.

In summary, therefore, our dataset covers 35 pairs of articles and reviews, where the reviews are split into smaller snippets that allow us now to further simulate the application of our model by manual annotation. This dataset is accessible online[15].

2.4.2 *Ground-Truth and Manual Annotation*

In order to find out how well the Linkflows model would have worked had it been applied to the papers of our dataset in the first place, we approached the specific peer reviewers who wrote the reviews in our dataset (which is why it was important to exclude anonymous reviewers) and asked them in a questionnaire to categorize one of their existing review comments according to our new model. We can then use their responses as ground-truth data. Specifically, they were asked whether a review comment was about syntax, style or content, whether they raised a positive or negative point, whether an action was mandatory or suggested, what the impact level was for the overall quality of the article, and whether the author eventually addressed the point.

With our envisaged approach, the various actors (most importantly the reviewers) would in the future directly contribute to the network of semantically represented snippets. This kind of annotation activity is therefore only needed for our evaluation methodology but would not be needed in the future when our approach is applied.

Considering the limited time people have for filling in questionnaires, we selected just one review comment for each peer reviewer. For that, we again applied SHA-256 hashing on the review text and chose the one with the smallest hash value. For simplicity, we only considered paragraph-level comments here. To optimize the questionnaire that asks reviewers to provide the classifications according to our model, we conducted a small pilot study in the Semantic Web groups at the Vrije Universiteit Amsterdam. We received 12 answers and incorporated the feedback in the further design of the questionnaire before we sent it to the peer reviewers. Out of 35 contacted peer reviewers (corresponding to the 35 selected articles in our dataset), eleven replied. This therefore resulted in a ground-truth dataset of eleven review comments.

In order to further assess the applicability of the model, the three authors of this chapter independently annotated the review comments

15 https://github.com/LaraHack/linkflows_reviewing_study

of the ground-truth datasets as well, considering all review snippets of all reviewers for these articles. This will allow us to compare our answers to the ones of the ground truth, but also to quantify the agreement among us model experts by using Fleiss' Kappa, which can be seen as an indication of the clarity and quality of the model.

For comparison, we also calculate a random baseline of equally distributed and fully uncorrelated responses that correspond to the outcome of random classification. This will serve as an upper baseline for the disagreement with the ground truth.

2.4.3 *Peer Questionnaire*

Next, we wanted to find out how important it is to let reviewers themselves semantically express their reviews, instead of extracting this representation afterwards from classical plain-text based reviews. Specifically, we wanted to find out to what extent the interpretation of a review comment by a peer researcher would differ from the meaning that the reviewer had in mind. We hypothesize that there is a substantial level of disagreement, which would demonstrate that we cannot reliably reconstruct the reviewer's intention in detail if we don't let him or her express these explicitly when authoring the review. This in turn, can then be seen an indication of the benefit and value of our approach.

For this part, we used the same kind of questionnaire that we also used for the ground-truth collection, now covering all eleven review snippets that are covered by the ground-truth data. We first conducted again a small pilot study, to see how long it will take to fill in such a questionnaire. Due to the assumed time constraints of potential respondents, we decided to split the questionnaire randomly in two parts, such that filling in one part of the questionnaire would take somewhere between 10–15 minutes. We then sent the two parts of the questionnaire to various mailing lists in the field of Computer Science.

2.4.4 *Automated Sentiment Analysis*

Finally, we wanted to compare the results we get from human experts with what we can achieve with fully automated methods. As one of the main objectives of this research is to make reviewing more structured, we investigated the extent to which this can be achieved automatically by using the Linkflows fine-grained model for reviews. The positivity/negativity dimension of this model appears to be well-suited for automated tools and various automated sentiment detection tools can be used for this purpose.

Specifically, we applied off-the-shelf sentiment analysis tools on our review texts and compared their resulting sentiment scores to the positivity/negativity values as annotated by the original reviewers.

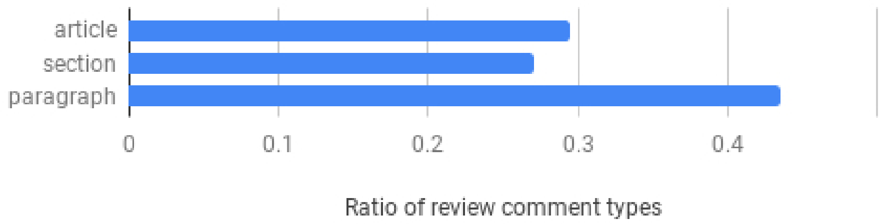


Figure 3: The part of the article that a review comment targets.

Sentiment, however, is not exactly the same as positivity/negativity: Sentiment analysis focuses on 'how' things are expressed (i.e. whether positive/polite language is used) while our positivity/negativity dimension captures 'what' is said (i.e. whether the raised point is a positive one). Reviews can be very politely phrased yet contain very serious criticism, thus expressing a negative point with positive or neutral language. Moreover, sentiment analysis tools are typically trained on Web documents such as news items and not on scientific texts or reviews, which have a different vocabulary and language use. We therefore hypothesize that the performance of sentiment analysis tools to detect positivity/negativity in our sense will in general not be satisfactory.

We used a benchmark sentiment analysis platform called iFeel[16] [126] to automatically detect whether the review comments in our ground-truth dataset express negative, neutral/balanced, or positive points. We ran all the 18 lexicon-based sentiment analysis methods available of the benchmark platform and collected the results. Finally, we calculated the accuracy of these automated sentiment analysis methods versus the ground-truth data.

2.5 RESULTS

We can now have a look at the results we obtained by conducting the studies outlined above.

2.5.1 *Descriptive Analysis*

Figure 3 shows the distribution of the total of 421 review comments in our dataset with respect to the type of article structure they target. 29% of them are about the article as a whole, 27% are at the section level, and the remaining 44% target a paragraph or an even smaller part of the article. This indicates that the review comments indeed often work at a granular level and that the application of a finer-grained model, such as the Linkflows model, therefore indeed seems appropriate and valuable.

16 http://blackbird.dcc.ufmg.br:1210/

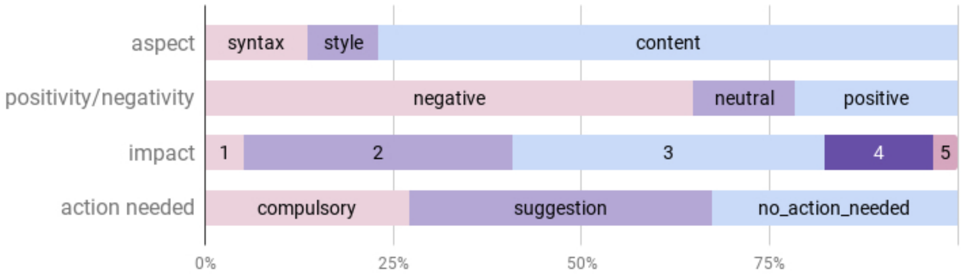


Figure 4: The results of the model expert annotations.

Table 2: Percentages of questions where it was not possible to answer (Peers experiment).

	aspect	positivity/ negativity	action needed	impact	action taken	Overall
"more context needed"	0.75%	1.50%	2.64%	4.90%	2.64%	2.49%
"confusing"	1.89%	1.70%	1.32%	2.07%	0.94%	1.58%
Total	2.64%	3.20%	3.96%	6.97%	2.58%	4.07%

Figure 4 shows the results of the manual annotation by us model experts on the 450 review comments that correspond to the articles included in the ground-truth subset (articles for which we have one review snippet annotated by one of its peer reviewers). We notice that most of the review comments are about the content of the text, while comments about style and syntax were less common. In terms of the positivity/negativity, we see that — unsurprisingly — most of the review comments were rated as negative. The impact of the review commentx for the overall quality of the article is rarely assigned the extreme values of 1 or 5, but the remaining three values of 2–4 are all common (2 and 3 more so than 4). The action that needs to be taken according to the reviewer, finally, shows a balanced distribution of the three categories (compulsory, suggestion, and no-action-needed).

For all of the 450 review snippets that we rated, the average degree of agreement based on Fleiss' kappa [50] had a value of 0.42 which indicates moderate agreement between raters for all dimensions of the Linkflows model [87]. We notice, however, a considerable variation across dimensions: the aspect and positivity/negativity dimensions had the highest agreement with values of 0.68 and 0.62 respectively (substantial agreement), the action needed had a value of 0.37 (fair agreement), while the impact dimension gave a value of just 0.03 (slight agreement). This low value for the impact dimension is not so surprising given that it consists of just a numerical scale without precise definitions of the individual categories. Moreover, it has a larger number of categories which is known to lead to worse kappa values. But, altogether, the substantial inter-annotator agreement shows that our model is well applicable and sufficiently precise.

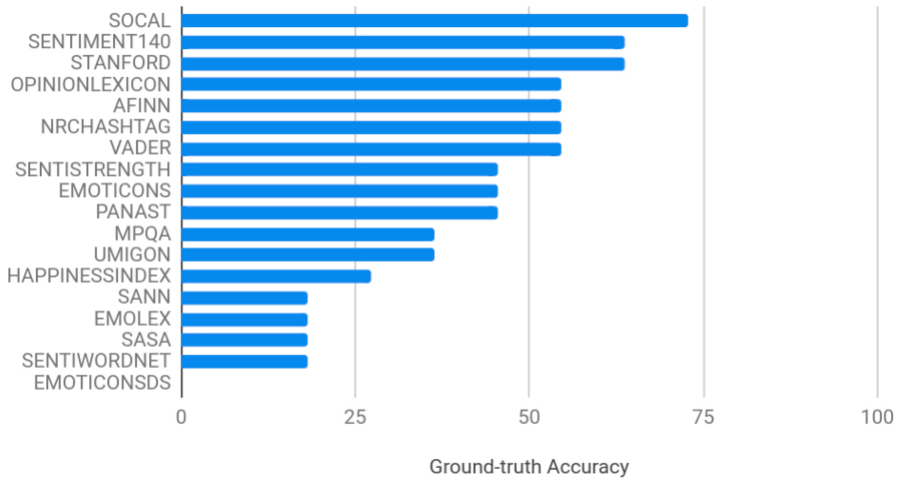


Figure 5: Accuracy of sentiment analysis methods.

For the peer questionnaire study, we gathered 79 responses in total: 43 responses for the first part of the questionnaire containing 5 review comments (215 answers in total) and 36 answers for the second part with 6 review comments (216 answers in total). The participants had to answer a few questions about their background, which revealed that 58.2% of the respondents have a university degree and that most of them are working in academia (89.9%). The majority (75.9%) has advanced knowledge of Computer Science, while only 5% consider themselves beginners in this field. This confirms that they can indeed be seen as peers of the reviewers and authors of our dataset of Computer Science papers.

When answering the questions, these peers could also choose the options "More context would be needed; it is not possible to answer" or "The review comment is confusing; it is not possible to answer", when they didn't feel confident to give an answer. We can now take this as an indication of how well the model worked out from their point of view. Table 2 shows these results. We see some variation across the different dimensions, but at a very low level. The two options combined were chosen in less than 7% of the cases for any of the dimension, and overall in just a bit over 4% of the cases was one of these two options chosen. This indicates that the dimensions of the Linkflows model are overall very well understood and easy to apply.

2.5.2 *Accuracy of Sentiment Analysis Methods*

Before we move on to compare the results from the different sources, we first analyze the performance of the different sentiment analysis tools here, in order to be able to chose the best ones for inclusion in the comparative analysis described below.

Table 3: Disagreement scores for four of the Linkflows model dimensions.

	aspect			action needed			impact			action taken		
	R	M	P	R	M	P	R	M	P	R	M	P
Reviewer (R)	0			0			0			0		
Model experts (M)	0.32	0		0.26	0		0.21	0		0.38	0	
Peers (P)	0.29	0.12	0	0.29	0.17	0	0.31	0.14	0	0.34	0.21	0
Random baseline	0.67			0.67			0.50			0.67		

Table 4: Disagreement scores for the positivity/negativity dimension.

	R	M	P	SA1	SA2	SA3
Reviewer (R)	0					
Model experts (M)	0.32	0				
Peers (P)	0.16	0.24	0			
SOCAL (SA1)	0.26	0.41	0.24	0		
SENTIMENT140 (SA2)	0.30	0.16	0.22	0.34	0	
STANFORD (SA3)	0.30	0.16	0.22	0.34	0	0
Random baseline	0.58					

The results we get from the iFeel platform are already normalized to the values of negative, neutral, and positive, which directly map to the positivity/negativity dimension of our model. We then define accuracy in this context as the ratio of correctly classified instances when compared to the ground truth we collected from the reviewers.

In Figure 5 we summarized the results of the 18 sentiment analysis methods used to rate the eleven review comments for which we have ground-truth data. We see a wide variation of performance with a maximum accuracy of 72.8% for the SOCAL method [153]. As expected, most of these methods perform quite poorly. In general, we can observe that the sentiment analysis methods that have more complex rules perform best, even if their lexicon size is not large. Apart from SOCAL, we also select the tools ranked second and third for our comparative analysis below: Sentiment140 [57] and Stanford Recursive Deep Model [149] (both 63.6% accuracy).

2.5.3 *Comparative Analysis*

We can now have a closer look into the differences between the answers we got from the different sources, to better understand the nature and extent of the observed disagreement. We quantify this disagreement by applying a variation of the Mean Squared Error metric. To compare the responses of two groups, we calculate normalized (0–1) within-group averages of their responses and then take the square root of the mean squared differences between the groups. For the nominal dimensions, we calculate the squared differences from the ratio for each category separately.

These disagreement scores between reviewers, model experts, and peers for all dimensions of the model are shown in Table 4 (for the positivity/negativity dimension, which includes the sentiment analysis tools) and Table 3 (for all other dimensions).

Looking at Tables 3 and 4 we see that the agreements between the reviewers and the groups of model experts and peers range from 0.12 to 0.38, well above perfect agreement (0) but also well below the random baseline (0.5 for the ordinal dimension *impact*, 0.58 for the ordinal dimension *positivity/negativity*, and 0.67 for the nominal dimensions). It is notable that the model experts and the peers always agree with each other more than they agree with the ground truth in the form of the original reviewer. This seems to indicate that they misinterpret the review comments in a relatively small but consistent manner. The highest disagreement scores for all pairs of groups come from the *action taken* dimension, which seems in general slightly more difficult than the others.

With respect to the positivity/negativity dimension (Table 4), we see that the best automated sentiment analysis tool surprisingly performed a bit better than the model experts (0.26 versus 0.32 disagreement score, compared to the ground truth). A further surprising outcome is that the peers performed much better than the model experts and the sentiment analysis tools. We will further investigate this effect below. Moreover we see that the second and third best sentiment analysis tools provide identical results (and therefore zero disagreement), but that they have considerable disagreement with the best tool (SOCAL), which might hint at complementary information given by these tools and therefore the potential for an ensemble method.

Table 5 focuses on the comparison of the two main groups of model experts and peers. We see that for three of the five dimensions, the peers had a higher agreement (i.e. lower disagreement) with the ground truth than the model experts. Testing these observed differences for statistical significance with a two-tailed Wilcoxon signed-rank test, we see however that none of these differences is significant.

From the somewhat surprising result that model experts do not perform significantly better than peers, we could conclude that expertise on the model doesn't actually help with the task. It could also be that the benefit of this expertise was offset by the fact that the model experts shared more detailed information and acquired background knowledge through in-depth discussions with each other, and therefore they had information neither the reviewers nor the peers had. This might have made the reviewers and the peers settle on similar choices that are different from the model experts'. Because the group of peers consists of more than ten times as many individuals as compared to

Table 5: Model experts versus peers.

	aspect	positivity/ negativity	action needed	impact	action taken
Model experts	0.32	0.32	0.26	0.21	0.38
Peers	0.29	0.16	0.29	0.31	0.34
Difference	0.03	0.16	−0.03	−0.10	0.04
p-value (Wilcoxon signed-rank test)	0.21	0.28	0.42	0.09	0.66
Average of groups of 3 peers	0.34	0.22	0.31	0.34	0.44
Standard deviation of groups of 3 peers	0.026	0.010	0.119	0.055	0.016

the model expert group, an alternative explanation is that a wisdom of the crowd effect is involved, where averaging over a larger number of individual responses gets us closer to the truth.

In order to test this wisdom of the crowd hypothesis, we divided the answers from the peers into twelve groups of three peers each (thereby matching the size of the model expert group) and calculated the disagreement scores for each of these smaller groups. The bottom part of Table 5 shows the average results. We see that the average disagreement of the smaller peer groups is always larger than the disagreement of the large peer group, confirming that there seems to have been a wisdom of the crowd effect at play. Except for positivity/negativity, where the difference was largest to start with, the average disagreement of the smaller peer groups is larger than for the model experts. Therefore, there seems to be indeed a wisdom of the crowd effect at play that makes larger groups perform better than smaller ones.

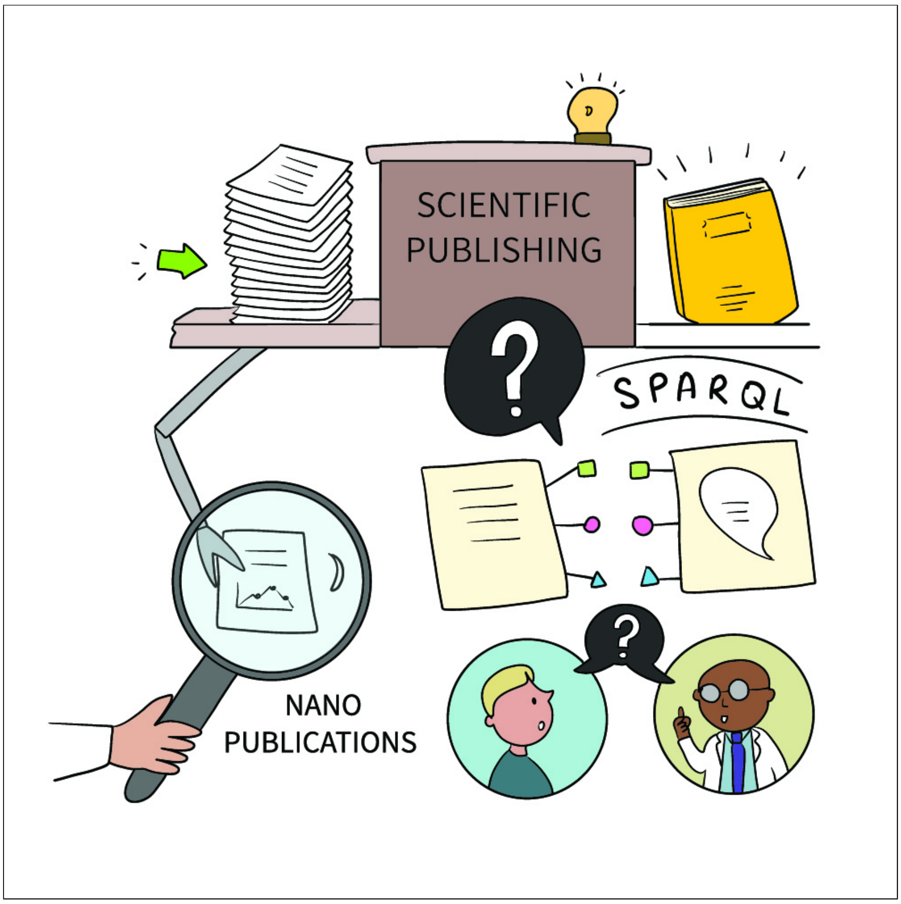
2.6 DISCUSSION AND CONCLUSION

The results of this study suggest that the Linkflows model is indeed able to express at a finer-grained level the properties of review comments and their relation to the article text. We propose to let reviewers create semantically represented review comments from the start. Our evaluation showed that a substantial level of disagreement arises if other actors like peer researchers or model experts try to reconstruct the intended nature of these review comments afterwards, thereby underlying the importance of capturing this information right at the source. Through a wisdom of the crowd effect, larger groups of peers can achieve lower disagreement with the ground-truth provided by the original reviewers, but the disagreement is still substantial and larger groups of annotators require more collective effort. Existing automated methods like sentiment analysis tools can reach the same level of agreement like model experts, but a substantial level of disagreement remains for both of them. In summary, this confirms the value of our approach to precisely capture the network and type of review comments directly from the reviewers.

While we have provided a simple mock-up of an interface for reviewers, an actual prototype still has to be implemented and evaluated as future work. Specifically, we plan to experiment with applying the Dokieli environment [27] and Linked Data Notifications (LDN) [26]. To publish reviews (and also the articles themselves) as a network of semantically annotated snippets in a decentralized way, we will furthermore investigate the application of the nanopublication technology and infrastructure [79] for reliable semantic publishing.

3

A UNIFIED NANOPUBLICATION MODEL FOR THE ELEMENTS OF SCIENTIFIC PUBLISHING



MAIN CLAIM FORMALIZATION: In the context of all things of type *scholarly communication*, things of type *application of unified and semantic publication model based on nanopublications can generally* have a relation of type *increases* to things of type *effectiveness* that are in the same context (i.e. *scholarly communication*).

FORMALIZATION NANOPUBLICATION: (link to nanopublication)[1]

PUBLICATION: This chapter was previously published as *A Unified Nanopublication Model for Effective and User-Friendly Access to the Elements of Scientific Publishing* in the Proceedings of the 22nd International Conference on Knowledge Engineering and Knowledge Management, 2020, and was co-authored by Tobias Kuhn, Davide Ceolin and Jacco van Ossenbruggen [23].

1 https://w3id.org/np/RAOduiP8tUDRwuX5WSFo1IwSV89YMq63_lqE2NM9ldP4s

ABSTRACT

Scientific publishing is the means by which we communicate and share scientific knowledge, but this process currently often lacks transparency and machine-interpretable representations. Scientific articles are published in long coarse-grained text with complicated structures, and they are optimized for human readers and not for automated means of organization and access. Peer reviewing is the main method of quality assessment, but these peer reviews are nowadays rarely published and their own complicated structure and linking to the respective articles are not accessible. In order to address these problems and to better align scientific publishing with the principles of the Web and Linked Data, we propose here an approach to use nanopublications as a unifying model to represent in a semantic way the elements of publications, their assessments, as well as the involved processes, actors, and provenance in general. To evaluate our approach, we present a dataset of 627 nanopublications representing an interlinked network of the elements of articles (such as individual paragraphs) and their reviews (such as individual review comments). Focusing on the specific scenario of editors performing a meta-review, we introduce seven competency questions and show how they can be executed as SPARQL queries. We then present a prototype of a user interface for that scenario that shows different views on the set of review comments provided for a given manuscript, and we show in a user study that editors find the interface useful to answer their competency questions. In summary, we demonstrate that a unified and semantic publication model based on nanopublications can make scientific communication more effective and user-friendly.

3.1 INTRODUCTION

Scientific publishing is about how we disseminate, share and assess research. Despite the fact that technology has changed how we perform and disseminate research, there is much more potential for scientific publishing to become a more transparent and more efficient process, and to improve on the age-old paradigms of journals, articles, and peer reviews [13, 122]. With scientific publishing often stuck to formats optimized for print such as PDF, we are not using the advances that are available to us with technologies around the Semantic Web and Linked Data [34, 144].

In this work, following our "Linkflows" vision [21], we aim to address some of these problems by looking at the scientific publishing process at a more finer-grained level and recording formal semantics for the different elements. Instead of treating big bulks of text as such, we propose to represent them as small snippets — e.g. paragraphs — that have formal semantics attached and can be treated as independent publication units. They can link to other such units and therefore form a larger entity — such as a full paper or review — by forming a complex network of links. With that approach, we can ensure that provenance of each snippet of information can be accurately tracked together with its creation time and author, and therefore allow for more flexible and more efficient publishing than the current paradigm. A process like peer-reviewing can then be broken down into small snippets and thereby take the specialization of reviewers and the detailed context of their review comments into account, and these review comments can formally and precisely link to exactly the parts of the paper they address. Each article, paragraph and each review comment thereby forms a single node in a network and is each identified by a dereferenceable URI.

We demonstrate here how we can implement such a system with the existing concept and technology of nanopublications, a LD format for storing small assertions together with their provenance and meta-data. We then show how this approach allows us to build powerful and user-friendly interfaces to aggregate and access larger numbers of such small communication elements. In order to assess the concrete benefits, we zoom in to just one out of the countless scenarios in which we can expect substantial advantages from such fine-grained semantic representations. We chose here the concrete case of a system for editors to assess manuscripts based on a set of review comments, and based on this concrete case we demonstrate and assess our approach.

In this research we aim to answer the following research questions:

1. Can we use nanopublications as a unifying data model to represent the structure and links of manuscripts and their assessments in a precise, transparent, and provenance-aware manner?

2. Is a fine-grained semantic publishing and reviewing model able to provide us with answers to common competency questions that journal editors face in their work as meta-reviewers?

3. Can we design an intuitive and effective interface based on a fine-grained semantic publishing and reviewing model that supports journal editors in judging the quality of manuscripts based on the received reviews?

We address these research questions with the following contributions:

- A general scheme of how nanopublications can be used to represent and publish different kinds of interlinked publication elements

- A dataset of 627 nanopublications, implementing this scheme to represent exemplary articles and their open reviews

- A set of seven competency questions for the scenario of journal editors meta-reviewing a manuscript, together with SPARQL representations of these questions
- A prototype of a fine-grained semantic analysis interface for the above scenario and dataset, powered by nanopublications
- Results from a user study on the perceived importance of the above competency questions and the perceived usefulness of the above prototype for answering them

The rest of this chapter is structured as follows. In Section 3.2 we describe the current state of the art in the field of scientific publishing and the reviewing process in particular. In Section 3.3 we describe our approach with regard to performing the reviewing process in a fine-grained manner based on nanopublications. In Section 3.4.1 we describe in detail how we performed the evaluation of our approach, while we report and discuss the results of this evaluation in Section 3.4.2. Future work and conclusion of the present research are outlined in Section 3.5.

3.2 BACKGROUND

Before we move on to describe our approach, we give here the relevant background on scientific publishing, semantic papers, and the specific concept and technology of nanopublications.

Scientific publishing is at the core of scientific research, which has moved in the last decades from print to online publishing [151]. It is, however, still mostly following the paradigm from the print age, with narrative articles being published in journals and assessed by peer reviewers, only the printed volumes having been replaced by PDF files that are made accessible via search engines [92]. Considering the ever increasing number of articles and the increasing complexity of research methods, this old paradigm of publishing seems to have reached its limit, and scientists are struggling to stay up to date in their specific fields [86]. Slowly but steadily, these old paradigms are shifting with open access publishing, semantically enriched content, data publication, and machine-readable metadata gaining momentum and importance [141, 162]. Opposition is also growing against the use of impact factor [53, 54, 106] or h-index as metrics for assessment of the participants in this publication process, and it has been shown that these metrics can be tampered with easily [2, 44, 129, 134].

Advances in Semantic Web technologies like RDF, OWL, and SPARQL have allowed for the semantic enhancement of scholarly journal articles when publishing data and metadata [140, 142]. As such, semantic publishing was proposed as a way to make scholarly publications discoverable, interactive, open and reusable for both, humans and machines, and to release them as Open Linked Data [67, 102, 131]. In order to extract formal semantics from already published papers in an automated manner, sophisticated methods such as the Compositional and Iterative Semantic Enhancement (CISE) [115], conceptual frameworks for modelling contexts associated with sentences in research articles [4] and semantic lenses were developed [40]. Furthermore, HTML formats like RASH have been proposed to represent scientific papers that include semantic annotations [117], and vocabularies like the SPAR suite of ontologies have been introduced to semantically model all aspects relevant to scientific publishing [119]. These approaches mostly work on already published articles, but it has been argued that scientific findings and their contexts should be expressed in semantic representations from the start by the researchers themselves, in what has been named *genuine semantic publishing* [82].

In our previous work [22], we applied the general principles of the Web and the Semantic Web to promote this kind of genuine semantic publishing [82] by applying it to peer reviews. We proposed a semantic model for reviewing at a finer-grained level called Linkflows and argued that Linked Data principles like dereferenceable URIs using open standards like RDF can be used for publishing small snippets of information, such as an individual review comment, instead of big chunks of text, such as an entire review. These small snippets of text can be represented as nodes in a network and can be linked with one another with semantically-annotated connections, thus forming distributed and semantically annotated networks of contributions. The individual review comments are semantically modeled with respect to what part of the paper they target, whether they are about syntax or content, whether they raise a positive or negative point, and whether they are a suggestion or compulsory, and what their impact on the quality of the paper is. We showed on this model that it is indeed beneficial if we capture these semantics at the source (i.e. the peer reviewer in this case).

Nanopublications [58] are a specific concept and technology based on Linked Data to publish scientific results and their metadata in small publication units. Each nanopublication has an assertion that contains the main content (such as a scientific finding), and comes with provenance about that assertion (e.g. what study was conducted to derive at the assertion; or which documents it was extracted from) and with publication information about the nanopublication as a whole (e.g. by whom and when it was created). All these three parts are represented in RDF and thereby machine-interpretable.

It has been shown how nanopublications can also be used for other kinds of assertions, including meta-statements about other nanopublications [78], and in order to make nanopublications verifiable and immutable, *trusty URIs* [81] can be used as identifiers, which include cryptographic hash values that are calculated on the nanopublication's content. A decentralized server network has been established

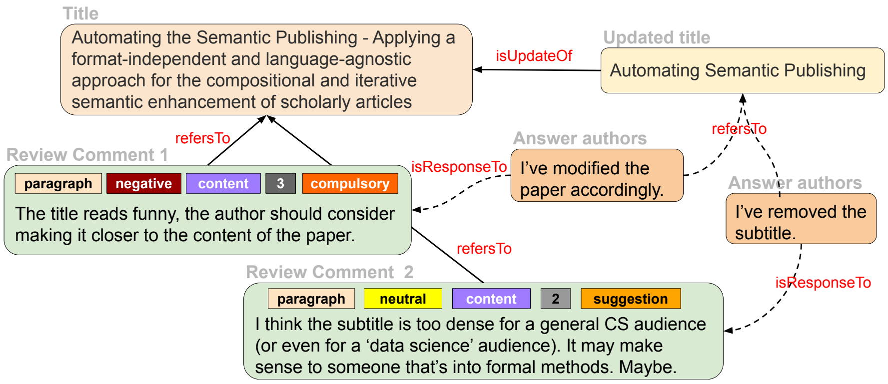


Figure 6: An example of a nanopublication-style communication interaction.

based on this, through which anybody can reliably publish and retrieve nanopublications [79]. In order to group nanopublications into larger collections and versions thereof, index nanopublications have been introduced [84]. With these technologies, small interconnected Linked Data snippets can be published in a reliable, decentralized, provenance-aware manner.

3.3 APPROACH

Our general approach is to investigate the benefits of using using the philosophy and technology of nanopublications as a unifying publishing unit to establish a new paradigm of scientific communication that is better aligned with the principles of the Web and Linked Data. We investigate how such an approach could allow us to communicate in a more efficient, more precise, and more user-friendly manner.

3.3.1 *Semantic Model and Nanopublications*

Our unifying semantic model based on nanopublications uses a number of existing ontologies like SPAR, Provenance Ontology (PROV-O), FAIR* reviews, the Web Annotation Data Model (WADM), and our own Linkflows model [22] to break the big bulks of article and review texts into smaller text snippets. An example of a nanopublication-style communication interaction during the reviewing process is illustrated in Figure 6, where the title of a paper is addressed by several review comments that come with semantic classes (e.g. *suggestion*), which are themselves referred to by the authors' answers that link them to the updated version. Each node in this network is represented as a separate nanopublication and all the attributes and relations are formally represented as Linked Data.

As we can see in Figure 6, the properties *refersTo*, *isResponseTo*, *isUpdateOf* play the key role of linking different nodes in this network. *refersTo* is a property that links a review comment to the text snippet in the article it refers to. *isResponseTo* links the answer of the authors to the review comments of the reviewer and also to new versions of the text snippets that these review comments triggered. *isUpdateOf* links a version of the text snippet to another. In our approach, snippets of scientific articles (mostly corresponding to paragraphs) as well as their review comments (corresponding to individual review comments) are semantically represented as nanopublications [58], and thereby they each form a node in the network described above. A complete example of such a nanopublication containing a review comment is shown in Figure 7.

Each of the three main parts of a nanopublication — assertion, provenance, and publication info — is represented as an RDF graph. In the example of Figure 7, the assertion graph describes a review comment using the classes and properties of the Linkflows model [2]. It raises a negative point with an importance of 2 out of 5, and is marked as a suggestion for the authors. Furthermore, we see that this review comment refers to an external element, with a URI ending in `#paragraph`, as the target of this comment. This external element happens to be a paragraph of an article described in another nanopublication, which we can find out by following that trusty URI link.

Moreover, the nanopublication contains information regarding the creator of the assertion and the creator of the nanopublication that contains this assertion. These pieces of information can be found in the *provenance* and *publication info* graphs. As illustrated in Figure 7, the author of the review comment is indicated by his Open Researcher and Contributor IDentifier (ORCID) and the source of the original source of the review comment is indicated by the the Uniform Resource Locator (URL) pointing to a link of the Semantic Web Journal. From the publication info graph, we can see who created the whole nanopublication together with the date and time of its creation.

As provenance and immutability of scientific contributions are crucial, we use trusty URIs [80, 81] to enforce these properties. As such, for every nanopublication, in order for it to be published, a unique immutable URI is generated to refer to the node that holds the nanopublication. Any change of this nanopublication results in the generation of a new nanopublication, thus of a new node that is linked to the previous one. Such nanopublications can then be published in the existing decentralized nanopublication network [79].

3.3.2 *Use Case with Competency Questions*

In the scientific publishing context, editors of journals play a key role, being an important link between content providers for journals (authors), the people who assess the quality of the content (peer reviewers) and the consumers of such content (the readers). While the

2 https://github.com/LaraHack/linkflows_model

peer reviewers are the ones that can recommend the acceptance or rejection of an article, it is up to the editors to make the final decision. We will look here into how our approach can benefit the specific scenario of editors assessing a manuscript based on given reviews and having to write a meta-review.

Performing such a meta-review is not a trivial task. As classical reviews are mainly comprised of a large bulks of text in natural language, it is hard to provide a tool with quantitative information about the reviews and their collective implications on the manuscript. As such, an editor needs to spend a lot of time just to read these reviews fully to even get an overview of the nature and range of the raised issues.

In order to apply our approach to this chosen use case, we first define a set of Competency Question (CQ)s, which are natural language questions that are created with the objective to assess the practicality and coverage of an ontology or model [14]. After consulting with publishing experts at IOS Press[3] and the Netherlands Institute of Sound and Vision[4] during an informal session, we came up with the following seven quantifiable competency questions from an editor's point of view during meta-reviewing:

- **CQ1**: *What is the number of positive comments and the number of negative comments per reviewer?*
- **CQ2**: *What is the number of positive comments and the number of negative comments per section of the article?*
- **CQ3**: *What is the distribution of the review comments with respect to whether they address the content or the presentation (syntax and style) of the article?*
- **CQ4**: *What is the nature of the review comments with respect to whether they refer to a specific paragraph or a larger structure such as a section or the whole article?*
- **CQ5**: *What are the critical points that were raised by the reviewers in the sense of negative comments with a high impact on the quality of the paper?*
- **CQ6**: *How many points were raised that need to be addressed by the authors, as an estimate for the amount of work needed for a revision?*
- **CQ7**: *How do the review comments cover the different sections and paragraphs of the paper?*

3 https://www.iospress.nl/
4 https://www.beeldengeluid.nl/en

3.3.3 *Dataset*

In order to evaluate our approach on the given use case, we need some data first. For this, we selected three papers that were submitted to a journal that has open reviews (Semantic Web Journal[5]). Therefore, we could also access the full text of the reviews these papers received. We then manually modelled all the article, paragraphs, review comments, their interrelations, as well as their larger structures — in the form of sections and full articles and reviews — as individual nanopublications according to our approach. All these elements were thereby semantically modeled, and we could reuse part of our earlier dataset of manually assigned Linkflows categories [22]. Figure 7 shows an example of a nanopublication that resulted from this manual modeling exercise. We would like to stress here that according to the vision underlying our approach, these semantic representations would in the future be generated as such from the start, and therefore this manual effort is only for evaluation purposes. This should be integrated in the future in smart tools, such that this approach does not come at an additional burden for reviewers but in fact leads to a more efficient way of reviewing.

Apart form nanopublications at the lowest level, such as the one shown in Figure 7, higher-level ones combine them (by simply linking to them) to form larger structures, such as entire sections, papers, and reviews. Section nanopublications, for example, point to their paragraphs and define their order among other metadata. We also created a nanopublication index [84] that refers to this set of manually created nanopublications such that we can retrieve and even reuse parts of this dataset for new versions incrementally. All the nanopublications from our dataset are in an online repository[6].

3.3.4 *Interface Prototype for Use Case*

In order to apply and evaluate our approach on the chosen use case, we developed a prototype of an editor interface that accesses the nanopublications in the dataset presented above to provide a detailed and user-friendly interface to support editors in their meta-reviewing tasks.

5 http://www.semantic-web-journal.net
6 https://github.com/LaraHack/linkflows_model_implementation

```
@prefix this: <http://purl.org/np/RA28QwaD22DK7U1Xq94UESAd3Y2jpGmVMCAk6_7aB_ivc> .
@prefix sub: <http://purl.org/np/RA28QwaD22DK7U1Xq94UESAd3Y2jpGmVMCAk6_7aB_ivc#> .
@prefix dcterms: <http://purl.org/dc/terms/> .
@prefix xsd: <http://www.w3.org/2001/XMLSchema#> .
@prefix prov: <http://www.w3.org/ns/prov#> .
@prefix pav: <http://purl.org/pav/> .
@prefix np: <http://www.nanopub.org/nschema#> .
@prefix linkflows: <https://github.com/LaraHack/linkflows_model/blob/master/Linkflows.ttl#> .

sub:Head {
  this: np:hasAssertion sub:assertion ;
    np:hasProvenance sub:provenance ;
    np:hasPublicationInfo sub:pubinfo ;
    a np:Nanopublication .
}

sub:assertion {
  sub:comment-7 a linkflows:ContentComment , linkflows:NegativeComment , linkflows:ReviewComment , linkflows:SuggestionComment ;
    linkflows:hasCommentText "In section 3.2) you give background information on the Find-Fix-Verify pattern (2nd paragraph). This information (in
which scenario it was used first, etc.) is not really necessary for the rest of the paper." ;
    linkflows:hasImpact "2"^^xsd:positiveInteger ;
    linkflows:refersTo <http://purl.org/np/RAoFc19T2f1F2j25Q6xBtfclZ1LyAZrlD5IU6F_yBozGE#paragraph> .
}

sub:provenance {
  sub:assertion prov:hadPrimarySource <http://www.semantic-web-journal.net/content/detecting-linked-data-quality-issues-crowdsourcing-dbpedia-study> ;
    prov:wasAttributedTo <https://orcid.org/0000-0001-7069-9804> .
}

sub:pubinfo {
  this: dcterms:created "2019-11-26T09:05:11+01:00"^^xsd:dateTime ;
    pav:createdBy <https://orcid.org/0000-0002-7114-6459> .
}
```

Figure 7: Example nanopublication of a review comment.

This prototype comes with two views: one where the review comments are shown per reviewer in a bar chart broken down into the different dimensions and classes, as shown in Figure 8 and another view that focuses on the distribution of the review comments to the different sections of the article, as shown in Figure 9. The interface for an exemplary article with three reviews can be accessed online[7]. The shown content is aggregated from nanopublications stored in a triple store and displayed by showing color codes for the different Linkflows classes for the individual review comments.

In the reviewer-oriented view (Figure 8), we can see in a more quantitative way the set of review comments and their types represented in different colors, where the checkboxes in the legend can be used to filter the review comments of the given category. To see the content of the review comments that are in a certain dimension, it is sufficient to just click on a bar in the chart.

The section-oriented view (Figure 9), aggregates all the finer-grained dimensions of the review comments at the level of sections in an article. Again, clicking on one cell in the table, thus selecting one specific dimension of the review comments, will show the content of those review comments underneath the table in the interface.

When data from the triple store is required, the server (implemented in NodeJS with the Express web application framework[8]) sends a request to the Virtuoso triple store where the nanopublications are stored. This request executes a SPARQL query on the stored nanopublications and returns the result to the server that, in turn, passes it further to the client, in the web browser, where the results are postprocessed and visualized. The code for the prototype can be found online[9].

3.4 EVALUATION

Here we present the evaluation of our approach in the form of a descriptive analysis, the analysis of the SPARQL implementations of our competency questions, and a user study with editors on our prototype interface.

3.4.1 *Evaluation Design*

First, we run a small descriptive analysis on the nanopublication dataset that we created. We can quantify the size and interrelation of the represented manuscripts and reviews in new ways, including the

7 http://linkflows.nanopubs.lod.labs.vu.nl
8 https://nodejs.org, https://expressjs.com/
9 Interface: https://github.com/LaraHack/linkflows_interfaces
Backend application: https://github.com/LaraHack/linkflows_model_app
Data: https://github.com/LaraHack/linkflows_model_implementation

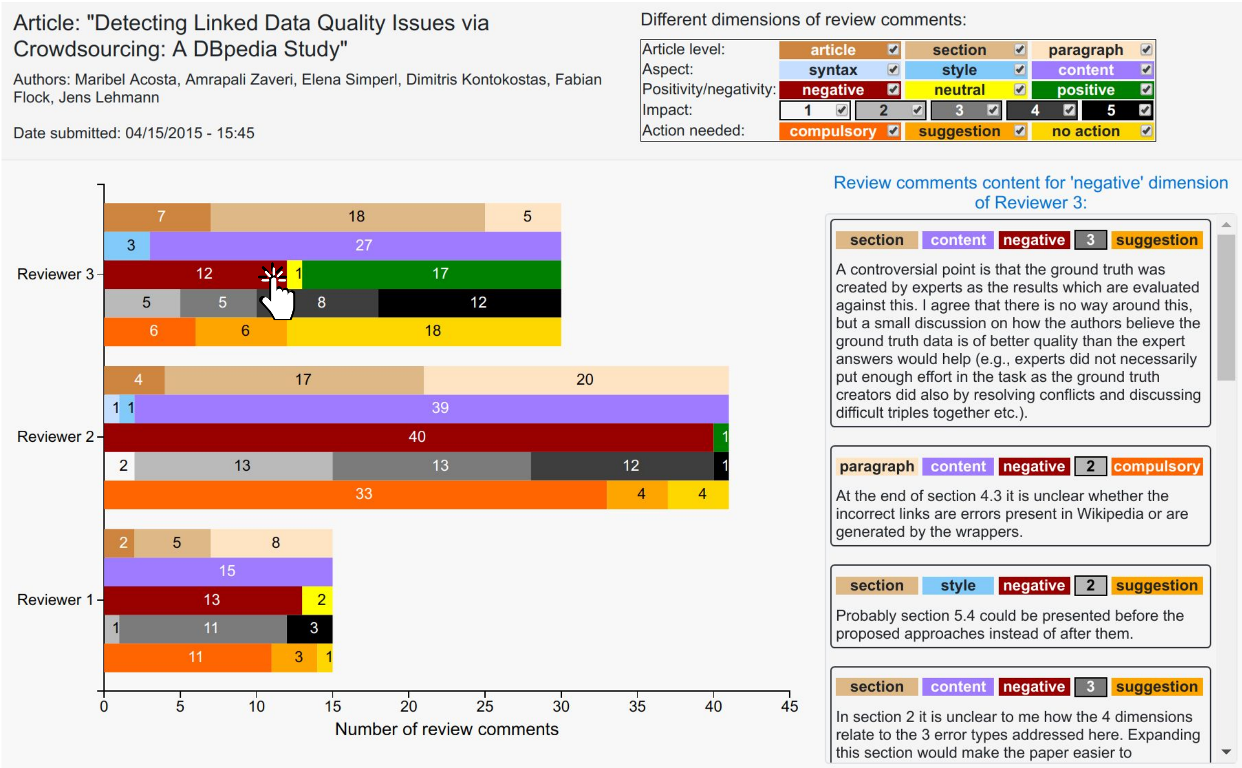


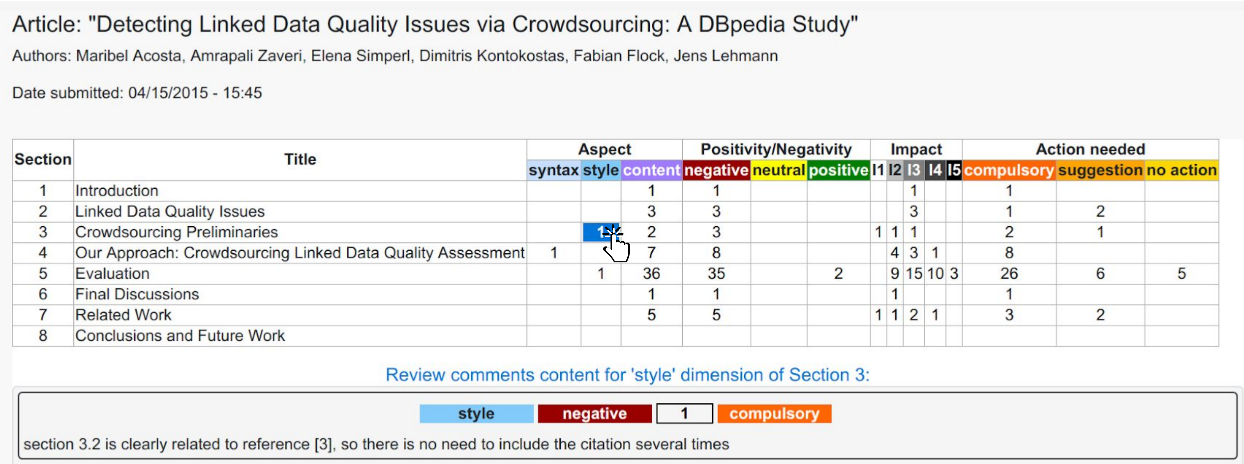
Figure 8: The reviewer-oriented view for the editor study.

number of nanopublications, triples, paragraphs, review comments, and links between them.

As we are dealing with hundreds of individual nanopublications instead of just a hand-full of submission and review files, the performance of downloading them might pose a practical problem. For that reason, we also tested how long it takes to download all 627 nanopublications from the server network, using `nanopub-java` [76] as a command-line tool and giving it only the URI of the index nanopublication. This small download test was performed on a personal computer via a normal home network. For this, we retrieved them all via the library's `get` command and measured the time. We performed this 50 times, in five batches of 10 executions.

Next, we used our dataset to see if we are able to answer the seven competency questions that we defined above, in order to help editors in their meta-reviewing task. With this, we want to find out whether the combination of ontologies and vocabularies we used in our approach is sufficient to cover them, and then whether we can use the SPARQL query language to operationalize them and make them automatically executable on our nanopublication data.

Finally, we perform a user experiment involving editors to find out whether they indeed consider our competency questions important, and how useful they find our interface for getting an answer to these questions. For this study, we created a form that had two parts cor-



Article: "Detecting Linked Data Quality Issues via Crowdsourcing: A DBpedia Study"

Authors: Maribel Acosta, Amrapali Zaveri, Elena Simperl, Dimitris Kontokostas, Fabian Flock, Jens Lehmann

Date submitted: 04/15/2015 - 15:45

Section	Title	Aspect			Positivity/Negativity			Impact					Action needed		
		syntax	style	content	negative	neutral	positive	I1	I2	I3	I4	I5	compulsory	suggestion	no action
1	Introduction			1	1					1			1		
2	Linked Data Quality Issues			3	3					3			1	2	
3	Crowdsourcing Preliminaries		1	2	3			1	1	1			2	1	
4	Our Approach: Crowdsourcing Linked Data Quality Assessment	1		7	8				4	3	1		8		
5	Evaluation		1	36	35		2		9	15	10	3	26	6	5
6	Final Discussions			1	1				1				1		
7	Related Work			5	5			1	1	2	1		3	2	
8	Conclusions and Future Work														

Review comments content for 'style' dimension of Section 3:

style | negative | 1 | compulsory

section 3.2 is clearly related to reference [3], so there is no need to include the citation several times

Figure 9: The section-oriented view for the editor study.

responding to the two parts of the study. We chose an article from our dataset that had a large number of review comments. For the first part, we asked for the importance of the competency questions using a Likert scale (from 1 to 5). For the second part, we provided static screenshots of our tool (the reviewer-oriented or the section-oriented view, depending on the question) together with a link to the live demo and asked about how useful the participants would find such a tool to answer the given competency question. The answers were on the same kind of a Likert scale from 1 to 5. We sent this questionnaire (details online[10]) to a total of 401 editors of journals that support open reviews, specifically Data Science[11], the Semantic Web Journal[12] and PeerJ in Computer Science Journal[13].

3.4.2 *Evaluation Results*

We can now turn to the results of these three parts of our evaluation. Details about the dataset and how it was generated and further queries and results can be found online[14].

3.4.2.1 *Descriptive Analysis*

Our representation of the three papers of our dataset together with their reviews leads to a total of 10 437 triples in 627 nanopublications, 279 text snippets and 213 review comments (85 for article 1, 59 for article 2 and 69 for article 3). Each of the three articles had three reviews: first article - 17, 18 and 50 review comments provided by the

10 https://github.com/LaraHack/linkflows_editor_survey/
11 https://content.iospress.com/journals/data-science
12 http://www.semantic-web-journal.net/
13 https://peerj.com/computer-science
14 https://github.com/LaraHack/linkflows_model_implementation

Table 6: Descriptive statistics dataset.

part of article	number
articles	3
sections	89
paragraphs	279
figures	11
tables	10
formula	8
footnote	2
review comments	213

Table 7: Statistics nanopublications.

	number	average
Nanopublications:	627	
Head triples:	2508	4.00
Assertion triples:	5420	8.64
Provenance triples:	1254	2.00
Publication info triples:	1255	2.00
Total triples:	10 437	16.65

three reviewers, second article - 16, 21, 22 review comments each and third article - 11, 42, 16 review comments each.

In Table 6 some general statistics of the dataset are presented, while Table 7 shows general statistics about the nanopublications corresponding to the three articles and their reviews. Overall, this demonstrates the working of our approach of representing the elements of scientific communication in a fine-grained semantic manner. Of course, more complex analyses are possible, including network analyses of the complex interaction structure, and the queries for the competency questions that we defined above, to which we come back below.

Our small test on the performance of retrieving all nanopublications from the decentralized nanopublication network showed an average download time of 11.66 seconds overall (with a minimum of 8.39 and a maximum of 13.34 seconds). This operation retrieves each of the 627 nanopublications separately and then combines them in a single output file. The time per nanopoublication is thereby just 18.6 milliseconds, which is achieved by executing the request in parallel to several servers in the network at the same time.

3.4.2.2 *Competency Question Execution*

In order to answer the competency questions in Section 3.3.2, we managed to implement each of them as a concrete SPARQL query. We can't go into them here in detail due to space limitations, but the complete queries and all the required data and code can be found online[15].

This shows that our model is indeed able to capture the needed aspects for our competency questions, but we still need to find out whether these competency questions are indeed considered important by the editors, and whether the results from the SPARQL queries allow

15 https://github.com/LaraHack/linkflows_model_implementation/tree/master/queries

us to satisfy these users' information needs. These two aspects are covered in our user study.

3.4.2.3 *User Study Results*

Out of the total 401 questionnaire requests sent, we received a total of 42 answers (10.5%). The importance of the seven competency questions for editors and the usefulness of the interface presented to answer these competency questions, assessed on a Likert scale from 1 to 5 where 1 is *not important at all* and 5 is *very important* can be seen in Table 8. We marked with * the competency questions that had a significant p-value (< 0.05) and without, the ones that were not significant. We calculate significance with a simple binomial test by splitting the responses into the ones that assign at least medium importance or usefulness (≥ 3) and the ones that assign low importance or usefulness (< 3).

We see the respondents declared high importance to five of the seven competency questions in a significant manner with average values from 3.05 to 4.58 (CQ1, CQ3, CQ4, CQ5 and CQ6), while the remaining two (CQ2 and CQ7) were not considered important in the editors' view (average values of 2.36 and 2.79, respectively). Apparently, the number of positive and negative comments per section of the article (CQ2) and how the review comments cover the different parts of the article such as sections (CQ7), seem to have mixed reviews from editors, not being considered significantly important. The critical points that were raised by the reviewers (negative comments with a high impact on the paper) seems to be considered the most important competency question for the editors that responded (CQ5) with an average value of 4.58. Also important, in decreasing order, are the distribution of review comments with respect to whether they address the content or the presentation (syntax and style) of the article (CQ3), the number of points raised to be addressed by authors as an estimate for the amount of work needed for a reviewer (CQ6), the number of positive and negative comments per reviewer (CQ1), and the nature of the review comments with respect to whether they refer to a paragraph or a larger structure such as a section or the whole article (CQ4). For CQ2 and CQ7, we can say that editors did find it on average less important which sections of the article the reviews comments addressed. In general, however, we can conclude that most of competency questions are found to be important by most editors. However, we also observe a quite large standard deviation (SD) as seen in Table 8, ranging from 0.93 to 1.36 on our Likert scale that has a maximum distance of 4.0.

Next, we evaluated the usefulness of our prototype interface. Here the Likert scale went from 1 standing for *not useful at all* to 5 standing for *very useful*. As we can see from Table 8, this interface was on average considered useful for all of the seven competency questions, with averages ranging from 3.21 to 3.83. The preference for scores of 3

Table 8: Results of the user study with editors.

Question	importance of question						usefulness of interface					
	AVG	MED	SD	count <3	count ≥ 3	Δcount p-value	AVG	MED	SD	count <3	count ≥ 3	Δcount p-value
CQ1	3.17	3	1.36	15	27	0.044 *	3.48	4	1.17	9	33	1.36e-4 *
CQ2	2.36	2	1.10	24	18	0.860	3.83	4	1.03	5	37	2.22e-7 *
CQ3	3.64	4	0.93	5	37	1.36e-4 *	3.40	3.5	1.04	9	33	1.47e-3 *
CQ4	3.05	3	1.19	14	28	0.022 *	3.26	3	1.20	14	28	0.022 *
CQ5	4.58	5	0.63	0	42	$<$ e-12 *	3.21	3	1.16	9	33	1.36e-4 *
CQ6	3.57	4	1.02	6	36	1.41e-6 *	3.43	4	1.06	8	34	3.44e-5 *
CQ7	2.79	3	1.12	18	24	0.220	3.62	4	1.03	5	37	2.22e-7 *

or larger is clearly significant for all of them. A substantial minority of respondents, however, didn't find our interface useful leading again to relatively large standard deviation values between 1.06 and 1.19.

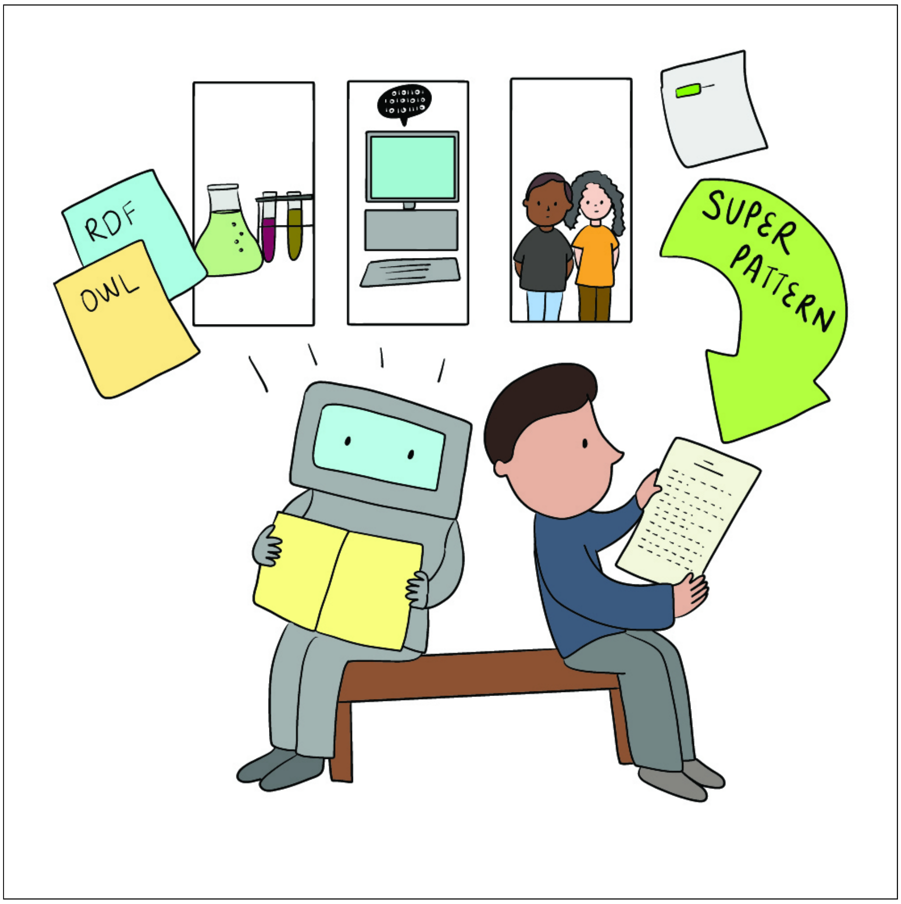
The free-text feedback field at the end of the questionnaire, finally, gave us a variety of suggestions for improvement (some of the editors argued that the interface used too many colors, others suggested other ways of grouping the data) but without clear overall tendencies. These responses also did not hint at any competency questions they found to be missing.

3.5 DISCUSSION AND CONCLUSION

Our results show that we can practically represent the different elements of scientific communication, such as articles and reviews, in a fine-grained and semantic way with nanopublications. We could show that we thereby can automatically answer a wide range of competency questions in the concrete scenario of editors in their meta-reviewing task. We found, however, that some of these were not found to be important, on average, by the editors who participated in our user study. Specifically, the questions about how well the review comments cover the different parts of the paper were not found to be important by a majority of editors. This could indicate that the article structure in terms of its different sections is not a good target for measuring the coverage of reviews. For all the questions, a relatively high variation is observed, which might be hinting at a lack of agreement among editors with respect to how scientific manuscripts should be assessed. This in turn could highlight the importance of more structured and more open reviewing processes. Irrespective of whether the competency questions are important, the majority of editors found our prototype to be useful to answer them, although again with a large variation. With our approach focusing on interoperability and openness, however, it is not necessary to design a single interface that suits everybody, but we could allow editors to choose from several alternatives in the future.

In summary, we could show that nanopublications might be a suitable format not just for scientific findings but also for their reviewing processes. Their open and semantic nature can moreover allow other participants outside of the assigned editor and invited reviewers to contribute with their suggestions and comments, both before and after publication, while all the provenance needed to understand the context of each contribution is recorded. In this way, publication and reviewing as a whole might become more fluid, more inclusive, and more powerful.

4 EXPRESSING HIGH-LEVEL SCIENTIFIC CLAIMS WITH FORMAL SEMANTICS



MAIN CLAIM FORMALIZATION: In the context of all things of type *scientific finding*, things of type *application of the super-pattern can generally* have a relation of type *enables* to things of type *formal representation* that are in the same context (i.e. *scientific finding*).

FORMALIZATION NANOPUBLICATION: (link to nanopublication)[1]

PUBLICATION: This chapter was previously published as *Expressing High-Level Scientific Claims with Formal Semantics* in the Proceedings of the 11th International Conference on Knowledge Capture, 2021, and was co-authored by Tobias Kuhn, Davide Ceolin and Jacco van Ossenbruggen [24].

1 https://w3id.org/np/RA1uQyi8UrLXoCS_oJ1VJa3SenQ8XaR0d2ftymAFDurgk

ABSTRACT

The use of semantic technologies is gaining significant traction in science communication with a wide array of applications in disciplines including the life sciences, computer science, and the social sciences. Languages like RDF, OWL, and other formalisms based on formal logic are applied to make scientific knowledge accessible not only to human readers but also to automated systems. These approaches have mostly focused on the structure of scientific publications themselves, on the used scientific methods and equipment, or on the structure of the used datasets. The core claims or hypotheses of scientific work have only been covered in a shallow manner, such as by linking mentioned entities to established identifiers. In this research, we therefore want to find out whether we can use existing semantic formalisms to fully express the content of high-level scientific claims using formal semantics in a systematic way. Analyzing the main claims from a sample of scientific articles from all disciplines, we find that their semantics are more complex than what a straight-forward application of formalisms like RDF or OWL account for, but we managed to elicit a clear semantic pattern which we call the "super-pattern". We show here how the instantiation of the five slots of this super-pattern leads to a strictly defined statement in higher-order logic. We successfully applied this super-pattern to an enlarged sample of scientific claims. We show that knowledge representation experts, when instructed to independently instantiate the super-pattern with given scientific claims, show a high degree of consistency and convergence given the complexity of the task and the subject. These results therefore open the door on the longer run for allowing researchers to express their high-level scientific findings in a manner they can be automatically interpreted. This in turn will allow for automated consistency checking, question answering, aggregation, and much more.

4.1 INTRODUCTION

Semantic Web technologies have started to be more widely applied in science communication to address, among other things, the accelerating growth of scientific literature [142]. This growth makes it increasingly difficult for individual researchers to follow and be up-to-date with all the current developments in their fields [86]. However, despite semantic technologies being applied in many ways in science, there still is a big gap between the existing (and continuously developing) formal ontologies and the content of scientific articles expressed in natural language that is only human-readable [140]. As such, the content of scientific articles of today is not accessible to algorithms. Methods of semantic interlinking and metadata enrichment of scientific articles exist, especially in life sciences [32, 33, 52, 137], but they are mostly about annotating the existing natural language text in a relatively shallow manner. Other successful applications of semantic technologies focus only on the metadata level, such as provenance representation and versioning of scientific works [33].

Applying semantic technologies to not just annotate or describe (on a metalevel) but actually express scientific knowledge is a much more ambitious and much harder task. While technologies like RDF and OWL grounded in first-order logic are now mature and well-tested, full semantic representation of natural language has remained a task that is too complex for automated approaches. Even the extraction of simpler RDF-based structures from the content of scientific articles typically requires manual curation to achieve sufficient level of quality [37, 52, 136]. In fact, even for knowledge representation experts it is very challenging to fully represent typical high-level scientific findings in formal logic, as no general scheme or procedure exists that is known to apply to findings across specific fields.

In order to make progress on these challenges, we focus here on how general high-level scientific claims can be represented in formal logic. We assume in this work that these representations are manually created, but it could also lay the basis for future research on how to automatically generate these representations from scientific texts (though we assume this to be a very hard problem, as it has been in the past). Concretely, we introduce below our approach to express the content of high-level scientific claims with formal semantics with a semantic template that we call the "super-pattern".

In this research we therefore aim to answer the following research questions:

1. To what extent can our super-pattern be used to formalize the main claims of scientific articles from different disciplines?

2. How reliably can the super-pattern be applied to formalize existing claims by knowledge representation experts?

We envisage that in the longer-term future researchers could themselves express their findings with the super-pattern and thereby make their work directly add to a complex knowledge graph of scientific findings. On that basis, it is easy to imagine how technologies like graph query languages and logic reasoners can be used to find similar research, corroborate scientific claims, spot contradictions, provide aggregations and visualizations, answer questions, and many other kinds of tasks. Among countless other applications, this could save months of literature study for first-year PhD students to get an overview of their fields.

The rest of this chapter is structured as follows. In Section 4.2 we describe the current state of the art regarding the representation of the structure and content of scientific articles. In Section 4.3 we describe our methods and approach with regard to the representation of the content of scientific claims, mainly describing the super-pattern ontology we developed. In Section 4.4 we describe in detail how we performed the evaluation of our approach, while we report and discuss about the design and the results of the formalization study done to evaluate our approach in Sections 4.4.2 and 4.4.3. Future work and conclusion of the present research are outlined in Section 4.5.

4.2 BACKGROUND

Despite being in the digital era, the medium in which the scientific articles are still published are journals, which are now electronically available, but otherwise still follow the old paper paradigm [112, 113, 140]. On top of this, the growing number of articles that are published every day makes it imperative to make the articles machine-readable as well, as human researchers are reaching their limits in terms of how much they are able to process and read every day [86, 127]. With current developments in the Semantic Web, with technologies like RDF, OWL, and SPARQL [140, 142], it is currently possible to enrich the meaning of a traditional article in the digital publishing environment and facilitate its automatic discovery, to have access in a semantic way within the article and link to other related articles or other related parts of articles [67, 102, 131].

Based on these Semantic Web technologies, there are various approaches that try to automatically extract such semantic information from articles to make them more machine-readable, covering the metadata [32] or the article structure [154], providing summarizations of the main concepts and ideas (human assisted and curated) [136], using annotations to semantically enhance articles [33, 137, 154] or semantically interlink them [52], and describing bibliographic references and citations [118]. On the one hand, we have rich vocabularies and ontologies like the SPAR suite [116] that support the classical publishing process starting from formal descriptions of the structure of articles and article discourses [1], to the traditional publishing workflows [51], to accounting for all roles involved in the publishing process [120], and even to the current imperfect peer-reviewing process [130]. On the other hand, there are AI tools that can identify plagiarism and predict the impact an article might have [125], to natural language processing techniques that make use of strategies such as entity extraction to identify concepts and ideas in articles, assisted and curated by humans [136].

Data infrastructures such as the Linked Open Data (LOD) allow for connecting data between scientific publications [144], but this needs additional steps as concepts and relations need to be identified beforehand, and the content of the articles need to be interconnected semantically, usually by using humans in the loop [32, 137]. Moreover, more complex approaches that attempt to automate the process of extracting formal semantics from traditionally-structured scientific articles include techniques like the CISE [115], semantic lenses [40], and modelling the context of sentences from scientific articles in conceptual frameworks [4]. Semantic Web technologies are therefore extensively used but the underlying publishing paradigm has stayed the same [151].

New initiatives that try to change this old paradigm of publishing are proposed, together with new publishing workflows. This paradigm shift entails to move from the textual representation of information to a more data-centric one, similar to the proposal for the next-generation Web [13], where instead of documents, the interest shifts from the syntactic (e.g. HTML) to the semantic level (e.g. RDF, OWL). On the semantic level, we would be able to express the content and not just the structure of what is now in narrative documents. Moreover, formats that are based on HTML, like RASH [117], have been proposed, where scientific articles that include semantic annotations can be represented. Other initiatives involve using semantic representations from the start, written by the actual authors of the research in what is named *genuine semantic publishing* [82], and we have proposed in our earlier work to move from the monolithic form structure of scientific articles to smaller, more granular interconnected parts that each contain single scientific claims or statements from the start [23].

The Open Research Knowledge Graph (ORKG) [70] is another initiative that aims to make research articles machine-readable. With this approach, scientific entities are expressed as a semantically interconnected knowledge graph, populated by methods such as extracting scientific concepts from the abstracts of scientific articles with the help of annotators [18]. In the life sciences, especially in the biomedical fields, there have been many initiatives to create controlled vocabularies that can serve as the foundation to represent scientific knowledge in a structured way in order to capture evidence and scientific findings from research in a computable form [30, 95, 146] with specific markup languages [64] and exchange formats [39]. Some of these approaches target high-level claims, but only with restricted coverage, for example Biological Expression Language (BEL) [145] for specific kinds of biological relations.

Nanopublications [58] are a concept and technology of LD containers that can be used to represent and share different kinds of scientific knowledge. They are expressed in a way that is fully formal, thus machine-interpretable and can be regarded as "minimal publications" due to the fact that they contain just three basic elements (represented in RDF): an assertion containing a small unit of information that is its main content (such as a scientific finding), provenance of the assertion (e.g. linking to the scientific methods used to derive the scientific finding in the assertion) and a publication information part about the nanopublication as a whole (e.g. when it was created and by whom).

Nanopublications can not only be an openly accessible structured data container for scientific findings or claims, but they can also be used for different types of meta-level assertions as well, for example for statements or assessments about other nanopublications [78].

4.3 APPROACH AND METHODS

We present here our approach to formalize high-level scientific claims by proposing what we call a "super-pattern". This super-pattern is a general template of a logical statement that can be instantiated to represent scientific claims in formal logic. To aid the development of this super-pattern, we created a dataset of 50 claims from scientific publications (Dataset A). After the super-pattern had been finalized, we created a second set of another 25 claims (Dataset B), which did not influence our design decisions of the super-pattern and is therefore not biased towards it.

4.3.1 *Super-Pattern*

The super-pattern is a template with five slots, which are to be filled in when it is instantiated. It has the following structure with five elements with expected types in square brackets:

- **Context class** ("in the context of all ..."): [class identifier]
- **Subject class** ("things of type ..."): [class identifier]
- **Qualifier** (e.g. "mostly"): [qualifier from closed list]
- **Relation type** ("have a relation of type ..."): [relation type from closed list]
- **Object class** ("to things of type ... that are in the same context"): [class identifier]

The context class is optional, whereas all other slots are mandatory. The phrases in quotes above give some hints towards how to interpret it, but we will provide a thorough formal definition below. But let us first have a look at a concrete example.

In Figure 10 we see an example where the context class is named "person", the subject class is named "obesity together with metabolic abnormality", the qualifier is "frequently", the relation is "co-occurs with" and the object class has the name "knee osteoarthritis". By just filling in the informal phrases above, with a little bit of editing we get the following sentence, hinting at how to interpret it:

> In the context of all *persons*, things of type *obesity together with metabolic abnormality frequently* have a relation of type *co-occurs with* to things of type *knee osteoarthritis* that are in the same context (i.e. the same *person*).

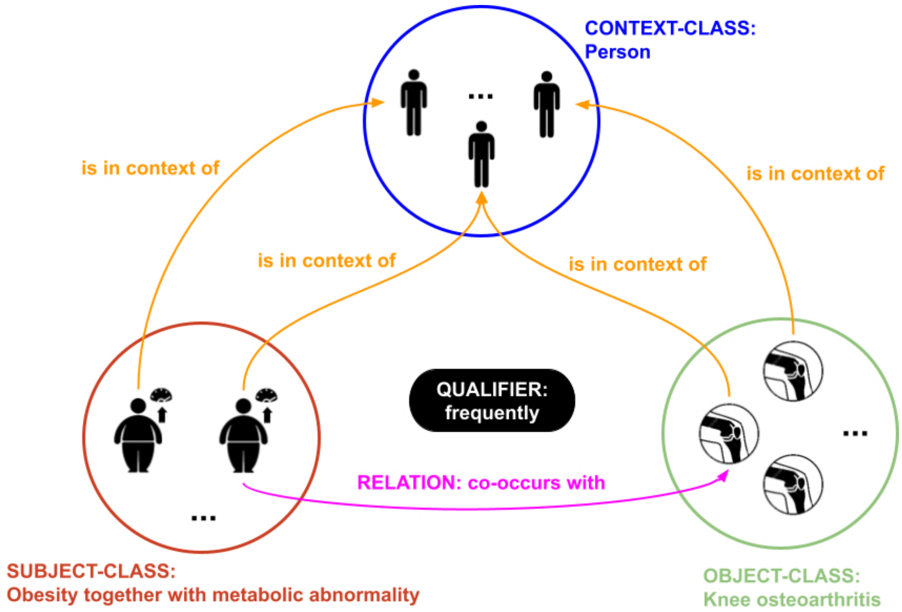


Figure 10: Diagram with an example for the Super-pattern instantiation.

It is easy to see that this is saying that when people have obesity together with metabolic abnormality then these people frequently also have a thing called knee osteoarthritis.

To arrive at a more formal representation grounded in logic, we can start by looking at the qualifiers like "generally". Our super-pattern defines the qualifiers shown in Table 9. There are the five base qualifiers *always*, *generally*, *mostly*, *frequently*, and *sometimes*, each with its negative counterpart, such as *never* for *always* and *frequently not* for *frequently*. On top of that, each of these ten qualifiers can be amended by the operator *can*, such as *can generally*, which will be interpreted in a modal logic manner. The table shows how each qualifier is mapped to a quantified interpretation, such as *at most 90%* for *frequently not*. These interpretations are necessarily a bit arbitrary, but to reach our goal of fully formal semantics we had to define a precise value for these otherwise vague terms.

For our formalization to be described below, each of the 20 qualifiers is given a relation predicate $\lesseqqgtr$, a probability value q, and a modality value m. The values for $\lesseqqgtr$ and q are shown in Table 9. The modality value m is $\Diamond$ for the qualifiers with *can* and the empty symbol ϵ otherwise.

We define the semantics of an instantiated super-pattern by the following logic expression template:

$$P(\ m(\exists z(o(z) \wedge i(z,x) \wedge r(y,z)))\ |\ s(y) \wedge c(x) \wedge i(y,x)\ ) \lesseqqgtr q$$

Here, c, s, and o stand for the context, subject, and object classes, respectively, of the instantiated super-pattern. Them being classes they map to unary predicates in the expression above. The super-pattern

Table 9: Super-pattern qualifiers.

qualifier	**interpretation**	$\lesseqgtr$	**q**
(can) always	100%	$=$	1
(can) generally	at least 90%	$\geq$	0.9
(can) mostly	at least 50%	$\geq$	0.5
(can) frequently	at least 10%	$\geq$	0.1
(can) sometimes	at least 0.1%	$\geq$	0.001
(can) never	0%	$=$	0
(can) generally not	at most 10%	$\leq$	0.1
(can) mostly not	at most 50%	$\leq$	0.5
(can) frequently not	at most 90%	$\leq$	0.9
(can) sometimes not	at most 99.9%	$\leq$	0.999

qualifier is expressed with the three parameters $\lesseqgtr$, q, and m, as explained above. r, finally, represents the relation of the instantiated super-pattern, mapped to a binary predicate. This relation comes from a fixed list of possible relations, consisting of sameness, numerical comparison, causality, and spatio-temporality relations. The full list of relations can be found in the online super-pattern ontology documentation[2] and in Table 11.

The remaining symbols are interpreted as follows. The symbol i stands for the general relation "is in the context of" that maps instances in the subject and object classes to instances of the context class.[3] $P(\ldots|\ldots)$ represents the conditional probability function of standard statistics, but with the slightly non-standard convention of having logical expressions as elements. They are interpreted as the sets of tuples of all contained free variables (here, x and y) that satisfy the expression. z, finally, is a bound variable in the first of these expressions with the normal first-order semantics.

We can also phrase the $P(\ldots|\ldots)$ part as a matter of ratios instead of probabilities: It is the ratio of things that satisfy the left-hand side out of all the things that satisfy the right-hand side. This can be defined as such in higher-order logic, but we chose here the notationally simpler expression with conditional probabilities.

To get a better understanding of how the super-pattern works, let us walk through two cases. First, we can have a look at the effect of the *always* qualifier, for which we set $\lesseqgtr$ to $=$, q to 1, and m to ϵ. The above expression then becomes (omitting the empty symbol ϵ, as it has no effect by definition):

$$P(\ (\exists z(o(z) \wedge i(z,x) \wedge r(y,z)))\ |\ s(y) \wedge c(x) \wedge i(y,x)\ ) = 1$$

2 https://larahack.github.io/linkflows_superpattern/doc/sp/index-en.html

3 A future version of the super-patterns might also make this relation configurable with more specific relations when the super-pattern is instantiated.

This stating that the probability is 1, this is equivalent to saying that the right-hand side part of the conditional probability expression implies the left-hand side:

$$\exists x \exists y(\ s(y) \wedge c(x) \wedge i(y,x) \rightarrow \exists z(o(z) \wedge i(z,x) \wedge r(y,z))\ )$$

We see that this is expressing that for all things y in class s that are in the context of a thing x in class c, there exists another thing z in class o that is in the context of the same thing x such that there is a relation of type r from y to z, which matches how we introduced the super-pattern above.

As a second case, let us pick the more complex case of the qualifier *can generally not*, corresponding to the parameter values $\lesseqgtr = \leq$, $q = 0.1$, and $m = \Diamond$. This gives us:

$$P(\ \Diamond(\exists z(o(z) \wedge i(z,x) \wedge r(y,z)))\ |\ s(y) \wedge c(x) \wedge i(y,x)\ ) \leq 0.1$$

This is stating that given a thing y of type s that is in the context of a thing x of type c, the probability (or "ratio") that it is possible that there exists a thing z of type o in the context of the same thing x such that there is a relation of type r from y to z is less or equal 0.1. Or in other words, thing z as specified is possible under the given conditions in at most 10% of the cases.

The "it is possible that" here corresponds to the "can" in the qualifier and to the logical operator $\Diamond$. It is interpreted according to standard modal logic with possible world semantics with the meaning of "there is an accessible possible world where the following is true". As a more intuitive variation, we can also say "it can be made true that". As we will see below, we often use these "can" qualifiers for statements from applied sciences like computer science or medicine, where findings are often stating that a certain kind of treatment or kind of software *can* solve a particular problem, but only if it is actually applied to it. As an extreme example, there might be strong evidence that a novel kind of medical treatment can treat a certain kind of condition, without a single instance of that actual condition having been treated in this way to that point in time. In that situation we need to express the possibility that it could treat these conditions without necessarily treating them in reality.

For the case when the optional context slot is omitted, we can define a *universal context* class, which contains just one element that by definition has everything in its context. Leaving the context slot empty thereby corresponds to setting it to the universal context. In the expression above, x therefore corresponds to the universal context instance and the $i(\ldots,\ldots)$ components are always true and can be omitted. We therefore arrive at this simpler expression without context:

$$P(\ m(\exists z(o(z) \wedge r(y,z)))\ |\ s(y)\ ) \lesseqgtr q$$

4.3.2 *Dataset A*

In order to see if the content of high-level scientific claims from different disciplines can be expressed with our super-pattern, we created a set of randomly selected scientific articles from Semantic Scholar. All data can be found online[4]. The general methodology for creating this Dataset A was as follows: (1) selecting a random sample of articles from different disciplines; (2) identifying a high-level scientific claim from each article; (3) applying the super-pattern on this claim with informal classes; and (4) formalizing the classes by finding existing identifiers or defining new ones.

STEP 1: SELECT ARTICLES. We selected a set of 50 random scientific articles by creating random numbers and then consulting Semantic Scholar[5] to retrieve the article with that number as Semantic Scholar corpus ID, if it exists. We then checked whether it matched our inclusion criteria of being written in English, published in 2000 or later, having an abstract, and being a research paper. We repeated this process until we collected 50 such matching articles.

STEP 2: IDENTIFY MAIN CLAIM. We check the abstract to find the main scientific claim, finding, or conclusion, e.g. "In a cross-sectional study in Korean women, obesity showed closest association with knee osteoarthritis when accompanied by metabolic abnormality". When several candidate claims existed, we chose the most high-level one, or if that did not resolve it, we simply chose the one mentioned first. We discarded articles that failed to clearly mention their main claim in the abstract. To arrive at a clear understanding of what exactly we want to formalize, we rephrased the found claims as AIDA sentences [77, 78]. These are single English sentences that are Atomic, Independent, Declarative, and Absolute. These properties make them a good tool to delineate what should go into the formalization to be created. For the claim above, the corresponding AIDA sentence could be "Obesity when accompanied by metabolic abnormality is closely associated with knee osteoarthritis".

STEP 3: APPLY THE SUPER-PATTERN. Next, we tried to apply the super-pattern by figuring out what classes would have to go into the three class slots, and which qualifier and relation would be most suitable. At this step, we do not yet check whether the needed classes are defined by any of the existing ontologies, but we simply refer to them by their (likely) class name. For the example given above, we would arrive at something similar to what is shown in Figure 10.

4 https://github.com/LaraHack/linkflows_claims_dataset
5 https://www.semanticscholar.org/

STEP 4: FIND IDENTIFIERS IN EXISTING ONTOLOGIES. Next we tried to find identifiers in existing ontologies for the classes needed for the super-pattern instantiation from the last step. For this, we used (1) Wikidata[6], the free knowledge base in which new concepts and properties can be added in a structured way, (2) BioOntology[7], one of the biggest repositories containing biomedical ontologies and (3) Linked Open Vocabularies (LOV)[8], a big curated data collection of vocabularies that are reviewed and added continuously. All these chosen sources contain indexed vocabularies and ontologies and as such permit full text searches of concepts and relations. If we found multiple applicable candidate identifiers, we selected the one that seemed most suitable or best documented. For the cases where we could not find a matching identifier, we tried to find identifiers for parts of the required concept. These parts could then be used to construct a (partial) class definition, such as "obesity together with metabolic abnormality", which can be defined as the intersection of the conditions "obesity" and "metabolic abnormality". We mint new identifiers for such complex classes and also for the cases where we could not find any existing identifiers at all.

The process described above was itself the result of an iterative process during which the super-pattern was defined in its current form.

4.3.3 *Dataset B*

As the super-pattern was developed based on the claims found during the generation of Dataset A, we created an additional dataset that did not influence the super-pattern and can therefore provide a more reliable picture about what the range of scientific claims the super-pattern is able to express. For this Dataset B, we selected an extra set of 25 claims from another set of 25 randomly selected articles from Semantic Scholar, following the same procedure and criteria as for Dataset A above. We therefore arrived at another 25 formalizations in the form of instantiated super-patterns (to the extent it applied; more on this later).

In contrast to Dataset A, Step 3 for Dataset B was performed several times independently by different knowledge representation experts. Only after the independent super-pattern instances were created, the experts discussed and tried to reach an agreement. This will be explained in more detail below.

---

6 https://www.wikidata.org
7 https://bioportal.bioontology.org/
8 https://lov.linkeddata.es/

Table 10: Usage of qualifiers in the dataset.

qualifier	sometimes	frequently	mostly	generally	always
positive	4	3	3	29	4
negative			1	5	
can positive	1			17	1
can negative					

4.4 EVALUATION

In this section we will present the design and results of the evaluation for our approach. We performed a descriptive analysis and a vocabulary use analysis for the datasets introduced above. In order to assess how well and how consistently the super-pattern can be applied by knowledge representation experts, we ran a *formalization study* where such experts independently perform super-pattern based formalizations.

4.4.1 *Descriptive Analysis*

We performed a descriptive analysis on the combined datasets used in this research, dataset A and dataset B, one with 50 claims and the other with 25 other claims respectively.

91% of the claims (68 out of the total 75 claims) could be expressed with the super-pattern. The remaining ones were rather *too simple* and not too complex to be expressed, as they could be formalized with a simple *subject–predicate–object* triple in RDF style. The super-pattern was therefore successful in covering all the non-trivial scientific claims encountered. 76% of the claims (57 out of the total 75 claims) used the optional context slot of the super-pattern, whereas the remaining ones 15% (11 claims) used the simpler version of the pattern without context.

Next we can look at the distribution of qualifiers and relations for both datasets, which is shown in Tables 10 and 11. As we can see from Table 10, the most used qualifier is "generally" in almost 39% of cases (29 claims), together with its modal counterpart, "can generally" in 23% of cases (17 claims). The modal negative of qualifiers was never used, while using the negative for qualifiers seems to be less common, in just 6% of cases (6 claims), while the most used qualifiers are positive with 57% (43 claims) and modal positive with 25% (19 claims).

In terms of the relations used in the datasets, Table 11 shows that relations that express causal relations are the most common with 57% (43 claims), then the equivalency relation "is same as" is the next most used with 16% (12 claims), then in a smaller ratio, the relations

Table 11: Usage of relations in the dataset.

relation	**count**	**total per group**
is same as	12	12
compares to	0	7
has similar value as	0	
has same value as	0	
has different value from	0	
has smaller value than	1	
has larger value than	6	
has causal relationship with	0	43
affects	2	
contributes to	11	
enables	6	
inhibits	2	
prevents	1	
increases	4	
decreases	3	
requires	4	
causes	5	
is necessary and sufficient for	2	
is caused by	3	
has spatio-temporal relationship with	0	6
includes	5	
is included in	0	
co-occurs with	1	
is followed by	0	
follows	0	

marking numerical comparisons (the "compares to" relations) and the relations about spatio-temporal relationships (the "has spatio-temporal relationship with") are used in about 6-7% of cases (7 and 6 claims, respectively).

Overall, we see a broad but far from uniform distribution. Certain qualifiers and relation types were much more often used than others, which is in general not very surprising. That most claims are of a positive and causal nature is also consistent with expectations.

For Dataset A we can do an analysis of the classes used from the existing ontologies. We only restrict ourselves here to the top-level classes that could directly be used in the slots of the super-pattern, and we exclude here the classes and relations from existing ontologies that can be used to build a complex class definition.

In terms of the vocabulary usage for the classes for the subject, context, and object slots of the super-pattern, we noticed that most of these classes were not present in existing vocabularies, but had to be newly minted. In only 18 out of the 128 classes (14%), we could directly use an existing class identifier. Most of the time (16 cases), these identifiers came from Wikidata. Therefore we can say that formalizations that use the super-pattern mostly depend on defining new classes, which can typically only be partially defined from existing concepts and relations.

4.4.2 *Design of Formalization Study*

In order to evaluate the proposed method and our super-pattern, we designed a formalization study where several knowledge representation experts independently apply the super-pattern. The goal of this study was to find out how reliably a super-pattern could be created from a given scientific claim. The extent to which users who are not knowledge representation experts would be able to create such formalizations is beyond the scope of this work.

We have therefore designed a three-stage formalization study where the four co-authors of this chapter participated as knowledge representation experts. In the first stage, the experts independently instantiate the super-pattern for the given claims. In the second stage, each expert is asked to review all four formalizations (which are anonymized and randomly shuffled) from the first stage and select the best ones. In the last stage, all experts meet, discuss their choices, and are given the opportunity to adjust their selection. All data can be found online[9].

STAGE 1. The first stage started with an introduction session where the participating knowledge representation experts learned about the details of the super-pattern and could discuss any questions. The formalizations of Dataset A were used as examples for all the experts to have a common and consistent understanding of the super-pattern. After this joint introduction session, all experts worked independently to apply the super-pattern on the 25 claims of Dataset B. For each of these claims, the participants also had to rate (on a scale from 1 to 5) how confident they were in their formalization. For the class slots, the experts were only asked to provide class names, but not existing class identifiers.

STAGE 2. In this stage, the participants were asked to independently inspect and review all four formalizations (i.e. the own one as well as the ones from the other participants). These formalizations were randomly shuffled and it was not visible who created which of them. The participants were asked to select the best formalization (in their

9 https://github.com/LaraHack/linkflows_formalization_study

Table 12: Levels of agreement in the formalization study.

Agreement level: before and after discussion	before (Stage 2)		after (Stage 3)	
	abs.	rel.	abs.	rel.
Level A: Full agreement on best formalization	2	0.08	21	0.84
Level B: Majority agreement on best formalization	12	0.48	3	0.12
Level C: Full agreement on absence of formalization mistakes	9	0.36	1	0.04
Level D: No agreement of the types above	2	0.08	0	0.00

opinion) and to indicate which formalizations in their view contained clear formalization mistakes. If several equally good formalizations were present, participants were allowed to select more than one "best" formalization. After this stage, we can assess how much the participants agreed on the best formalization, and how much they agreed on the presence of formalization mistakes. As the task was complex and multiple valid solutions are possible, we can expect a substantial level of disagreement at this stage.

STAGE 3. The goal of this last part of the formalization study was to see the extent to which disagreement was resolved when the experts were allowed to directly discuss the cases. In several discussion sessions, the experts went through all 25 claims and their candidate formalizations. The experts could explain why they chose a particular formalization as the best one and what kind of mistakes they identified. During this discussion, the experts could update their previous decisions, e.g. by choosing a different formalization as the best one, or by recognizing a formalization mistake they previously missed. After this stage, we can assess the degree of agreement when the experts have the chance to consult the others' opinion and have the chance to react on that. This agreement can however also include the effect of social dynamics, such as participants trying to achieve (or prevent) consensus for social rather than conceptual reasons.

4.4.3 *Formalization Study Results*

For each of the 25 claims from Dataset B we have four different possible formalizations, therefore 100 formalizations in total. After Stage 2 (before the joint discussion), 38% of the individual formalizations were marked as containing a mistake by at least one of the experts. This dropped to just 6% after the discussion, showing agreement being increased by the discussion. Given the complexity of the task, these mistake ratios seem reasonably low. Overall, just 2% of the individual formalizations had both, at least one mistake mark as well as at least one best mark (this value didn't change after discussion).

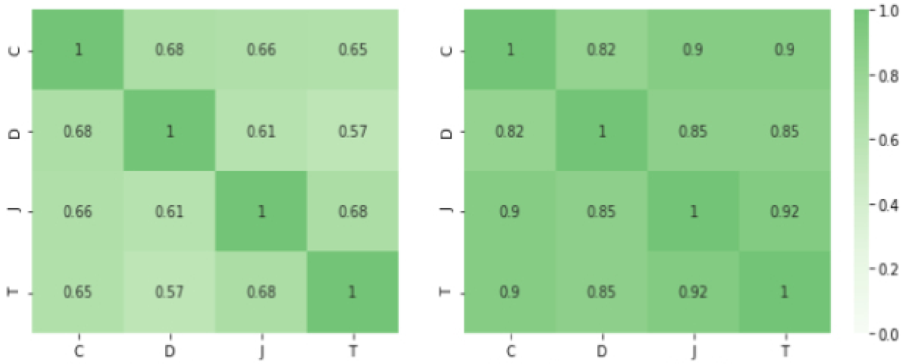
Table 12 summarizes the most important results from the Stages 2 and 3 of the formalization study. In order to analyze the level of agreement between the participants before and after the discussion meeting in Stage 3, we devised an agreement score of four levels A to D. Based on the participants' best and mistake marks given, each of the 25 claims is assigned a level from A (most agreement) to D (least agreement).

In Level A, we include all scientific claims with a formalization that everybody agreed it was the best (or one of several "best"). In Level B, there is no full agreement but a majority agreement of three out of four participants on the best formalization, with no mistake mark from the remaining participant. Level C includes all scientific claims that are not in levels A or B, but for which there exists at least one formalization for which all participants agree that it contains no mistakes. Level D, finally, applies to all claims that are not in any of the categories above.

We see in Table 12 that full agreement on the best formalization was rare after Stage 2, with only 8% (2 claims of 25). The slightly less restrictive majority agreement of Level B, however, was achieved in 48% cases, excluding the perfect agreement cases above. Therefore, in a majority of 56% of cases full or majority agreement on the best formalization was achieved, which seems quite remarkable given that this happened without discussion. Agreement on a formalization without mistakes (Level C) was achieved in another 36% of the cases, leaving only 8% to the lowest level D of no agreement at all. Again, considering the complexity and non-deterministic nature of the task, these seem very favorable outcomes that show that the super-pattern can be applied by knowledge representation experts in a reliable and consistent manner.

Looking at the agreement levels after the discussion meeting (Stage 3), we can observe that for most of the formalizations of the scientific claims, full agreement (84%) or majority agreement (12%) is reached. In all cases, there is at least a formalization on which all participants agree that it does not contain any mistakes. As expected, such a discussion session drives up agreement values substantially, which underlines the importance of checking the independent assessments of Stage 2 first. The results show that knowledge representation experts can agree on the best formalization in the vast majority of cases.

The individual pair-wise agreements between the four experts are visualized in Figure 11. To calculate these agreements, we use a value of 1 when the two participants fully agree on the status of an individual formalization (either "best", "mistake", or no mark), a value of 0.5 when they are different but without conflict (i.e. not "best" and "mistake" at the same time), and a value of 0 in the case of conflicting "best" and "mistake" marks. We then take the average of these values



a. Initial agreement b. Agreement after discussion

Figure 11: Agreement between participants on the formalizations before and after discussion.

of the 100 data points per expert pair (4 formalization candidates for each of 25 claims).

We see that the average values before the discussion (Stage 2) are all above the 0.5 threshold (0.57 being the lowest value), and move quite close to full agreement after the discussion in Stage 3 (all values being over 0.82). Also, these pair-wise disagreement values are quite uniform across the pairs, and disagreement was therefore not driven by just one of the experts. Before as well as after discussion, the uniformity of the disagreement can be seen as an indication that it was rather due to the difficulty of the task and not to fundamental misunderstandings by some of the experts.

These pair-wise results on the individual formalization candidates therefore confirm the more general picture of Table 12. In the future, following the discussion in Stage 3 of the study, minor changes were proposed to the Super-Pattern ontology. As such, a future version might contain small changes to the definitions of some relations where, for example, the definition for the "includes" relation, will be "spatio-temporally or conceptually includes" instead of just "spatio-temporally includes". In summary, we can conclude that writing a formalization that expresses the content of a high-level scientific claim using the super-pattern is an achievable task with a reasonably high success rate.

4.5 DISCUSSION AND CONCLUSION

While Semantic technologies have been extensively applied to a broad range of applications in various science disciplines, their use is quite limited with regards to science communication. The rapid increase in the number of scientific articles and the outlook to be able to communicate, link, and validate scientific discoveries in an automated manner should make it a priority to allow for the content of scientific articles to become readable not only by humans, but by machines as well. However, the use of these semantic technologies have mainly targeted the structure of the scientific publications themselves and not the content of the core scientific discoveries of these articles. In this research, we managed to show how we can create a framework in which we can represent these scientific claims with formal semantics.

Taking a random sample of scientific claims extracted from scientific articles from all disciplines, we succeeded to create formalizations of these scientific claims by using our super-pattern approach. It appears that this approach works for scientific claims from a variety of disciplines, confirming the cross-disciplinary nature and applicability of the super-pattern. Most of the classes used in the super-pattern formalizations had to be minted and could not be fully defined from existing vocabularies, but we argue that we managed to solve the more difficult part of the problem, as expressing the high-level logical structure of scientific claims is a more complex task than defining new classes. We could also demonstrate that knowledge representation experts can create such formalizations using the super-pattern in a fairly consistent and reliable manner, despite the inherent complexity and difficulty of this task.

In future work, we plan to investigate how our formal super-pattern representations can be used in practice to let researchers themselves, who are not necessarily knowledge representation experts, publish their own scientific claims. Furthermore, it will be interesting to investigate how advanced reasoning can be applied, considering that super-pattern formalizations use higher-order logic with its theoretical and practical problems around efficiency and decidability. In any case, any kind of reasoning with partial results or incomplete heuristics would constitute a huge advance compared to what we can currently do in terms of reasoning on scientific knowledge. We can then imagine countless novel applications, such as overarching aggregations, finding supporting or conflicting claims, answering high-level questions, making visualization of scientific knowledge, and much more. In summary, it would allow us to harness computers in a more effective way to increase our understanding of scientific discoveries and thereby amplify their impact.

5 NANOPUBLICATION-BASED SEMANTIC PUBLISHING AND REVIEWING

MAIN CLAIM FORMALIZATION: In the context of all things of type *scholarly submission and reviewing workflow*, things of type *application of nanopublications can generally* have a relation of type *enables* to things of type *fully machine-interpretable result and review* that are in the same context (i.e. *scholarly submission and reviewing workflow*).

FORMALIZATION NANOPUBLICATION: (link to nanopublication)[1]

PUBLICATION: This chapter was previously published as *Nanopublication-based Semantic Publishing and Reviewing: A Field Study with Formalization Papers* in the PeerJ in Computer Science Journal, 2023, and was co-authored by Tobias Kuhn, Davide Ceolin and Jacco van Ossenbruggen [25].

1 https://w3id.org/np/RAYoO6zK9BW82YlPq4mIMD99BZ48igGFlHx9vxz_FQmqM

ABSTRACT

With the rapidly increasing amount of scientific literature, it is getting continuously more difficult for researchers in different disciplines to keep up-to-date with the recent findings in their field of study. Processing scientific articles in an automated fashion has been proposed as a solution to this problem, but the accuracy of such processing remains very poor for extraction tasks beyond the most basic ones (like locating and identifying entities and simple classification based on predefined categories). Few approaches have tried to change how we publish scientific results in the first place, such as by making articles machine-interpretable by expressing them with formal semantics from the start. In the work presented here, we propose a first step in this direction by setting out to demonstrate that we can formally publish high-level scientific claims in formal logic, and publish the results in a special issue of an existing journal. We use the concept and technology of nanopublications for this endeavor, and represent not just the submissions and final papers in this RDF-based format, but also the whole process in between, including reviews, responses, and decisions. We do this by performing a field study with what we call formalization papers, which contribute a novel formalization of a previously published claim. We received 15 submissions from 18 authors, who then went through the whole publication process leading to the publication of their contributions in the special issue. Our evaluation shows the technical and practical feasibility of our approach. The participating authors mostly showed high levels of interest and confidence, and mostly experienced the process as not very difficult, despite the technical nature of the current user interfaces. We believe that these results indicate that it is possible to publish scientific results from different fields with machine-interpretable semantics from the start, which in turn opens countless possibilities to radically improve in the future the effectiveness and efficiency of the scientific endeavor as a whole.

5.1 INTRODUCTION

Considering the abundance of scientific articles that are published every day [157], keeping up with the latest research is becoming a significant challenge for researchers in many fields. This is at least partially due to the fact that we are still holding on to an archaic paradigm of scientific publishing: the canonical way to publish scientific results is by writing them up in long English texts called articles, which are in the best case easy to read by human experts but remain mostly inaccessible to automated approaches (except on a very superficial level with text mining approaches) [15, 143, 164]. These articles then undergo peer reviewing, which is typically done in a way that is secretive and not standardized, with the effect that the reviewing process may lack transparency and that valuable comments from the reviewers cannot be reused or built upon. There have been studies on the effectiveness of peer-reviewing in its current form [75, 91, 147] that showed not only systematic biases among peer-reviewers, but also a lack of transparency in the general peer-reviewing process as a whole [12, 89, 148]. Making reviews open might alleviate some of these concerns by ensuring higher-quality reviews, while at the same time increasing the trust in the reviewing process and the quality of the scientific publications themselves.

A range of approaches have been proposed to address some of these problems by making scientific texts machine-readable, allowing for automatic summarising, finding and retrieving information easier and even the ability to (partially) reason on the scientific texts themselves. Text mining approaches work reasonably well when it comes to simple entity extraction with techniques like named-entity recognition to extract the main concepts from a text (e.g. [104, 167]), but accuracy dramatically drops with more complicated tasks like relation extraction or identifying links between entities [45, 166, 170].

The vast majority of existing approaches of making scientific texts machine-readable have one thing in common: they take the current paradigm of scientific articles for granted and therefore take them as their starting point to extract information. While it is important to try to process the vast amount of existing scientific literature that has the form of long English texts (and sometimes long texts in other languages), we should also think about how we can improve the way how we publish scientific insights in the first place. An important aspect of this is the vision of semantic publishing, which we mean here in the sense of *genuine semantic publishing* [82], where the machine-interpretable formal semantics cover the main scientific claims the work is making. Nanopublications [58], which are small RDF-based semantic packages, have emerged as a powerful concept and technology for enabling such genuine semantic publishing. We should note here that we are using the term "semantic" in its more narrow sense of "with meaning represented in a formal computer-interpretable manner", and not in the more general meaning of "with respect to meaning".

In previous research we have applied nanopublications to implement a semantic and fine-grained model for reviewing [22], and have extended this to semantically represent the full structure of (classical) scientific articles with their reviews and review responses as a single network of nanopublications [23]. In order to get closer to our vision of genuine semantic publishing, however, we need to represent not just the structure but also the main content of these articles, most importantly their main scientific claims. To that aim, we proposed in subsequent work the *super-pattern*, a semantic template to represent the meaning of scientific claims in formal logic [24].

Taking an example from our previous study as an illustration of the super-pattern, it has been stated in scientific literature [48] that in particular kinds of cells in the rat brain (specifically, endothelial cells) some sort of stress called transient oxidative stress affects the expression of a protein called Pgp. The super-pattern consists of five slots that would in this example be filled in as follows:

- Context class: rat brain endothelial cell
- Subject class: transient oxidative stress
- Qualifier: generally
- Relation: affects
- Object class: Pgp expression

Informally, we can read this in the following way: whenever there is an instance of transient oxidative stress in the context of an instance of a rat brain endothelial cell, then generally (meaning in at least 90% of the cases), that instance of stress has the relation of affecting an instance of Pgp expression. Formally, it directly maps to this logic formula:

$$P(\ \exists z(\textsf{pgp-expression}(z) \land \textsf{in-context}(z,x) \land \textsf{affects}(y,z)) \mid \\ \textsf{transient-oxidative-stress}(y) \land \textsf{rat-brain-endothelial-cell}(x) \land \\ \textsf{in-context}(y,x)\ ) \geq 0.9$$

This is stating in logic terms (in slightly non-standard notation using conditional probability as a shorthand) that given a thing y of type *transient-oxidative-stress* in the context of a thing x of type *rat-brain-endothelial-cell*, the probability of there being a z of type *pgp-expression* that is in the same context x is at least 90%. We have shown that this pattern can be applied to formalize most high-level claims found in scientific literature across disciplines [24].

The representation above is in a certain way more precise than what the article was stating, by making the percentage of 90% explicit. For this example, we made a best guess, but ideally this decision should of course come from the original researchers. In another sense, the statement is probably less specific than what can be read in the paper in certain other respects, as any formal representation trades to a certain extent nuance and detail for precision. As a further side remark, it is important to realize that the above number of 90% is part of what the statement expresses (namely the minimum ratio of how often the *affects*-relation holds in the given condition), and does *not* stand for the certainty or validity of the statement as a whole. Expressing the (un)certainty of the statement can be done with models such as Ontology of Reasoning, Certainty and Attribution (ORCA) [161], but this is outside of the scope for the super-pattern.

In the work to be presented below, the main research goal is to combine all the elements we have previously worked on— namely semantic representation of reviews [22], scientific works as networks of nanopublications [23], and representing the main claims with the super-pattern [24] — in order to implement genuine semantic publishing and putting it to the test in a field study. Whereas our current scientific publishing process works with narrative-based, natural language texts in the form of scientific articles that are later made more machine-interpretable by means of semantic annotations and semantic interlinking to enhance their semantic integration and discovery, we propose a different approach, one that considers semantics from the beginning. Therefore, the main aim of this research is to make a first step in the direction of publishing with formal semantics from the start, showing that it is possible to represent *semantically* not only scientific claims, but also the entire scientific publishing process without going through other intermediary semantic processing stages. For practical reasons, we did not require the scientific claims in this field study to be novel ones, but they were selected from existing publications. This field study led to the publication of a special issue in an established journal (Data Science[2]) at an established publisher (IOS Press[3]). This special issue consists of what we call *formalization papers*, which are nanopublication-based semantic publications whose novelty lies in the formalization of a previously published scientific claim.

In this research we therefore aim to answer the following research question:

- Can nanopublications and the super-pattern enable a new paradigm of scientific communication where authors publish their scientific findings with formal semantics from the start?

The rest of this chapter is structured as follows. In Section 5.2 we describe the current state of the art in the field of scientific publishing with regard to scientific knowledge representation, semantic publishing and semantic articles and also alternative proposed machine-readable approaches like nanopublications. In Section 5.3 we describe our approach with regard to a new way of publishing, starting from a formal way of representing the content of scientific claims and ending with the representation of the publication process itself in what we call "formalization papers". We then report and discuss the results of the field study we performed using formalization papers in Section 5.4. Future work and conclusion of the present research are outlined in Section 5.5.

2 https://www.iospress.com/catalog/journals/data-science
3 https://www.iospress.com

5.2 BACKGROUND

We provide here the background on scientific knowledge representation, scientific publishing, nanopublications, and genuine semantic publishing in particular.

5.2.1 *Scientific knowledge representation*

Novel proposals for the current "Disruption Era" [124] include scientific publication management models that connect abstract knowledge with actual world problems in the constantly growing body of scientific knowledge [29], and the use of decentralized publication systems for open science using, for example, existing technologies like Blockchain and InterPlanetary File System (IPFS) [155].

A range of methods have been proposed to make scientific articles more machine-readable: from structuring scientific works as Research Object (RO)s [10, 11] to using facets in order to uncover the main methods, data, code and other objects that are used in scientific articles [98, 121]. Most approaches, however, have focused on automated content extraction from scientific articles as they are currently available. Recent machine learning techniques, for example, can after training with large sets of scientific articles extract the main concepts and structure of scientific articles [166, 170]. While the results can be very valuable there are also clear limitations, with the resulting data needing almost always manual curation to achieve decent quality [37, 52, 136].

A significant number of vocabularies and ontologies in many various domains have been developed, which are now ready to be used for scientific knowledge representation. But, even though practical problems like finding, accessing and versioning among other things have been reported [55, 59, 61, 68], these vocabularies and ontologies have proven to be extremely useful for example for biomedical literature curation [105, 145]. A considerable amount of attention has also been given to the datasets accompanying scientific articles. The Data Set Knowledge Graph (DSKG), for example, covers datasets from over 600k scientific publications [46]. An important development in this respect is the strong momentum behind the FAIR initiative to make research data Findable, Accessible, Interoperable, and Reusable [165]. A large amount of research is ongoing on how these FAIR principles can be put into practices (e.g. [55]). Many other aspects of scientific communication have been approached with more formal representations, such as declaring authorship contributions with the Contributor Roles Taxonomy [99] to mention just one of them.

Semantic technologies have been used extensively in the Life Sciences, e.g. for the representation and discovery of concepts, their relations and associated supporting evidence in order to integrate distributed repositories [60]. A variety of controlled vocabularies exist in these fields that can serve as the foundation to represent scientific knowledge in a structured way in order to semantically capture the context of scientific findings [30, 95, 146].

The BEL language [145] is one of the few attempts to represent the high-level scientific claims themselves, with coverage for specific kinds of biological relations. One of the first attempts that follows the "genuine semantic publishing" vision with a focus on scientific findings from the life sciencies field is the BEL [145]. BEL is a language that was developed to express in a computable format scientific findings, being used initially mainly for curation purposes of biological data and later for more complex tasks. Recent research has shifted the initial curation purpose of BEL into multiple development directions, showing the full potential that such computable representations can entail, despite being still limited to the life sciencies field: from software and visualization [62], to algorithms and analytical frameworks [171], data integration [43], natural language processing [139], curation workflows [63], to content and applications [73].

Also many other domains besides the Life Sciences have adopted the principles and technologies of Linked Data and the Semantic Web, for example to build interlinked, heterogeneous, and semantically rich datasets in Cultural Heritage [65] and to find, address, and sometimes even solve research problems in Digital Humanities in interactive ways [66].

5.2.2 *Semantic publishing and semantic papers*

Semantic publishing applies semantic technology to scientific publishing, and comes in many forms and does not always align with what we have introduced above as genuine semantic publishing [82]. Under this umbrella of semantic publishing, there are approaches that generate semantically-enriched data models from digital publications for the integration, sharing, management and data comparison between publications [111], study the semantic annotation and enhancement of scholarly articles [140], provide dynamic visualizations in semantically enhanced papers [135], assess the versioning aspect of semantic publishing [107], create a global-scale platform with a dataset metadata for automated ingestion, discover, and linkage [67], and propose semantic and web-friendly HTML-based alternatives to the currently PDF-focussed scientific writing process [117]. Semantic enhancements of scientific articles can be used for semantic interlinking, interactive figures, re-orderable references and even summary creation [142], and workflows to convert regular scientific articles into LOD have also been investigated [131]. Other approaches like the CISE advocate for a process of automatic semantic enhancement with semantic annotations [115]. A key role in most of these approaches is played by the variety of existing ontologies covering many different aspects of scientific publishing, most importantly the SPAR [114].

Further note-worthy approaches include the work to semantically represent the setup and results of scientific studies, which then allows for running meta-analyses in a semi-automated way, better research replication, and automated hypothesis generation [156], and the development of the ORKG [70]. The latter is an initiative that aims to make research articles machine-readable by expressing their main scientific entities as a semantically interconnected knowledge graph. This graph is populated by methods such as extracting scientific concepts from the abstracts of scientific articles with the help of annotators [18].

5.2.3 *Nanopublications*

Nanopublications [58] are a specific concept and technology that deserves special attention here. They have been proposed to express scientific (and other kind of) knowledge in Linked Data (LD) as small independent publication packages. They allow for rich provenance and metadata and are structured as follows: the assertion part contains the main content of the nanopublication, such as a scientific claim, expressed as RDF triples. The provenance part of a nanopublication describes how the assertion came about, e.g. by linking to the scientific methods used to arrive at the finding. The publication information part, finally, contains metadata about the nanopublication as a whole, such as by whom and when it was created. Nanopublications can be used for scientific findings, but also for representing the other elements of the scientific workflow, such as reviews and method descriptions, and more generally any kind of small coherent set of RDF triples [78]. It has been shown how nanopublications can be made reliable and immutable by identifying them with cryptographic Trusty URIs [80, 81], and how this allows for a decentralized network of services and template-based user interfaces such as Nanobench [83].

5.2.4 *Genuine semantic publishing*

In the "genuine semantic publishing" vision [82] semantics are taken into account before, during and after publication and these pertain integrally to the publication itself. As such, genuine semantic publishing means not only to formally represent from the start the structure and content of authentic fine-grained representations of research work, but also to publish these semantic representations (directly by the authors themselves) as main elements of a published entity without the need for a separate narrative article to exist. For most current approaches, in contrast, semantics are considered only after the publication of scientific articles, with semantic annotations, semantic interlinking and semantic integration used to semantically enrich and extract information from research that is still being published in coarse-grain texts in natural language.

One of the first attempts that aligns with some of the elements of genuine semantic publishing is the markup-language TaxPub [109]. TaxPub enables the publication of specific structured information for biological systems, but this is meant only as a tool that is able to provide semantic enhancements for already published scientific articles, hence not the publication of formal fine-grained authentic work by itself. More recently, a set of more advanced and complex tools for biodiversity literature have been created, like the OBKMS [108]. This system is able to not only link reusable data with its provenance, but also to provide a platform for the complete publishing process of biodiversity data from its submission to its reviewing, to its publication and dissemination. In this way, the complete publication process of scientific findings where facts and claims are linked to their original publications is supported by using a community accepted interoperable open format for biodiversity data. Furthermore, OpenBiodiv [110], an OBKMS that uses a Linked Open Dataset generated from scientific literature is able to provide an infrastructure where biodiversity knowledge can be managed, but this is also based on data extracted from already published scientific articles.

Another project that comes quite close to the genuine semantic publishing vision, even though it does not comply with it completely, is the creation of the SKGO [47], an initiative that seeks to organize the scholarly information published online in terms of its content, with a focus on the representation of scientific findings from various fields. As such, workflows that use semi-automatic methods to capture the contents of research findings in a structured manner have been proposed [159], but again, their core assumption is that this knowledge needs to have been previously published in research articles.

To conclude, apart from very few exceptions such as the work done in the biodiversity data field and the SKGO project, hence on restricted fields and restricted kinds of claims, genuine semantic publishing remains a vision for which we have little practical evidence as of now on how it can work in practice. So, there is still a huge gap with regard to making scientific knowledge machine interpretable, despite all these useful attempts and approaches that aim towards machine interpretability. As such, for a publication to be genuinely semantic, not only the semantic representation from the start of the structure and content of a fine-grained and authentic primary component of a publication entity needs to be considered, but also all the aspects that pertain to its publication. And, as we noticed in the projects mentioned above, the combination of all these requirements together is not present in current research.

5.3 Methodology

In this section we describe the approach and methods we followed to investigate whether nanopublications and the super-pattern are suitable to achieve genuine semantic publishing.

5.3.1 *Approach*

In our approach, we committed to a number of features. First, we wanted the final contributions to be published as "real" papers in a real established journal. They should be fully semantically represented (in RDF) but also have classical views that makes them look like other papers. Like that, they should also seamlessly integrate with the existing bibliometric system and it should be straightforward to cite them in the classical way.

Second, we decided to fully focus on arguably the most interesting element of scientific articles, which happens to also be one of the most challenging to formally represent: the main scientific claims the article is making. Scientific articles have a large number of other interesting pieces of information, e.g. information about the used methods among many other things, but for the purpose of the study to be presented, we focus only on the main claims.

Third, in order to retain the flexibility and power of nanopublications, we decided to refrain from providing a custom-built and optimized user interface that hides the complexity and limits the flexibility. By using generic template-based nanopublication tools and by customizing them solely by providing the templates, we hoped to get a better understanding of how the nanopublication technology works for such kinds of content and workflows in general, and not just for our specific case. On the other hand, this also means that we were looking for a bit more technically minded authors who can handle interfaces that do not come with all the comfort of polished specific applications.

Fourth, we wanted to test a system that *could* be used to publish novel claims, but decided for practical reasons to focus on formalizing claims from previously published articles. Our approach is therefore based on what we call *formalization papers* that contribute novel formalizations of existing claims.

Finally, we wanted to cover not just these main claims, but the whole publishing workflow that involves the initial submission of contributions, their reviewing, the responses to the reviews, the updated versions, and the final decision, and represent these as independent but interlinked nanopublications.

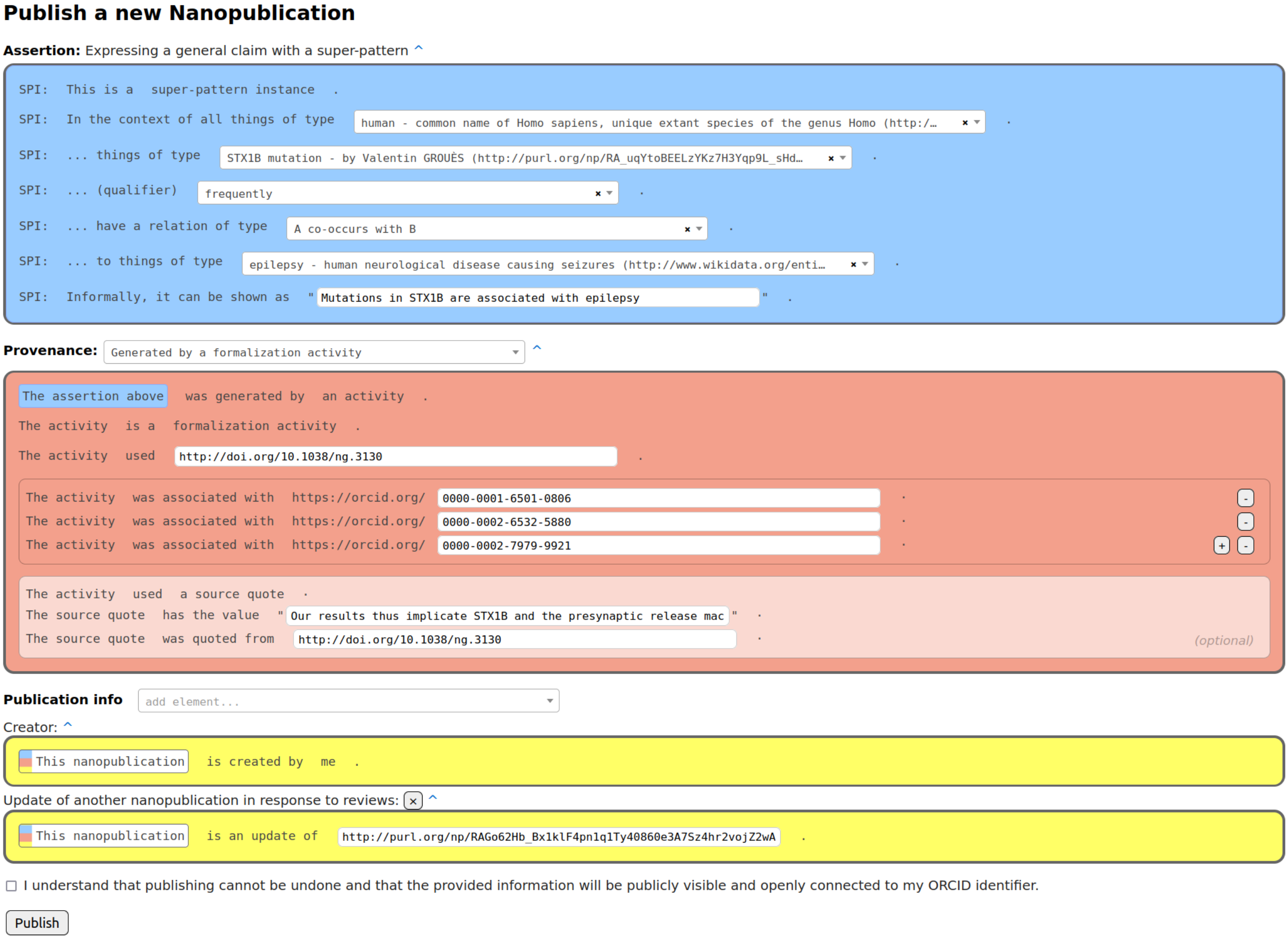


Figure 12: Formalization paper template from Nanobench as used by the participants of our study.

5.3.2 *Formalization Papers*

Our approach builds upon our new concept of formalization papers. A formalization paper contributes a semantic formalization of one of the main claims of an already published scientific article. Its novelty therefore lies solely in the formalization of a claim, not the claim itself. The authors of such formalization papers consequently take credit for the way how the formalization is done, but not for the original claim (unless that claim happens to come from the same authors).

The content of a formalization paper is fully expressed in RDF in the form of nanopublications. Such a formalization paper can be shown in other formats to users, e.g. in HTML or PDF, but these are just views of the same underlying RDF content. Our formalization papers consist of nanopublications in which the assertion contains the formalization of the scientific claim using the super-pattern [24], the provenance points to the original paper of the claim, and the publication information attributes the author of the formalization. Figure 12 shows an example of such a nanopublication in the interface the participants of our study used to create them. The instantiated super-pattern in the assertion part refers to a context class, a subject class, a qualifier, a relation type, and an object class according to the super-pattern ontology[4]. In the process of coming up with such a formalization, one often realizes that for some of the class slots of the super-pattern (i.e. context, subject, and object class) the class that should be filled in to arrive at a correct formalization is not directly defined in any existing vocabulary or ontology and as such, this class might need to be minted as well. The provenance part of the nanopublication describes the "formalization activity" that was conducted in order arrive at this formalization from what is written in the source publication. The precise phrase from that source publication that was used can be quoted too.

5.3.3 *Tools*

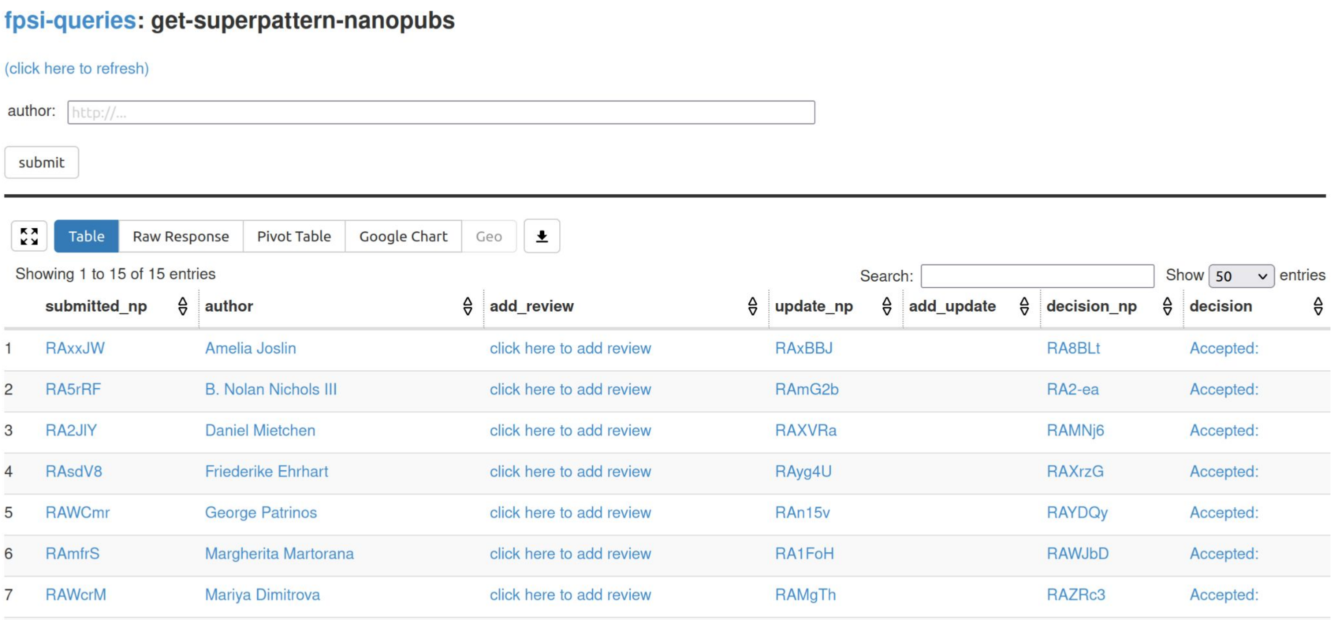
In order to publish formalization papers, class definitions, and all the other kinds of nanopublications (submissions, reviews, responses to reviews, and decisions), we use Nanobench[5] [83]. Figure 12 introduced above shows a screenshot of the publishing page of Nanobench. Publishing in Nanobench is based on templates, which are themselves expressed in nanopublications. The form shown in the screenshot is automatically generated based on the information found in several template nanopublications that we created and published for that purpose. All the application-specific behavior is therefore semantically represented in the templates, and Nanobench can flexibly be used for any other kind of data and workflow.

The second tool that we are using, Tapas[6] [93], is equally generic. It is a simple user interface component built on top of grlc[7] [100] that allows to run template-based SPARQL queries on RDF triple stores. In our case, we run it on SPARQL endpoints provided by the nanopublication service network [83]. We use Tapas to show aggregations and overviews of submissions and reviews. Figure 13 shows a screenshot of the main submission overview. Tapas by itself is read-only, but we connect to the Nanobench tool with links that lead to partially filled-in forms (e.g. "click here to add review" in the screenshot).

5.3.4 *Field study design*

In order to test our approach, we devised a field study where interested authors could submit formalization papers, which upon acceptance were published as a special issue in the journal Data Science[8] by IOS

4 https://larahack.github.io/linkflows_superpattern/doc/sp/index-en.html
5 https://github.com/peta-pico/nanobench
6 https://github.com/peta-pico/tapas
7 https://grlc.io
8 https://datasciencehub.net



fpsi-queries: get-superpattern-nanopubs

(click here to refresh)

author: http://...

submit

Table | Raw Response | Pivot Table | Google Chart | Geo

Showing 1 to 15 of 15 entries

Search: Show 50 entries

	submitted_np	author	add_review	update_np	add_update	decision_np	decision
1	RAxxJW	Amelia Joslin	click here to add review	RAxBBJ		RA8BLt	Accepted:
2	RA5rRF	B. Nolan Nichols III	click here to add review	RAmG2b		RA2-ea	Accepted:
3	RA2JIY	Daniel Mietchen	click here to add review	RAXVRa		RAMNj6	Accepted:
4	RAsdV8	Friederike Ehrhart	click here to add review	RAyg4U		RAXrzG	Accepted:
5	RAWCmr	George Patrinos	click here to add review	RAn15v		RAYDQy	Accepted:
6	RAmfrS	Margherita Martorana	click here to add review	RA1FoH		RAWJbD	Accepted:
7	RAWcrM	Mariya Dimitrova	click here to add review	RAMgTh		RAZRc3	Accepted:

Figure 13: The Tapas interface listing submitted formalizations as results of SPARQL queries over the nanopublication service network.

Press. The goal of this was to demonstrate for the first time that scientific articles can be formalized and therefore machine-interpretable including the main scientific claims. As a secondary goal, we wanted to find out whether nanopublications are a good technology for that, and whether it is feasible to represent also the entire submission and reviewing process within the same framework.

Because the user interfaces we have at our disposal are still quite rough and technical, we restricted the set of possible authors and sent the call for papers on a by-invitation basis to selected groups of researchers who have previously worked or had experience with technologies like RDF and semantics. We expect to be able to build more accessible user interfaces in the future that can show the inherent complexity in a way that does not require technical skills, but how this can be achieved is out of scope for this work.

Participants to our field study, thus the authors of formalization papers, formalized their own previously published claim, or a claim from a paper published by others. In the latter case, the formalization paper authors take credit for the formalization of the claim but not for the claim itself. All submissions to this special issue were peer-reviewed (also as nanopublications) using our previously proposed reviewing ontology [22]. Upon acceptance, these formalization papers were published in a journal at IOS Press, thereby giving them the same bibliometric status as other scientific articles, which leads to regular indexing in scientific article databases, counting of citations, and so on.

The authors received close guidance on how to represent a claim of their choosing in RDF using the super-pattern and nanopublications, and on the various stages of the publication process. Authors took part in several information sessions and discussion meetings and were provided at each step with helper materials, videos, and even direct assistance if needed. In total, 24 such individual sessions were organized from May to December 2021.

In order to define a formalization, sometimes some of the class slots (i.e. context, subject, and object slots) of the super-pattern should be filled in with classes that are not yet defined in any existing vocabulary or ontology. In this case the authors first had to define these themselves, and they could do that also with the Nanobench tool loading a template for class definition. (Alternatively, they could also mint a new class identifier by other means, such as creating it on Wikidata.) The assertion of a nanopublication defining a new class may look for example as follows (link to full nanopublication)[9]:

```
sub:STX1B-mutation a owl:Class ;
  rdfs:subClassOf wd:Q42918 ;
  rdfs:label "STX1B mutation" ;
  skos:definition "mutation in STX1B" ;
  skos:relatedMatch wd:Q18048867 .
```

Here, "mutation" from Wikidata (Q42918) is declared as superclass of the newly minted class "STX1B mutation", and "STX1B" (Q18048867) is linked as a related class.

Then the authors can publish their formalization in the form of a nanopublication using Nanobench (see Figure 12), and afterwards they needed to submit it to the special issue using another Nanobench template, leading to an assertion like (link to full nanopublication)[10]:

```
<http://purl.org/np/RAGo62Hb_Bx1klF4pn1q1Ty40860e3A7Sz4hr2vojZ2wA>
  pso:withStatus pso:submitted ;
  frbr:partOf fpsi:DataScienceSpecialIssue .
```

All submitted formalizations were subsequently reviewed. All authors were encouraged to review other submissions, and these reviews were semantic, open, and non-anonymous. These reviews were again done in nanopublications with the Nanobench tool. Such an example of a nanopublication assertion that contains a review modeled using the reviewing ontology can be seen below (link to full nanopublication)[11]:

```
sub:comment a lfr:ReviewComment , lfr:ContentComment ,
              lfr:NeutralComment , lfr:SuggestionComment ;
  lfr:hasCommentText "Maybe the use of a causal relation like
                      \"contributes to\" can also be used here." ;
```

9 http://purl.org/np/RA_uqYtoBEELzYKz7H3Yqp9L_sHdU-kgL8R5EqmBsTVzE
10 http://purl.org/np/RAWI_6Wpnnvn5scKXazYTqMftavW-HW9S-Alqlh1lf6Eo
11 http://purl.org/np/RAio--7IbPa3_ZSG3GspUsXeWP2ZwMIzy4Kzos0yZ7NIw

```
    lfr:hasImpact "1" ;
    lfr:refersTo
     <http://purl.org/np/RAGo62Hb_Bx1klF4pn1q1Ty40860e3A7Sz4hr2vojZ2wA> ;
    lfr:refersToMentioningOf sp:hasRelation .
```

In such a structured review (see more details in our previous research [23]), it is possible to specify various aspects that the review addresses including the aspect it comments on (syntax, style or content), the positivity/negativity of the review, the impact and the action that needs to be taken by the authors as the reviews see it (whether it is compulsory to be addressed, a suggestion or no action needs to be taken by the authors) and the importance of the point made by the review for the overall quality of the formalization. In the above example, the review comment makes a neutral point about the content of the given formalization with an importance of 1 out of 5, and is marked as a suggestion for the authors. The specific part of the formalization that this review targets is the *sp:hasRelation* field, as indicated by the *refersToMentioningOf* relation.

Subsequently, authors of the submissions could respond to the received review comments, again in nanopublications, and update their submissions based on these review comments. This is an example of a response to a review comment (link to full nanopublication)[12]:

```
  sub:comment a , lfr:ResponseComment lfr:DisagreementComment ,
      lfr:PointNotAddressedComment ;
    lfr:hasCommentText "I don't think the original publication shows a
    causal relationship. It seems to me only a correlation is proven." ;
    lfr:isResponseTo
     <http://purl.org/np/RAio--7IbPa3_ZSG3GspUsXeWP2ZwMIzy4Kzos0yZ7NIw> ;
    lfr:refersTo
     <http://purl.org/np/RAeRSya2qIYymsBxiqOZP_oaQpHXUVXiydKvPCFM-7DDQ> .
```

This response registers the agreement with the point made by the reviewer (whether the author agrees totally, partially or not at all) and if that point was addressed, partially addressed or not addressed at all by the author. Moreover, a link to the respective review is given using the *isResponseTo* relation, while the updated version of the formalization is indicated using the *refersTo* relation. In our example, we see that the author does not agree with the point made by the reviewer and hence did not address the point raised by him, and also give a textual motivation on why this is the case.

Finally, the authors updated their formalizations with the same template as depicted in Figure 12. The full final formalization nanopublication[13] of the same example is shown in Figure 14.

12 http://purl.org/np/RAAgR5ZKIIvujTwNwwxr6-bsjF1GXk_W7Zx7qxEeLrOX0
13 http://purl.org/np/RA22JAQihYeiJkNIjvwnxLPmjuG74yPcRXpPyVX8DV6fA

```
@prefix this: <http://purl.org/np/RAeRSya2qIYymsBxiqOZP_oaQpHXUVXiydKvPCFM-7DDQ> .
@prefix sub: <http://purl.org/np/RAeRSya2qIYymsBxiqOZP_oaQpHXUVXiydKvPCFM-7DDQ#> .
@prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#> .
@prefix xsd: <http://www.w3.org/2001/XMLSchema#> .
@prefix dct: <http://purl.org/dc/terms/> .
@prefix np: <http://www.nanopub.org/nschema#> .
@prefix npx: <http://purl.org/nanopub/x/> .
@prefix nt: <https://w3id.org/np/o/ntemplate/> .
@prefix sp: <https://w3id.org/linkflows/superpattern/terms/> .
@prefix lfr: <https://w3id.org/linkflows/reviews/> .
@prefix wd: <http://www.wikidata.org/entity/> .
@prefix orcid: <https://orcid.org/> .
@prefix prov: <http://www.w3.org/ns/prov#> .

sub:Head {
  this: np:hasAssertion sub:assertion ;
    np:hasProvenance sub:provenance ;
    np:hasPublicationInfo sub:pubinfo ;
    a np:Nanopublication .
}

sub:assertion {
  sub:spi a sp:SuperPatternInstance ;
    rdfs:label "Mutations in STX1B are associated with epilepsy" ;
    sp:hasContextClass wd:Q5 ;
    sp:hasSubjectClass <http://purl.org/np/RAPVWYH0x-xyDa9PfBcGUFly3m1FNEO43KG9s0uH-y6yo#STX1B-mutation> ;
    sp:hasQualifier sp:frequentlyQualifier ;
    sp:hasRelation sp:cooccursWith ;
    sp:hasObjectClass wd:Q41571 .
}

sub:provenance {
  sub:assertion prov:wasGeneratedBy sub:activity .
  sub:activity a sp:FormalizationActivity ;
    prov:used <http://doi.org/10.1038/ng.3130> , sub:quote ;
    prov:wasAssociatedWith orcid:0000-0001-6501-0806 , orcid:0000-0002-6532-5880 , orcid:0000-0002-7979-9921 .
  sub:quote prov:value "Our results thus implicate STX1B and the presynaptic release machinery in fever-associated epilepsy syndromes" ;
    prov:wasQuotedFrom <http://doi.org/10.1038/ng.3130> .
}

sub:pubinfo {
  sub:sig npx:hasAlgorithm "RSA" ;
    npx:hasPublicKey "MIGfMA0GCSqGSIb3DQEBAQUAA4GNADCBiQKBgQCY36SLWPLee0SZGM108+7dyjGzKFYg9t09XuL3js13jO3CDzqAZygcrwbJsbLQMRHYvWf0Mkly1ePLgdb43NqEbXiDHC4o49nHjhi2bSWeRDJ4..." ;
    npx:hasSignature "QF+C9lXmczrn9cJWuimwLG45Mpntk2CcRIWbeWmKvfE9gmQ6MPKa/x6AfNgVQRnPWppJdDoWepK6m/+m8tWY1WQsXn0KZ8sER+graEHQYuE70Mz9JzuBTyYu0vpWj5jteoCve5fyvFkhkYVjoRK9..." ;
    npx:hasSignatureTarget this: .
  this: dct:created "2021-10-29T10:35:33.912+02:00"^^xsd:dateTime ;
    dct:creator orcid:0000-0001-6501-0806 ;
    npx:introduces sub:spi ;
    lfr:isUpdateOf <http://purl.org/np/RAGo62Hb_Bx1klF4pn1q1Ty40860e3A7Sz4hr2vojZ2wA> ;
    nt:wasCreatedFromProvenanceTemplate <http://purl.org/np/RAB_oy10D3XUP-zYlqGz7Uj58AsUXhEKeGqmRFg5LSgDM> ;
    nt:wasCreatedFromPubinfoTemplate <http://purl.org/np/RAA2MfqdBCzmz9yVWjKLXNbyfBNcwsMmOqcNUxkk1maIM> , <http://purl.org/np/RAOGu9Lh0BD4tbIRB9RG6RGRA_ObDh75NTbIqaWgxxs8M> ;
    nt:wasCreatedFromTemplate <http://purl.org/np/RAv68imZrEjfcp2rnEg1hzoBqEVc0cQMtp9_1Za0BxNM4> .
}
```

Figure 14: A nanopublication view of a formalization paper.

For all updated submissions then a decision was made by us as the special issue editors about their acceptance. This decision was also represented as a nanopublication that looked as follows (link to full nanopublication)[14]:

```
<http://purl.org/np/RAeRSya2qIYymsBxiqOZP_oaQpHXUVXiydKvPCFM-7DDQ>
  dct:description "All review comments were addressed and the
                                    formalization looks good." ;
  pso:withStatus pso:accepted-for-publication ;
  frbr:partOf fpsi:DataScienceSpecialIssue .
```

All formalizations reached a satisfactory level of quality, as indicated by the reviews and the authors' responses, and we therefore accepted all 15 submissions for publication.

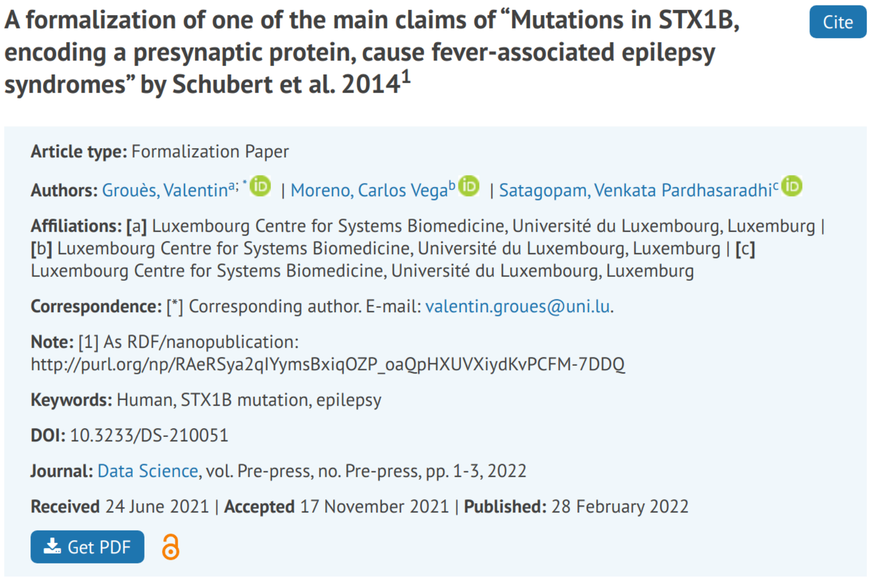
In order to show the accepted papers in the special issue as if they were classical papers, to integrate them in the publisher's content management system, and to make them connect to the existing bibliometric system, we semi-automatically created "classical views" in the form of HTML and PDF versions of the nanopublications, as can be seen in Figure 15 (link to the human-readable view)[15].

5.3.5 *User Feedback*

In order to evaluate the general idea of formalization papers, all participants to the field study were asked to give us their opinion and report on their experiences about the involved processes and concepts. This evaluation was performed by means of a structured questionnaire consisting of four main parts, each one evaluating different aspects of the workflow.

In the first part, we were interested in assessing the difficulty of conceptually understanding the formalization paper idea and the super-pattern, and of performing the formalization tasks. In part two, we focus on the difficulty of the technical aspects in the various submission, reviewing and revision stages. Part three addresses some more general aspects about the authors' experience and preferences. Authors were asked about their confidence in the formalization they published and about their interest of publishing such formalizations along their scientific publications in the future. We also asked them how important they think it is that all these steps are performed by the authors themselves (as they did). Moreover, in this part, authors could give us their opinion with regard to the importance of having a "classical view" along with the nanopublication representation of their formalization paper. The fourth and final part of the questionnaire asked for the technical background of the authors. At the very

14 http://purl.org/np/RAKNnwB9sUaOdqUz3vk6FvIY8ckt5NsEn3scZb0MLux00
15 https://content.iospress.com/articles/data-science/ds210051



A formalization of one of the main claims of "Mutations in STX1B, encoding a presynaptic protein, cause fever-associated epilepsy syndromes" by Schubert et al. 2014[1]

Cite

Article type: Formalization Paper

Authors: Grouès, Valentin[a; *] | Moreno, Carlos Vega[b] | Satagopam, Venkata Pardhasaradhi[c]

Affiliations: [a] Luxembourg Centre for Systems Biomedicine, Université du Luxembourg, Luxemburg | [b] Luxembourg Centre for Systems Biomedicine, Université du Luxembourg, Luxemburg | [c] Luxembourg Centre for Systems Biomedicine, Université du Luxembourg, Luxemburg

Correspondence: [*] Corresponding author. E-mail: valentin.groues@uni.lu.

Note: [1] As RDF/nanopublication: http://purl.org/np/RAeRSya2qlYymsBxiqOZP_oaQpHXUVXiydKvPCFM-7DDQ

Keywords: Human, STX1B mutation, epilepsy

DOI: 10.3233/DS-210051

Journal: Data Science, vol. Pre-press, no. Pre-press, pp. 1-3, 2022

Received 24 June 2021 | **Accepted** 17 November 2021 | **Published:** 28 February 2022

Get PDF

Figure 15: The human-readable view of a formalization paper, as it appears on the publisher's website.

end, the respondents could give further free-text feedback. The full questionnaire is available in our supplemental material[16].

5.4 RESULTS

In this section we present the formalizations that resulted from our field study. We present a descriptive analysis of the generated data and analyze it also with the help of a network visualization. Finally, we report on the results from the user feedback questionnaire.

All the nanopublications that were created for all the submissions, formalization paper versions, the review comments, the responses to the review comments and the newly minted classes used in the formalizations together with the decisions are accessible online[17], while the nanopublication index[18] containing all these nanopublications has also been published. Also, the final submissions for the special issue with formalization papers at the Data Science journal[19] that was released in March 2022 can be found online[20].

16 https://github.com/LaraHack/formalization_papers_supplemental/tree/main/questionnaire
17 https://github.com/LaraHack/formalization_papers_supplemental/tree/main/nanopubs
18 http://purl.org/np/RAkLJW7vIsnKKJDf1iswdgtFPQSo3lEG_z8DhHfD7dofE
19 https://content.iospress.com/journals/data-science/5/1
20 https://github.com/LaraHack/formalization_papers_supplemental/tree/main/accepted_submissions

5.4.1 *Analysis of Formalizations*

In total, we had an initial number of 20 people that replied to our call for papers[21] from 12 different institutions from the United States of America, Germany, Luxembourg, Bulgaria, and The Netherlands from fields like biomedicine, bioinformatics, health sciences, ecology, data science, and computer science. After an initial information session, out of the 20 authors that responded to the call for papers, 18 decided to continue their participation. All these 18 authors that responded to the call for formalization papers managed in the end to publish (upon acceptance) their articles in a special issue at the Data Science journal.

We had a total of 15 formalization paper submissions, 13 with individual authors and 2 with joint authorship. Out of the total of 18 authors, two of these have both an individual submission and a joint-authorship one. The super-pattern instantiations of the final accepted formalization paper submissions can be seen in Table 13. Here, the classes used to instantiate the super-patterns that comprise the formalizations are given for each submission: the context, subject and object classes for each submission are listed, together with the qualifier and relations selected from the SuperPattern ontology [24]. Each instantiation of the super-pattern can be interpreted as follows: *"Every thing of type [SUBJECT] that is in the context of a thing of type [CONTEXT] [QUALIFIER] has a relation of type [RELATION] to a thing of type [OBJECT] that is in the same context."*.

In the same Table 13 we can also take note of the distribution of qualifiers and relations that were used in the instantiated super-patterns of the accepted formalizations. As such, the most used qualifier is "generally" in almost 47% of cases (7 formalizations), while its modal counterpart, "can generally" is next in 20% of cases (3 formalizations). While all positive, non-modal qualifiers defined in the ontology seem to be used at least once (in at least one formalization), the only negative qualifier used was "never", in almost 7% of cases (1 formalization). The most used qualifiers are positive with about 73% (11 formalizations) and modal positive with 20% (3 formalizations), while the negative qualifiers seem to be less common with about 7% (1 formalization) and the modal negative qualifiers were never used. In terms of the relations used, we observe that relations that express causal relations are the majority with 80% (12 formalizations), then the next used is the equivalency relation with almost 13% (2 formalizations), then with the smaller ration, the ones about the spatio-temporal relations with only almost 7% (1 formalization), while the relations making numerical comparisons (the "compares to" relations) were never used.

Looking at Table 13, we see that the super-pattern instances exhibit quite a broad variety of scientific fields (bioinformatics, biomedicine,

21 https://github.com/LaraHack/formalization_papers_supplemental/tree/main/call_for_papers

pharmacology, data science, computer science) mostly linked to the life sciences. 7 out of the 15 submissions contain a formalization in which authors extracted a scientific claim from their own previously published article (submission number marked with ◊). Additionally, out of the total 44 classes used in the formalizations, 22 new classes were minted using Nanobench (marked with *), while 4 were newly minted Wikidata classes (marked with **). 13 already-existing classes were reused from Wikidata (their Wikidata identifier is specified next to the class name) and 4 classes were referenced from other ontologies.

5.4.2 *Analysis of Nanopublications*

In this field experiment we used nanopublications to embed, not only the formalizations created, but also the entire publication process that these formalizations underwent. As such, the entire formalization papers creation and publication process was thoroughly documented and published in a formal and machine-interpretable way, made possible by making use of nanopublications as "FAIR data containers". All the nanopublications pertaining to the special issue with formalization papers at the Data Science journal have been retrieved from the nanopublication network[22] and made available online[23] after serialization in *trig* files.

Table 14 shows the statistics about the nanopublications created during our field study. It shows a total of 15 submissions with their 15 corresponding super-pattern definitions; the content of these submissions is the one summarized in detail in Table 13. There are 25 updated super-patterns, indicating that some of the submissions were updated more than once. 34 new classes were minted in nanopublications as class definitions, which were subsequently used in the formalizations. With regard to the reviews received and the author responses, class definitions received an average number of around 3 reviews per class (46 review comments in total), while the super-pattern definitions had almost 8 review comments on average (119 review comments in total). In terms of the responses given to these reviews, the average responses to class definitions was a little over 2 (34 review comments in total), while the average number of responses to the review comments for the super-pattern definitions was about 6.7 (100 review comments in total).

In Figure 16 we can see a graphical representation of all the special issue nanopublications, where each node represents such a nanopublication and the arrows between the nodes show how the nanopublications are linked semantically with each other. The legend for the node types indicated by color and letter code can be found in Table 14.

22 https://monitor.petapico.org

23 https://github.com/LaraHack/formalization_papers_supplemental/tree/main/nanopubs

	CONTEXT ("in the context of all ...")	SUBJECT ("things of type ...")	QUALIFIER	RELATION	OBJECT ("to things of type...")
1	early human adipogenesis*	regulatory element within the first intron of FTO*	generally	affects	expression of genes IRX3 and IRX5*
2	human motor neuron (Q101404862)	TAR DNA binding protein (Q21133247)	can generally	contributes to	transcription of stmn2*
3 ◇	dejellied fertilizable stage VI Xenopus laevis oocyte**	strong static magnetic field**	generally	affects	cell cortex (Q5058180)
4 ◇	(no context class)	genes associated with CAKUT**	sometimes	is same as	targets of vitamin A**
5 ◇	patient undergoing PCI*	pharmacogenomics guided clopidogrel therapy*	generally	enables	cost-effective treatment*
6	human (Q5)	smoothened signaling pathway	mostly	affects	astrocyte development
7 ◇	biodiversity data (Q28946370)	license with non-commercial clause*	generally	inhibits	data reuse (Q58023280)
8 ◇	release of OpenBiodiv knowledge graph*	triple in OpenBiodiv knowledge graph*	generally	is same as	semantic triple extracted from biodiversity literature*
9	UNC13A (Q18036664)	TAR DNA binding protein (Q21133247)	generally	inhibits	inclusion of cryptic exon
10 ◇	data set (Q1172284)	adherence to the FAIR guiding principles*	can generally	enables	automated discovery*
11	human (Q5)	NGLY1 deficiency	always	is caused by	dysfunction of ERAD pathway*
12	social group (Q874405)	relative neocortex size*	never	affects	social group size*
13	ecm bound cancer cell*	glycocayx bulk*	generally	increases	integrin clustering*
14	human (Q5)	STX1B mutation*	frequently	co-occurs with	epilepsy (Q41571)
15 ◇	digital humanities research*	usage of Linked Data Scopes*	can generally	contributes to	transparency (Q535347)

Table 13: Instantiated super-patterns accepted for publication in formalization papers in the Data Science special issue. Submissions marked with ◇ are formalizations in which authors extracted a scientific claim from their own previously published article; classes minted using Nanobench are marked with *, while newly minted Wikidata classes are marked with **.

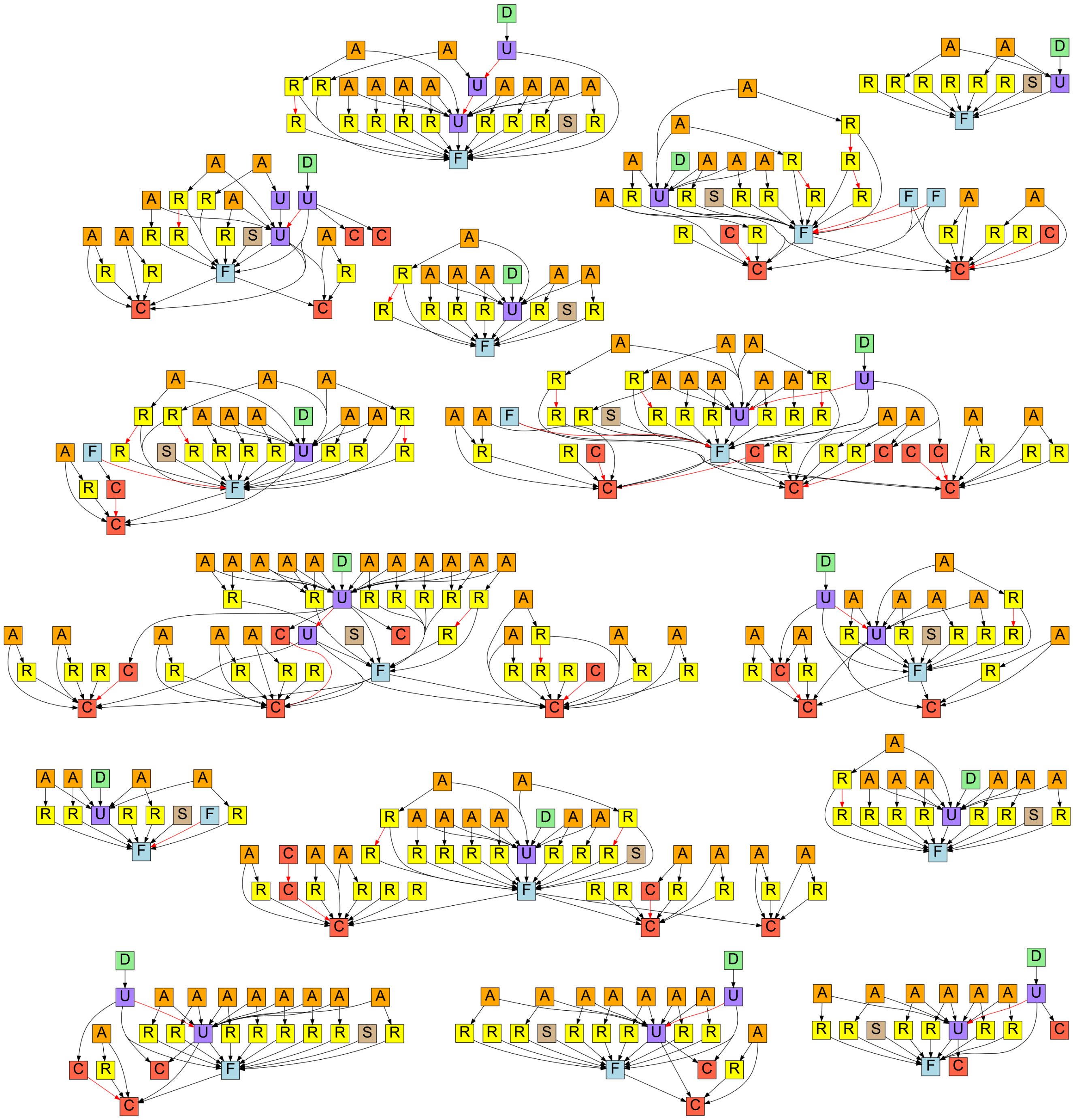


Figure 16: Graphical representation of all the submissions to the Data Science special issue with formalization papers. A click-able version with links to the nanopublications can be found online: https://raw.githubusercontent.com/LaraHack/fpsi_analytics/main/np-graph.svg.

Table 14: Nanopublications created during the field study of the special issue with formalization papers at the Data Science journal.

icon	type	average number per submission	total
S	submissions	1.00	15
F	super-pattern definitions	1.00	15
C	class definitions	2.27	34
R	reviews of super-patterns	7.93	119
R	reviews of class definitions	3.07	46
A	responses to super-pattern reviews	6.67	100
A	responses to class definition reviews	2.27	34
U	updated super-pattern definitions	1.67	25
D	decisions	1.00	15

For every formalization paper, we see a first formalization (F) together with a submission nanopublication (S). Later updated versions (U) of formalizations also link back to the initial formalization. The initial submissions received review comments (R), to which authors then answered with response nanopublications (A). Additionally, some of the formalization papers used newly minted classes (C), which then also received review comments and responses. The final decision (D) points to the finally accepted updated formalization. The edges (i.e. arrows) of the graph indicate when a nanopublication is referring to another one by using its identifier in the assertion. The edges shown in red are *superseding* relations, pointing from a new version of a nanopublication to its previous version. This is how nanopublications, being immutable, are dealing with representing new versions.

5.4.3 *User Feedback Analysis*

The 18 authors and co-authors of the formalization papers were asked to fill in the user feedback questionnaire. It was important for this questionnaire to be fully and reliably anonymous, as the authors needed to be able to give their honest opinions. This meant that we had to send reminders without knowing who already filled it in. After several rounds of reminders, we ended up getting 19 responses, meaning that at least one of the authors submitted two responses. Due to the anonymous nature of this questionnaire, it was not possible find out which responses were affected, and we have therefore to deal with such a dataset of slightly imperfect representation.

In Figure 17 we see the results for the first part of the questionnaire. Authors expressed that it was rather easy to understand what a formalization paper is (with a score of 4.32 out of 5). The elements of the super-pattern were found a bit harder to understand but still

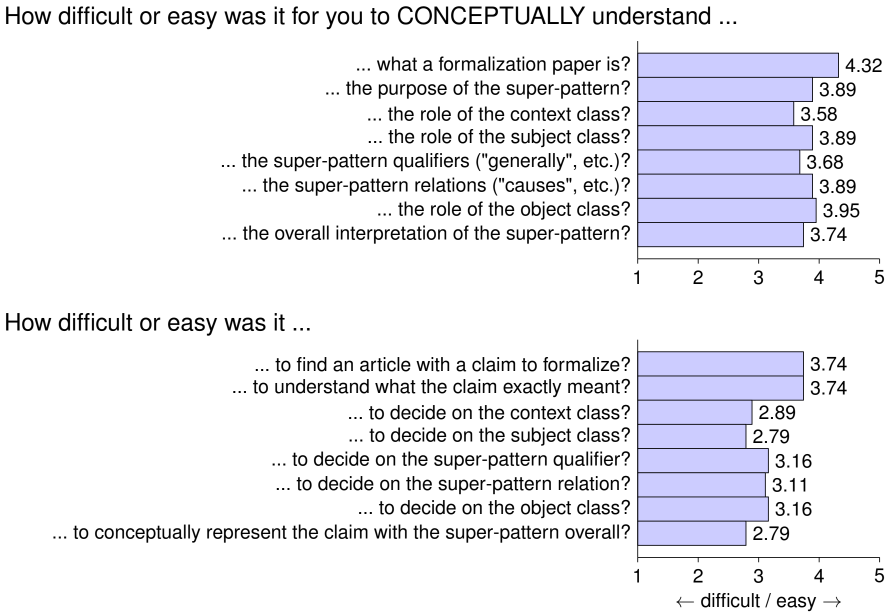


Figure 17: Questionnaire Part 1: Average answers from participants on conceptual aspects of formalization papers.

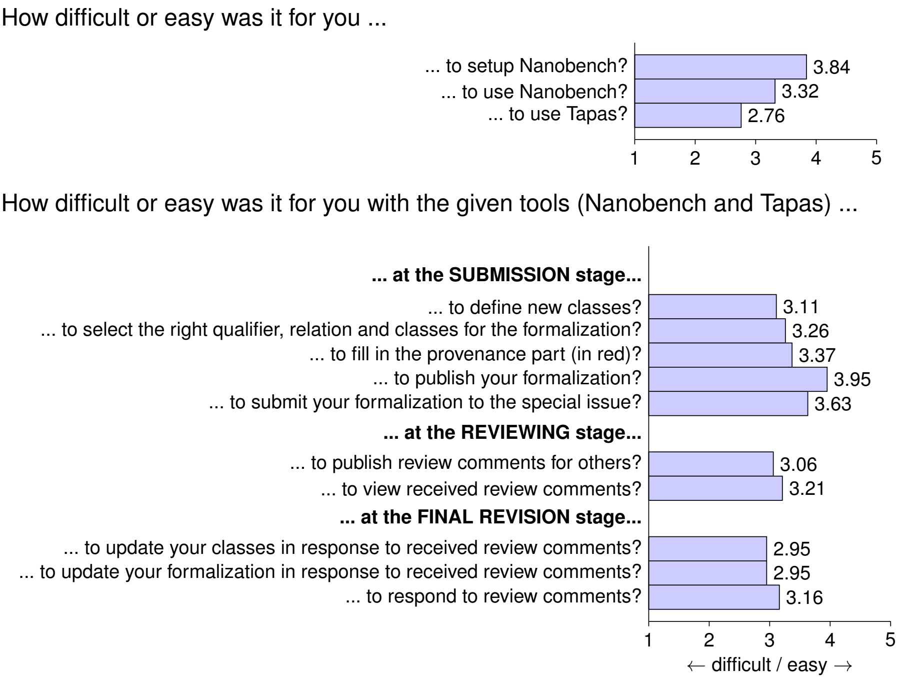


Figure 18: Questionnaire Part 2: Average answers from participants on technical aspects of formalization papers.

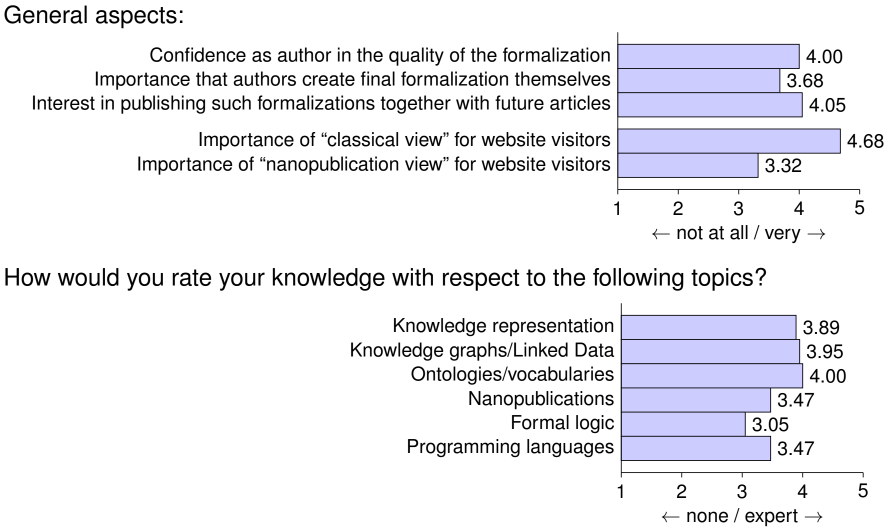


Figure 19: Questionnaire Parts 3 and 4: Average answers from participants on general aspects about formalization papers and average rating of participants with respect to their background knowledge on various topics.

quite easy, with scores above 3.50. Finding an article from which to select a claim to formalize and to understand what the chosen claim really meant was also deemed easy, with scores of 3.74. The actual instantiation of the super-pattern with all its fields given the chosen claim was considered a little more difficult, with scores around 3.0, indicating medium difficulty roughly in the middle of *very difficult* and *very easy*. These results seem to suggest that the authors were able to understand the main formalization papers idea together with the super-pattern that comprises it, but when it came to the actual instantiation of the super-pattern (especially concerning the context and subject class), this was considered a little more difficult, but still on average far from very difficult.

In Figure 18, we see the authors’ responses with respect to technical difficulty. In terms of the tools used, we see that setting up and using Nanobench was considered easy enough (with a score of 3.30), while the Tapas interface seems a little harder to use (with a score of 2.76). The different tasks in the different stages all seemed to be between medium and easy on average, with the exception of the tasks to provide responses to reviews, which scored slightly below 3.0. The response nanopublications are indeed among the most complex ones, as they refer not only to the affected review but also to the updated formalization. Overall, while these results show room for improvement, they still seem favorable given that we were building upon generic and powerful tools without specific user interface design or polishing.

Figure 19 summarizes the assessment of more general aspects of formalization papers and also contains information about the authors’ background. We see that authors have a high confidence in the quality of their formalization, with an average score of 4.0, and that they are interested in the future publication of such a formalization along their scientific publications, with a score of 4.05. The respondents very clearly stated that the classical view of formalization papers is important for website visitors, with a score of 4.68. Exposing also the "naked" nanopublications to the website visitors with a nanopublication view was found to be much less important (3.32).

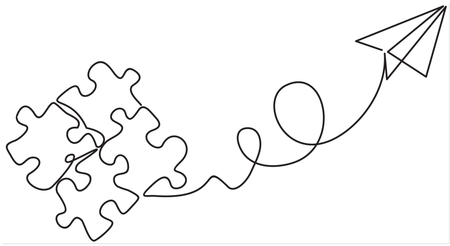
The authors indicated that they have, on average, a high level of knowledge on the topics of knowledge representation, knowledge graphs, Linked Data, and ontologies/vocabularies, with scores from 3.89 to 4.00. Their background in nanopublications, formal logic, and programming languages was significantly lower, on average, but still relatively high, with scores between 3.05 and 3.47.

10 out of the 18 authors used the free text feedback of the questionnaire. 8 of these 10 respondents expressed their excitement about the field study and found the formalization paper concept and the whole publication process interesting and useful. However, half of these respondents also mentioned that the overall process proved to be a little more difficult than they expected, due to the tools used maybe being too technical. One author also pointed out that multiple formalizations can be written for the same claim by choosing the context, subject and object classes differently and expressed the worry that this would decrease the interoperatibility or utility of formalizations especially when aggregating or mining them. This is a reasonable point to make, but due to the fully formal semantics, syntactic differences are in principle not hindering this kind of interoperability. Overall, the super-pattern, the formalization paper concept, and the nanopublication-based publication workflow seem to have been well-accepted and understood by the participants, and many of them showed an enthusiastic reaction.

5.5 DISCUSSION AND CONCLUSION

The publication of the special issue with formalization papers at the Data Science journal shows not only that nanopublications and the super-pattern can be used to implement the basic steps and entities of a journal workflow, but also that authors of such formalization papers can be taught to use these in order to publish in a novel journal publication workflow as the publication of the special issue demonstrates. Our results show that the super-pattern can be well understood conceptually and despite the fact that from a practical standpoint applying it seems to be more difficult, its application remains perfectly feasible. Furthermore, we saw in our field study that even if the current general-purpose tools can be considered a viable solution, these are not necessarily easy to use, but they still remain a good tool for the purpose of publishing formalization papers. Moreover, considering the formalization papers, authors seem confident with regard to the quality of their publications and seem interested in publishing such formalizations in the future.

In future work, we plan to take the next logical step by publishing novel claims in this way from the start, and not depend on claims from already-published papers. These contributions will then also have to be accompanied by statements about the methods, equipment, and all other relevant scientific concepts, and can include not just the high-level claim but more lower-level ones, possibly all the way down to the raw data. This representation would then ideally cover the entire scientific workflow, starting from a motivation, leading to the design and execution of a study, and ending in new scientific insights. Such fully formalized scientific contributions can be seen as a major step — even a breakthrough — for the Semantic Web and Open Science movements and will bring us closer to a world where machines can interpret scientific knowledge and help us organize and understand it in a reliable and transparent manner.



6 CONCLUSIONS

For now, what is important is not finding the answer, but looking for it.
— Douglas Hofstadter, *Gödel, Escher, Bach: an Eternal Golden Braid*

In this chapter we summarize the research questions addressed in this thesis in Chapters 2 to 5. We also discuss the limitations of our work and identify future research directions that can extend the contributions and findings presented in this thesis.

6.1 RESEARCH QUESTIONS REVISITED

Considering the main research question ***How can we apply the Semantic Web principles and nanopublications to make genuine semantic publishing possible to scientific findings as well as their assessments?*** introduced at the beginning of this thesis, we further reiterate and discuss the results of the four sub-research questions that were used to answer it.

> *1. Can an approach for scientific publishing based on a fine-grained semantic model help to make reviewing better structured and more accurate?*

We addressed this research question in Chapter 2. After using the general principles of the Web and the Semantic Web in the field of scientific publishing, concentrating on the reviewing process, we created a novel model titled Linkflows that we showed was able to support a fine-grained, more detailed and semantically rich view on the reviewing process by linking the review comments and their relation to the article text. The results of our user study evaluation showed that a substantial level of disagreement arises if other actors like peer researchers or model experts try to reconstruct the intended nature of these review comments afterwards. This shows the importance of capturing this information from the beginning and at the source. Moreover, larger groups of peers can achieve lower disagreement with the ground-truth provided by the original reviewers, but the disagreement is still substantial and larger groups of annotators require more collective effort. After also applying various automated methods like sentiment analysis tools, we noticed that these can reach the same level of agreement like model experts, but a substantial level of disagreement remains for both of them. In summary, our research showed that the Linkflows model for reviews would support fine-grained, detailed and structured reviews, while the full value of our approach could be reached if the review comments are captured directly from the reviewers themselves.

> *2. Can we use nanopublications as a unifying data model to represent the structure and links of manuscripts and their assessments in a precise, transparent, and provenance-aware manner?*

We address this research question in Chapter 3. We created a representation of the text snippets and review comments as nodes in a network with semantics attached. We used nanopublications as containers to embed these text snippets and review comments and by the use of nanopublications, these snippets of information can be represented in a decentralized network with Web-like links that have semantics attached to them. The finer-grained way to represent and structure the article and its corresponding reviews makes use of semantically enriched models and ontologies, like the Linkflows model proposed in Chapter 2. Our evaluation demonstrated that this representation is useful in answering interesting research questions from the different perspectives of the editors, while a proposed prototype interface is able to support the editors in their work. The proposed fine-grained and semantic representation with nanopublications allows to automatically answer a wide range of competency questions for editors, while the proposed prototype proved useful to answer them, despite having a large variation in the preferences of editors. This is expected, as it is not necessary to design a single interface that suits everybody and due to the interoperability and openness of our approach, several visualization alternatives can be constructed.

> *3. To what extent can a semantic template called the super-pattern be used to formalize the main claims of scientific articles from different disciplines?*

We address this research question in Chapter 4. We created a dataset of formalizations of scientific claims using our super-pattern approach by extracting a random sample of scientific claims from scientific articles from all disciplines. The formalization study performed with knowledge representation experts indicates that our approach works for scientific claims from a variety of disciplines, confirming the cross-disciplinary nature and applicability of the super-pattern. Our results showed that in terms of modelling, the majority of the classes used in the super-pattern formalizations could not be fully defined from existing vocabularies and had to be newly minted. However, a more complex task than defining new classes and the difficult part of the problem was the structured and formal representation of the content of the scientific claim itself, mainly expressing the high-level logical structure of scientific claims. The formalization study also demonstrated that the formalizations created using the super-pattern could be done in a fairly consistent and reliable manner, despite the inherent complexity and difficulty of this task. Overall, in this research we demonstrate that it is possible to express the content of high-level scientific claims by writing formalizations of such claims that make use of the super-pattern ontology we developed. Moreover, the formalization study we conducted indicates that even if pre-requisite knowledge is necessary to achieve this, this task is one that can be done in a reliable manner.

> *4. Are nanopublications and the super-pattern appropriate concepts to enable a new paradigm of scientific communication where authors publish their scientific findings with formal semantics?*

We address this research question in Chapter 5. In this chapter we used all of the contributions developed in the previous chapters and made use of existing semantic web technologies and tools like nanopublications to enable the publication of a special issue with formalization papers at the Data Science journal. By this, we wanted to show not only that nanopublications and the super-pattern can be used to implement the basic steps and entities of a journal workflow, but also that authors of such formalization papers can be taught to use these in order to publish in a novel journal publication workflow as the publication of the special issue demonstrates. The results of our field study indicate that the super-pattern can be well understood conceptually, despite the fact that from a practical standpoint applying it seems to be more difficult, its application is perfectly feasible. Furthermore, we saw that even if the current general-purpose tools can be considered a viable solution, these are not necessarily easy to use, but they have the needed features for the purpose of publishing formalization papers. In addition, the results of the evaluation made by the participants of the field study (and thus the authors of the articles of the special issue with formalization papers) suggest a high degree of confidence with regard to the quality of the formalization paper publications and show that there is a high interest in publishing such formalizations in the future.

6.2 DISCUSSION

We discuss several topics that might benefit from the work conducted as part of this thesis as five open problems: (1) perfecting the peer reviewing process, (2) designing intelligent user interfaces for semantic publishing, (3) expressing scientific findings with formal semantics, (4) increasing interoperability in scientific publishing and (5) transitioning to new publication systems. The first open problem is related to the peer reviewing process. Despite gathering much criticism along the years, peer reviewing is still the main way of assessing the scientific contributions made by researchers in the form of scientific articles. While various changes have been proposed to alleviate some of the issues that the reviewing process entails, many open problems remain. The second open problem of discussion concerns user interfaces, mainly building intelligent user interfaces for the publication process. These can support users in a timely and friendly manner, even in the reviewing process, and are able to abstract the technicalities used to gather or represent the data behind the scenes, hiding even semantic formalisms that make the data machine-readable. The third open problem discusses some different ways of expressing and representing scientific findings, not only as part of scientific articles, but as individual contributions themselves with their own (machine-interpretable) representations and where this can fit in the publication process itself. The forth open problem deals with the topic of interoperability in the field of scientific publishing and its importance when it comes to the adoption of new standards and how incompatibilities in interoperability can significantly reduce the potential of a widespread adoption of even the best ideas. The fifth and final open problem we bring under consideration is the more general theme of the future of new publication systems, mainly with regard to the economic implications for the publishing industry.

PERFECTING THE PEER REVIEWING PROCESS

The semantic reviewing model which we propose in Chapter 2 that makes use of review comments, a form of reviews that are smaller, finer-grained, interlinked and contain semantics from the start, can provide multiple advantages to the current (imperfect) peer reviewing process. As such, using such rich semantic representations of reviews and the reviewing process can make the complicated and complex reviewing process more transparent and reduce the inherent subjective nature of peer reviews and help to diminish the time length of the entire process as reviews would be given on a finer-grained scale. Moreover, the formal representation would decrease reviewer bias and the finer, more specialised review comments could limit and even prevent "superficial reviews" where the reviews fail to identify serious methodology or research flaws. However, even as shorter, and more specific targeted review comments would be easier and less time consuming to write for reviewers, thus being more convenient than classical long text-based ones, a new way of reviewing puts an additional complexity layer on top of an already non-trivial process where reviewers complain about the lack of incentives to write reviews in the first place. As such, in the context of heavy workloads as the volume of published papers increases and their own scientific duties, researchers find less and less incentives for writing reviews, while, using and learning a new way of writing reviews, like the model we propose, despite more organized and accurate then the one currently in use, is an extra step for which extra incentives would need to be provided for its adoption. Furthermore, it is not clear how such incentives should or could be implemented and also if a larger adoption of such a reviewing process is even possible.

Some other solutions have been proposed to the complicated and non-trivial problems raised by the peer reviewing process, ones that consider different approaches that cater more towards the reviewing process in its current form. Some of these solutions include offering fast-track peer reviewing to speed up reviewing, giving rewards to reviewers as an incentive for writing reviews (like free access to the publisher's journal for a limited period of time or the possibility to show the contributions made on platforms such as Publons). Additionally, open or crowdsourced peer reviews are meant to ensure more review transparency, while combining them with post-reviews (like on the F1000 platform) can further ensure a more objective assessment of the scientific contribution under scrutiny. Other approaches are more disruptive in the sense of proposing a different paradigm of reviewing altogether [28, 133, 168], one that is able to handle reviews in a more machine-interpretable way.

But, while all these alternatives try to address certain criticisms of the peer reviewing process, they come with challenges of their own. And, while it is true that peer reviewing cannot fully eliminate all its downsides, it still remains the favoured way of assessing the quality of new research. Despite all its criticism, it is a process that has proved nonetheless its value by showing in the majority of cases that it can help improve and correctly assess research accordingly. In the end, it all comes down to having dedicated researchers that act responsible in both their roles as authors themselves and then as reviewers for others, in the context of a peer reviewing process that constantly adapts to the new technology advancements and to the challenges it faces.

Designing Intelligent User Interfaces for Semantic Publishing

Providing good user interfaces is challenging in particular for complex models and complex situations like the ones involved in the scientific publishing process. However, good user interfaces can speed up the process of publication (an advantage not only for authors, reviewers, editors and readers, the publication companies themselves, but also for the world of science as a whole), and moreover increase its transparency, especially if using some of the models and approaches indicated in the chapters of this thesis. As such, using content that is machine-interpretable (both in the content of the article itself and also in the reviews as described in Chapter 3) would allow reviewers to review specific parts of the manuscript content based on their expertise and availability, skipping the parts of the content for which they lack either expertise, interest or time. This would in turn allow editors to choose from a larger pool of reviewers and to aggregate smaller reviews for particular parts of manuscript content for the final decision, as demonstrated by our field study described in Chapter 5.

Also, in in general, intelligent user interfaces can further be a way of hiding the technical intricacies that can reside behind structured and machine-interpretable content that make information and content available not only for semantic knowledge experts, but for all, humans and machine alike, increasing discoverability according to the FAIR principles [165]. User interfaces can be a versatile approach to access and present the same information in different forms based on the audience, for machines in a machine-readable format, and for humans in HTML, or other standards, for example. And, last but not least, structured data would allow for incorporating chatbot-like user interfaces in publication platforms, that can again tackle some of the challenges of the scientific publication in its current form.

EXPRESSING SCIENTIFIC FINDINGS WITH FORMAL SEMANTICS

Scientific publications are a way of sharing with the research community of the respective field the new scientific findings that result as a consequence of the work undergone by a researcher or a group of researchers, incorporating in the scientific manuscript not only the scientific finding per se, but also the materials, methods, approaches, experiments and evaluations that have lead to and which support the respective scientific finding. As such, we can say that the scientific findings presented in a scientific article represent "the essence" or the core of the research article, enabling and paving the way for newer discoveries and scientific findings.

While it is hard to formalize into a machine-interpretable format the entire text of a scientific article, this is possible for scientific findings. As we have shown in Chapter 4, expressing the scientific findings is possible using formal semantics and this can pave the way to using more advanced reasoning that is able to maybe find novel findings through aggregation, to find other findings that support or contradict the current ones and to even support question answering systems. However, seeing how this high-order reasoning can be applied in practice is still a matter of research due to the innate issues characteristic to automated high-order logic, like decidability. It is an exciting topic to further study, especially with the rise of chatbots like ChatGPT.

THE CHALLENGE OF INTEROPERABILITY IN SCIENTIFIC PUBLISHING

The approaches proposed in this thesis show, especially in the field study performed in Chapter 5, that is it possible to make connections in a structured, semantic way between publications and authors, between publications and reviews, between editors and the final publication decision, creating a possible interoperable semantic network of the whole scientific publication process, where individual contributions can be accurately tracked and registered. However, while we show that such accurate tracking is possible, navigating from publications to the datasets used, to the works cited, to the publication reviews and their respective authors, to the steps of the decisions along the publication process, needs an interoperability layer that can facilitate making these connections in other real-world systems and between them. Without this interoperability layer and without a larger adoption of standards that can support such ideas, such proof of concepts might prove incompatible for the future and their adoption lacking altogether.

But, integrating and linking data from various sources with different schemas or data representations is possible (making updates and tracking ownership of data possible as well) if standards for data accessibility and interoperability are put in place. This is crucial for the progress of science, and this is a complex problem that is addressed only partially. An example in this respect is the DataCite project [19] which allows publishers to link with data repositories, to share, reuse data, and software and get the credit for it - making Digital Object Identifier (DOI)s available for the research ecosystem and as such making scientific contributions discoverable, citable, accessible and overall FAIR [165].

Entity and dataset linking is partially supported and sometimes scientific articles may link to external data repositories which have their own (meta)data schemas and most of the times there are no common protocols for technical interoperability or a minimal set of standards even for metadata between entities. As such, integration can be among the biggest challenges as it would require different technical implementations for every data silo that is to be connected in the absence of a unified interoperability layer which would facilitate in making connections between different scientific contributions on various platforms or data silos.

Interoperability can enhance data reuse, data aggregation from various different sources, eliminating data silos and as such overcome organizational and disciplinary boundaries, while common interoperability standards would ensure scientific data is computable, improving research quality through reproducibility and replicability. The interoperability of data and scientific articles would support the development of science as a whole by making scientific publications accessible and increasing the reproducibility of results, which would make the scientific journals more sustainable for the future. Hence, interoperability frameworks sustained by the publication of accompanying interoperability standards would allow information to flow unencumbered among both humans and machines, and provide readers (whether human or automatic) better access to all that the scientific community has to offer. Systematic solutions are required and there is a growing acceptance that there is plenty of room for improvement in this respect.

TRANSITIONING TO NEW PUBLICATION SYSTEMS

In this thesis, throughout all the chapters, we present new approaches that incorporate the principles of Semantic Web in order to make the scientific publishing process more transparent and machine readable, focusing on the different actors of the scientific publishing process, whether they are the authors, reviewers, editors and readers. However, whether or not the strategies put forward will actually be adopted by the publishing industry or not depends on a lot of variables, some of which are very hard to estimate or even foresee, starting from the economical point of view of the publishing houses to practical things like incentives for adoption and the compatibility with the current publishing systems.

The implications of the different strategies adopted by the different publishing companies to cope with the necessary changes required by the advancements in technology and in the new requirements brought forth by the main actors of the scientific publishing process (authors, editors, reviewers, and even readers), coupled with having a competitive advantage over other publishing companies (whether it is in the form of more prestigious publications, striving for more open access, payment only for more advanced services or ultimately higher profits) is still something to be seen and hard to predict.

But, it appears as if there are two conflicting sides to the same problem: one the one hand, we have the actors involved in the scientific publication process that advocate for a more timely, transparent, unbiased process, with sufficient incentives on all sides that are able to provide with quality and sound work in terms of the research presented and its reviews, while, on the other hand, we have the publishing companies that first need to make their way in the publishing industry and ultimately be economically profitable in order to survive on the market. And we also have the advocates of open science and open access whose goal is to strive to make science accessible to all, in all forms and for free for everybody. There can be lots of options, some of which can seriously limit the value of some of the approaches proposed in this thesis, but the future is hard to predict in this regard.

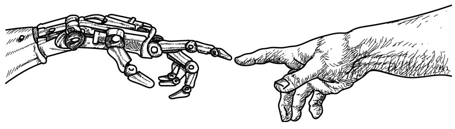
6.3 OUTLOOK

Finally, we can have a brief look into the future of scientific publishing and mention some possible impacts of this thesis. The current problems of the scientific publishing process are not expected to disappear in the near future. However, with the continuously increasing volume in the number of publications and the substantial criticism brought forth about the reviewing process, it is clear that the scientific publishing landscape will have to change in the future. Indeed, various steps have been taken to make the publication systems open-access and more open to automated systems, but the human layer is still predominant due to the current publication format that uses text as its main way of communication. Solutions that look at how to represent this text in structured forms have been considered, by using various vocabularies and ontologies to applying natural language processing techniques for entity extraction and linking to approaches that allow the publication of datasets, metadata, code and other scientific resources together with the text.

A promising direction is the one we also consider in this thesis, where a different way of publishing, the "genuine semantic publishing" approach, makes use of some of general principles of the Web and the Semantic Web in order to publish from the start scientific contributions (whether they are related to data, scientific findings, or the scientific publishing process itself) in a machine-interpretable format that is accessible to humans as well. However, despite very promising, many aspects of this approach need to be further studied as there have not been many real-world scenarios where these techniques have been applied.

While there are quite a few challenges concerning scientific publishing, some of which we discuss above (like challenges related to reviewing, user interfaces, scientific findings, interoperability and transitions to new publication systems), the work presented in this thesis has the potential to bring some important changes to scientific publishing. Using semantic models like the one proposed for reviewing could be the start for allowing science to be accessible equally to machines, helping in building a comprehensible and machine learning graph that can, in turn, further help humans in their scientific endeavors and discoveries, one that is supported and helped by machines as well. The value that this can bring to the future can be considerable, as humans are reaching their capacity when it comes to processing the large amount of information that is made available every day.

In the end, we should to be open to embrace a world where we allow machines to contribute to the scientific endeavors of the research community, preparing for a future where artificial intelligence and humans can work together to bring forth new scientific discoveries. This can be a first step in the direction of involving machines more in the research community, where a new way of publishing that considers the Semantic Web principles can help in creating a society where computers can assist the Open Science research community become more efficient and make scientific contributions FAIR: easier to find, more accessible, interoperable and reusable.



A APPENDIX

The detailed timeline of the field study that encompasses the special issue with formalization papers at the Data Science journal from Chapter 5 is shown in Figure 20.

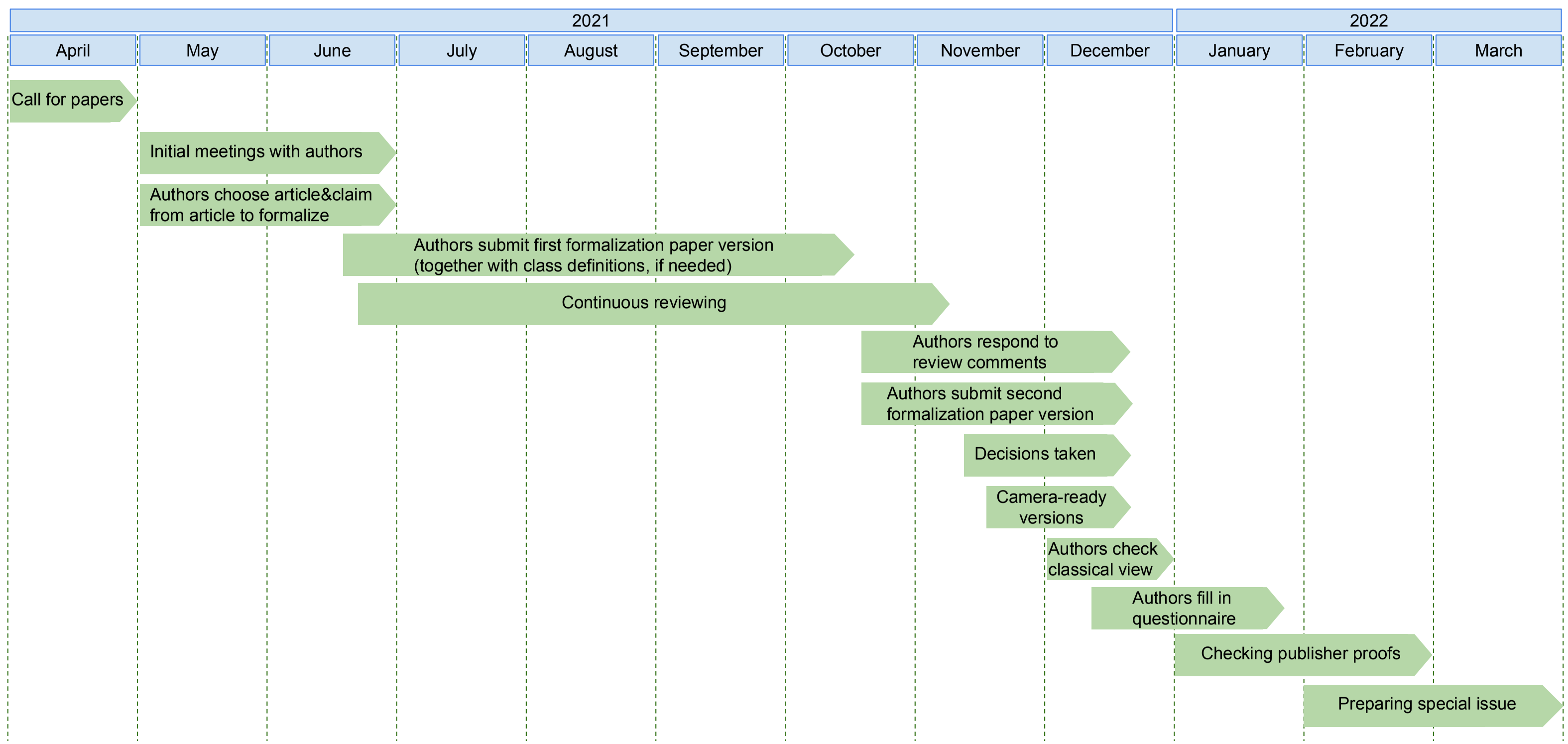


Figure 20: Timeline publication for the special issue with formalization papers at the Data Science journal.

BIBLIOGRAPHY

[1] Alexandru Constantin et al. "The Document Components Ontology (DoCO)." In: *Semantic Web* 7.2 (2016), pp. 167–181. DOI: 10.3233/SW-150177.

[2] Bruce Alberts. "Impact Factor Distortions." In: *Science* 340.6134 (2013), pp. 787–787. DOI: 10.1126/science.1240319.

[3] Brian S. Alper, Jason A. Hand, Susan G. Elliott, Scott Kinkade, Michael J. Hauan, Daniel K. Onion, and Bernard M. Sklar. "How Much Effort is Needed to Keep Up with the Literature Relevant for Primary Care?" In: *Journal of the Medical Library Association* 92.4 (2004), pp. 429–437. URL: https://ncbi.nlm.nih.gov/pmc/articles/PMC521514.

[4] M.A. Angrosh, Stephen Cranefield, and Nigel Stanger. "Contextual Information Retrieval in Research Articles: Semantic Publishing Tools for the Research Community." In: *Semantic Web* 5.4 (2014), pp. 261–293. DOI: 10.3233/SW-130097.

[5] Andrea Bagnacani, Paolo Ciancarini, Angelo Di Iorio, Andrea Giovanni Nuzzolese, Silvio Peroni, and Fabio Vitali. "The Semantic Lancet Project: A Linked Open Dataset for Scholarly Publishing." In: *Proceedings of the 19th International Conference on Knowledge Engineering and Knowledge Management Satellite Events*. EKAW'14. 2014, pp. 101–105. DOI: 10.1007/978-3-319-17966-7_10.

[6] Monya Baker. "1,500 Scientists Lift the Lid on Reproducibility." In: *Nature* 7604.533 (2016). DOI: 10.1038/533452a.

[7] Molood Barati, Quan Bai, and Qing Liu. "SWARM: An Approach for Mining Semantic Association Rules from Semantic Web Data." In: *Proceedings of the 14th Pacific Rim International Conference on Artificial Intelligence*. Vol. 9810. PRICAI'16. 2016, pp. 30–43. DOI: 10.1007/978-3-319-42911-3_3.

[8] Jonathan B.L. Bard and Seung Y. Rhee. "Ontologies in Biology: Design, Applications and Future Challenges." In: *Nature Reviews Genetics* 5.3 (2004), pp. 213–222. DOI: 10.1038/nrg1295.

[9] Elizabeth Beam, Christopher Potts, Russell A. Poldrack, and Amit Etkin. "A Data-Driven Framework for Mapping Domains of Human Neurobiology." In: *Nature Neuroscience* 24.12 (2021), pp. 1733–1744. DOI: 10.1038/s41593-021-00948-9.

[10] Sean Bechhofer et al. "Why Linked Data is Not Enough for Scientists." In: *Future Generation Computer Systems* 29.2 (2013), pp. 599–611. DOI: 10.1016/j.future.2011.08.004.

[11] Khalid Belhajjame et al. "Using a Suite of Ontologies for Preserving Workflow-Centric Research Objects." In: *Journal of Web Semantics* 32 (2015), pp. 16–42. DOI: 10.1016/j.websem.2015.01.003.

[12] Wim G.G. Benda and Tim C.E. Engels. "The Predictive Validity of Peer Review: A Selective Review of the Judgmental Forecasting Qualities of Peers, and Implications for Innovation in Science." In: *International Journal of Forecasting* 27.1 (2011), pp. 166–182. DOI: 10.1016/j.ijforecast.2010.03.003.

[13] Tim Berners-Lee and James Hendler. "Publishing on the Semantic Web." In: *Nature* 410 (2001), pp. 1023–1024. DOI: 10.1038/35074206.

[14] Camila Bezerra, Fred Freitas, and Filipe Santana. "Evaluating Ontologies with Competency Questions." In: *Proceedings of the ACM International Joint Conferences on Web Intelligence and Intelligent Agent Technologies*. Vol. 3. WI+IAT'13. 2013, pp. 284–285. DOI: 10.1109/WI-IAT.2013.199.

[15] Shitij Bhargava, Tsung-Ting Kuo, Ankit Goyal, Vincent Kuri, Gordon Lin, and Chun-Nan Hsu. "bioPDFX: Preparing PDF Scientific Articles for Biomedical Text Mining." In: *PeerJ Prepr.* 5 (2017). DOI: 10.7287/peerj.preprints.2993v1.

[16] Christian Bizer, Tom Heath, and Tim Berners-Lee. "Linked Data - The Story so Far." In: *International Journal on Semantic Web and Information* 5.3 (2009), pp. 1–22. DOI: 10.4018/jswis.2009081901.

[17] John Bohannon. "Who's Afraid of Peer Review?" In: *Science* 342.6154 (2013), pp. 60–65. DOI: 10.1126/science.342.6154.60.

[18] Arthur Brack, Jennifer D'Souza, Anett Hoppe, Sören Auer, and Ralph Ewerth. "Domain-Independent Extraction of Scientific Concepts from Research Articles." In: *Proceedings of the 42nd European Conference on Information Retrieval*. ECIR'20. 2020, pp. 251–266. DOI: 10.1007/978-3-030-45439-5_17.

[19] Jan Brase. "DataCite - A Global Registration Agency for Research Data." In: *Proceedings of the 4th International Conference on Cooperation and Promotion of Information Resources in Science and Technology*. 2009, pp. 257–261. DOI: 10.1109/COINFO.2009.66.

[20] Björn Brembs, Philippe Huneman, Felix Schönbrodt, Gustav Nilsonne, Toma Susi, Renke Siems, Pandelis Perakakis, Varvara Trachana, Lai Ma, and Sara Rodriguez-Cuadrado. "Replacing Academic Journals." In: *Royal Society Open Science* 10.7 (2023). DOI: 10.1098/rsos.230206.

[21] Cristina-Iulia Bucur. "Linkflows: Enabling a Web of Linked Semantic Publishing Workflows." In: *Proceedings of the 15th European Semantic Web Conference Satellite Events*. ESWC'18. 2018, pp. 262–271. DOI: 10.1007/978-3-319-98192-5_45.

[22] Cristina-Iulia Bucur, Tobias Kuhn, and Davide Ceolin. "Peer Reviewing Revisited: Assessing Research with Interlinked Semantic Comments." In: *Proceedings of the 10th International Conference on Knowledge Capture*. K-CAP'19. 2019, pp. 179–187. DOI: 3360901.3364434.

[23] Cristina-Iulia Bucur, Tobias Kuhn, Davide Ceolin, and Jacco van Ossenbruggen. "A Unified Nanopublication Model for Effective and User-Friendly Access to the Elements of Scientific Publishing." In: *Proceedings of the 22nd International Conference on Knowledge Engineering and Knowledge Management*. Vol. 12387. EKAW'20. 2020, pp. 104–119. DOI: 10.1007/978-3-030-61244-3_7.

[24] Cristina-Iulia Bucur, Tobias Kuhn, Davide Ceolin, and Jacco van Ossenbruggen. "Expressing High-Level Scientific Claims with Formal Semantics." In: *Proceedings of the 11th Knowledge Capture Conference*. K-CAP'21. 2021, pp. 233–240. DOI: 10.1145/3460210.3493561.

[25] Cristina-Iulia Bucur, Tobias Kuhn, Davide Ceolin, and Jacco van Ossenbruggen. "Nanopublication-Based Semantic Publishing and Reviewing: A Field Study with Formalization Papers." In: *PeerJ Computer Science* 9 (2023). DOI: 10.7717/peerj-cs.1159.

[26] Sarven Capadisli, Amy Guy, Christoph Lange, Sören Auer, Andrei Vlad Sambra, and Tim Berners-Lee. "Linked Data Notifications: A Resource-Centric Communication Protocol." In: *Proceedings of the 14th European Semantic Web Conference*. ESWC'17. 2017, pp. 537–553. DOI: 10.1007/978-3-319-58068-5_33.

[27] Sarven Capadisli, Amy Guy, Ruben Verborgh, Christoph Lange, Sören Auer, and Tim Berners-Lee. "Decentralised Authoring, Annotations and Notifications for a Read-Write Web with dokieli." In: *Proceedings of the 17th International Conference on Web Engineering*. ICWE'17. 2017, pp. 469–481. DOI: 10.1007/978-3-319-60131-1_33.

[28] Alessandro Checco, Lorenzo Bracciale, Pierpaolo Loreti, Stephen Pinfield, and Giuseppe Bianchi. "AI-Assisted Peer Review." In: *Humanities and Social Sciences Communications* 8.1 (2021). DOI: 10.1057/s41599-020-00703-8.

[29] Yang Chi, Yue Qin, Rui Song, and Haoyue Xu. "Knowledge Graph in Smart Education: A Case Study of Entrepreneurship Scientific Publication Management." In: *Sustainability* 10.4 (2018), pp. 1–21. DOI: 10.3390/su10040995.

[30] Marcus C. Chibucos, Christopher J. Mungall, Rama Balakrishnan, Karen R. Christie, Rachael P. Huntley, Owen White, Judith A. Blake, Suzanna E. Lewis, and Michelle Giglio. "Standardized Description of Scientific Evidence Using the Evidence Ontology (ECO)." In: *Database: The Journal of Biological Databases and Curation* v.2014 (2014). DOI: 10.1093/database/bau075.

[31] Paolo Ciancarini, Angelo Di Iorio, Andrea Giovanni Nuzzolese, Silvio Peroni, and Fabio Vitali. "Semantic Annotation of Scholarly Documents and Citations." In: *Congress of the Italian Association for Artificial Intelligence, Advances in Artificial Intelligence, Lecture Notes in Computer Science*. Vol. 8249. AI*IA'13. 2013, pp. 336–347. DOI: 10.1007/978-3-319-03524-6_29.

[32] Paolo Ciccarese, Marco Ocana, and Tim Clark. "Open Semantic Annotation of Scientific Publications Using DOMEO." In: *Journal of Biomedical Semantics* 3.1 (2012). DOI: 10.1186/2041-1480-3-S1-S1.

[33] Paolo Ciccarese, Marco Ocana, Leyla Jael Garcia Castro, Sudeshna Das, and Tim Clark. "An Open Annotation Ontology for Science on Web 3.0." In: *Journal of Biomedical Semantics* 2 (2011). DOI: 10.1186/2041-1480-2-S2-S4.

[34] Tim Clark, Paolo N. Ciccarese, and Carole A. Goble. "Micropublications: A Semantic Model for Claims, Evidence, Arguments and Annotations in Biomedical Communications." In: *Journal of Biomedical Semantics* 5.28 (2014). DOI: 10.1186/2041-1480-5-28.

[35] Joseph Paul Cohen and Henry Z. Lo. "Academic Torrents: A Community-Maintained Distributed Repository." In: *Proceedings of the 2014 Annual Conference on Extreme Science and Engineering Discovery Environment*. 2014. DOI: 10.1145/2616498.2616528.

[36] Sergio Copiello. "On the Skewness of Journal Self-Citations and Publisher Self-Citations: Cues for Discussion from a Case Study." In: *Learned Publishing* 32.3 (2019), pp. 249–258. DOI: 10.1002/leap.1235.

[37] Adrien Coulet, Yael Garten, Michel Dumontier, Russ B. Altman, Mark A. Musen, and Nigam H. Shah. "Integration and Publication of Heterogeneous Text-Mined Relationships on the Semantic Web." In: *Journal of Biomedical Semantics* 2 Suppl 2 (2011). DOI: 10.1186/2041-1480-2-S2-S10.

[38] Jeremy Debattista, Christoph Lange, and Sören Auer. "Representing Dataset Quality Metadata Using Multi-Dimensional Views." In: *Proceedings of the 10th International Conference on Semantic Systems*. SEMANTiCS'14. 2014, pp. 92–99. DOI: 10.1145/2660517.2660525.

[39] Emek Demir et al. "The BioPAX Community Standard for Pathway Data Sharing." In: *Nature Biotechnology* 28.9 (2010), pp. 935–942. DOI: 10.1038/nbt.1666.

[40] Angelo Di Iorio, Silvio Peroni, Fabio Vitali, and Jacopo Zingoni. "Semantic Lenses to Bring Digital and Semantic Publishing Together." In: *Proceedings of the 4th International Conference on Linked Science*. Vol. 1282. LISC'14. 2014, pp. 12–23. DOI: 10.5555/2878584.2878586.

[41] Eleftherios P. Diamandis. "The Current Peer Review System is Unsustainable-Awaken the Paid Reviewer Force!" In: *Clinical biochemistry* 50.9 (2017). DOI: 10.1016/j.clinbiochem.2017.02.019.

[42] "Dissecting Our Impact Factor." In: *Nature Materials* 645.10 (2011). DOI: 10.1038/nmat3114.

[43] Daniel Domingo-Fernándéz, Charles Tapley Hoyt, Carlos Bobis-Álvarez, Josep Marín-Llaó, and Martin Hofmann-Apitius. "ComPath: An Ecosystem for Exploring, Analyzing, and Curating Mappings Across Pathway Databases." In: *NPJ Systems Biology and Applications* 4 (2018). DOI: 10.1038/s41540-018-0078-8.

[44] Peng Dong, Marie Loh, and Adrian Mondry. "The 'Impact Factor' Revisited." In: *Biomedical Digital Libraries* 2.7 (2005). DOI: 10.1186/1742-5581-2-7.

[45] Oren Etzioni, Michael Cafarella, Doug Downey, Ana-Maria Popescu, Tal Shaked, Stephen Soderland, Daniel S. Weld, and Alexander Yates. "Unsupervised Named-Entity Extraction from the Web: An Experimental Study." In: *Artificial Intelligence* 165.1 (2005), pp. 91–134. DOI: 10.1016/j.artint.2005.03.001.

[46] Michael Färber and David Lamprecht. "The Data Set Knowledge Graph: Creating a Linked Open Data Source for Data Sets." In: *Quantitative Science Studies* 2.4 (2021), pp. 1324–1355. DOI: 10.1162/qss_a_00161.

[47] Said Fathalla, Sören Auer, and Christoph Lange. "Towards the Semantic Formalization of Science." In: *Proceedings of the 35th Annual ACM Symposium on Applied Computing*. SAC '20. 2020, pp. 2057–2059. DOI: 10.1145/3341105.3374132.

[48] Robert A. Felix and Margery A. Barrand. "P-glycoprotein Expression in Rat Brain Endothelial Cells: Evidence for Regulation by Transient Oxidative Stress." In: *Journal of Neurochemistry* 80.1 (2002), pp. 64–72. DOI: j.0022-3042.2001.00660.x.

[49] Imen Filali, Francesco Bongiovanni, Fabrice Huet, and Françoise Baude. "A Survey of Structured P2P Systems for RDF Data Storage and Retrieval." In: *Transactions on Large-Scale Data- and Knowledge-Centered Systems, LNCS* 6790.3 (2011), pp. 20–55. DOI: 10.1007/978-3-642-23074-5_2.

[50] Joseph L. Fleiss. "Measuring Nominal Scale Agreement Among Many Raters." In: *Psychological Bulletin* 76.5 (1971), pp. 378–382. DOI: 10.1037/h0031619.

[51] Aldo Gangemi, Silvio Peroni, David Shotton, and Fabio Vitali. "The Publishing Workflow Ontology (PWO)." In: *Semantic Web* 8.5 (2017), pp. 703–718. DOI: 10.3233/SW-160230.

[52] Leyla Jael García Castro, Rafael Berlanga, Dietrich Rebholz-Schuhmann, and Garcia Alexander. "Connections Across Scientific Publications Based on Semantic Annotations." In: *Proceedings of the 3rd Workshop on Semantic Publishing*. Vol. 994. SePublica'13. 2013, pp. 51–62. DOI: 10.5167/UZH-82214.

[53] Eugene Garfield. "Journal Impact Factor: A Brief Review." In: *Canadian Medical Association Journal (CMAJ)* 161.8 (1999), pp. 979–980. URL: www.cmaj.ca/content/161/8/979.

[54] Eugene Garfield. "The History and Meaning of the Journal Impact Factor." In: *Journal of the American Medical Association (JAMA)* 295.1 (2006), pp. 90–93. DOI: 10.1001/jama.295.1.90.

[55] Daniel Garijo and María Poveda-Villalón. "Best Practices for Implementing FAIR Vocabularies and Ontologies on the Web." In: *Applications and Practices in Ontology Design, Extraction, and Reasoning*. 2020. DOI: 10.3233/SSW200034.

[56] Fabio Giachelle, Dennis Dosso, and Gianmaria Silvello. "Search, Access, and Explore Life Science Nanopublications on the Web." In: *PeerJ Computer Science* 7 (2021). DOI: 10.7717/peerj-cs.335.

[57] Alec Go, Richa Bhayani, and Lei Huang. "Twitter sentiment classification using distant supervision." In: *Stanford CS224N Project Report* (2009). URL: https://cs.stanford.edu/people/alecmgo/papers/TwitterDistantSupervision09.pdf.

[58] Paul Groth, Andrew Gibson, and Jan Velterop. "The Anatomy of a Nanopublication." In: *Information Services and Use* 30.1-2 (2010), pp. 51–56. DOI: 10.3233/ISU-2010-0613.

[59] Harry Halpin, Patrick J. Hayes, James P. McCusker, Deborah L. McGuinness, and Henry S. Thompson. "When owl:sameAs Isn't the Same: An Analysis of Identity in Linked Data." In: *Proceeding of the 9th Internationl Semantic Web Conference*. ISWC'10. 2010, pp. 305–320. DOI: 10.1007/978-3-642-17746-0_20.

[60] Lance M. Hannestad, Vlado Dancík, Meera Godden, Imelda W. Suen, Kenneth C. Huellas-Bruskiewicz, Benjamin M. Good, Chris J. Mungall, and Richard M. Bruskiewich. "Knowledge Beacons: Web Services for Data Harvesting of Distributed Biomedical Knowledge." In: *PLoS ONE* 16.3 (2021). DOI: 10.1371/journal.pone.0231916.

[61] Pascal Hitzler and Frank van Harmelen. "A Reasonable Semantic Web." In: *Semantic Web* 1.1-2 (2010), pp. 39–44. DOI: 10.3233/SW-2010-0010.

[62] Charles Tapley Hoyt, Daniel Domingo-Fernándéz, and Martin Hofmann-Apitius. "BEL Commons: An Environment for Exploration and Analysis of Networks Encoded in Biological Expression Language." In: *Database: The Journal of Biological Databases and Curation* 2018 (2018). DOI: 10.1093/database/bay126.

[63] Charles Tapley Hoyt et al. "Re-curation and Rational Enrichment of Knowledge Graphs in Biological Expression Language." In: *Database: The Journal of Biological Databases and Curation* (2019). DOI: 10.1093/database/baz068.

[64] M. Hucka et al. "The Systems Biology Markup Language (SBML): A Medium for Representation and Exchange of Biochemical Network Models." In: *Bioinformatics* 19.4 (2003), pp. 524–531. DOI: 10.1093/bioinformatics/btg015.

[65] Eero Hyvönen. "Publishing and Using Cultural Heritage Linked Data on the Semantic Web." In: *Synthesis Lectures on Data, Semantics, and Knowledge*. SLDSK'12. 2012. DOI: 10.1007/978-3-031-79438-4.

[66] Eero Hyvönen. "Using the Semantic Web in Digital Humanities: Shift from Data Publishing to Data-Analysis and Serendipitous Knowledge Discovery." In: *Semantic Web* 11.1 (2020), pp. 187–193. DOI: 10.3233/SW-190386.

[67] Bryon Jacob and Jonathan Ortiz. "Data.world: A Platform for Global-Scale Semantic Publishing." In: *Proceedings of the ISWC Posters and Demonstrations and Industry Tracks co-located with 16th International Semantic Web Conference*. Vol. 1963. ISWC'17. 2017. URL: https://ceur-ws.org/Vol-1963/paper605.pdf.

[68] Prateek Jain, Pascal Hitzler, Peter Z. Yeh, Kunal Verma, and Amit P. Sheth. "Linked Data Is Merely More Data." In: *Proceedings of the Association for the Advancement of AI Spring Symposium*. AAAI*SS'10. 2010, pp. 82–86. URL: https://aaai.org/papers/01130-1130-linked-data-is-merely-more-data.

[69] Krzysztof Janowicz and Pascal Hitzler. "Open and Transparent: The Review Process of the Semantic Web Journal." In: *Learned Publishing* 25 (2012), pp. 48–55. DOI: 10.1087/20120107.

[70] Mohamad Yaser Jaradeh, Allard Oelen, Kheir Eddine Farfar, Manuel Prinz, Jennifer D'Souza, Gábor Kismihók, Markus Stocker, and Sören Auer. "Open Research Knowledge Graph: Next Generation Infrastructure for Semantic Scholarly Knowledge." In: *Proceedings of the 10th International Conference on Knowledge Capture*. K-CAP'19. 2019, pp. 243–246. DOI: 10.1145/3360901.3364435.

[71] Sumaiya Kabir, Shamim Ripon, Mamunur Rahman, and Tanjim Rahman. "Knowledge-based Data Mining Using Semantic Web." In: *Proceedings of the International Conference on Applied Computing, Computer Science, and Computer Engineering*. ICACC'13. 2013. DOI: 10.1016/j.ieri.2014.08.018.

[72] Tomi Kauppinen, Alkyoni Baglatzi, and Carsten Keßler. "Linked science: Interconnecting Scientific Assets." In: *Data-Intensive Science*. 2016, pp. 383–400. URL: www.oreilly.com/library/view/data-intensive-science/9781439881415/xhtml/27_Chapter15.xhtml#sec15_1.

[73] Sepehr Golriz Khatami, Daniel Domingo-Fernándéz, Sarah Mubeen, Charles Tapley Hoyt, Christine Robinson, Reagon Karki, Anandhi Iyappan, Alpha Tom Kodamullil, and Martin Hofmann-Apitius. "A Systems Biology Approach for Hypothesizing the Effect of Genetic Variants on Neuroimaging Features in Alzheimer's Disease." In: *Journal of Alzheimer's Disease* 80.2 (2021), pp. 831–840. DOI: 10.3233/JAD-201397.

[74] Tobias Kiesslich, Silke B. Weineck, and Dorothea Koelblinger. "Reasons for Journal Impact Factor Changes: Influence of Changing Source Items." In: *PLoS ONE* 11.4 (2016). DOI: 10.1371/journal.pone.0154199.

[75] Yasmine Kotturi, Andrew Du, Scott Klemmer, and Chinmay Kulkarni. "Long-Term Peer Reviewing Effort is Anti-Reciprocal." In: *Proceedings of the 4th ACM Conference on Learning at Scale*. L@S'17. 2017, pp. 279–282. DOI: 10.1145/3051457.3054004.

[76] Tobias Kuhn. "nanopub-java: A Java Library for Nanopublications." In: *Proceedings of the 5th Workshop on Linked Science*. Vol. 1572. LISC'15. 2015, pp. 19–25. URL: https://ceur-ws.org/Vol-1572/paper4.pdf.

[77] Tobias Kuhn. "Using the AIDA Language to Formally Organize Scientific Claims." In: *Proceedings of the 6th International Workshop on Controlled Natural Language*. Vol. 304. CNL'18. 2018, pp. 52–60. DOI: 10.3233/978-1-61499-904-1-52.

[78] Tobias Kuhn, Paolo Emilio Barbano, Mate Levente Nagy, and Michael Krauthammer. "Broadening the Scope of Nanopublications." In: *Proceedings of the 10th Extended Semantic Web Conference*. ESWC'13. 2013, pp. 487–501. DOI: 10.1007/978-3-642-38288-8_33.

[79] Tobias Kuhn, Christine Chichester, Michael Krauthammer, Núria Queralt-Rosinach, Ruben Verborgh, George Giannakopoulos, Axel-Cyrille Ngonga Ngomo, Raffaele Viglianti, and Michel Dumontier. "Decentralized Provenance-Aware Publishing with Nanopublications." In: *PeerJ Computer Science* 2 (2016). DOI: 10.7717/peerj-cs.78.

[80] Tobias Kuhn and Michel Dumontier. "Trusty URIs: Verifiable, Immutable, and Permanent Digital Artifacts for Linked Data." In: *Proceedings of the 11th Extended Semantic Web Conference*. Vol. 8465. ESWC'14. 2014, pp. 395–410. DOI: 10.1007/978-3-319-07443-6_27.

[81] Tobias Kuhn and Michel Dumontier. "Making Digital Artifacts on the Web Verifiable and Reliable." In: *IEEE Transactions on Knowledge and Data Engineering* 27.9 (2015), pp. 2390–2400. DOI: 10.1109/TKDE.2015.2419657.

[82] Tobias Kuhn and Michel Dumontier. "Genuine Semantic Publishing." In: *Data Science* 1-2 (2017), pp. 139–154. DOI: 10.3233/DS-170010.

[83] Tobias Kuhn, Ruben Taelman, Vincent Emonet, Haris Antonatos, Stian Soiland-Reyes, and Michel Dumontier. "Semantic Micro-Contributions with Decentralized Nanopublication Services." In: *PeerJ Computer Science* 7 (2021). DOI: 10.7717/peerj-cs.387.

[84] Tobias Kuhn, Egon Willighagen, Chris Evelo, Núria Queralt-Rosinach, Emilio Centeno, and Laura I. Furlong. "Reliable Granular References to Changing Linked Data." In: *Proceedings of the 16th International Semantic Web Conference*. Vol. 10587. ISWC'17. 2017. DOI: 10.1007/978-3-319-68288-4_26.

[85] Tobias Kuhn et al. "Nanopublications: A Growing Resource of Provenance-Centric Scientific Linked Data." In: *Proceedings of the 14th IEEE International Conference on e-Science*. e-Science'18. 2018, pp. 83–92. DOI: 10.1109/eScience.2018.00024.

[86] Esther Landhuis. "Scientific Literature: Information Overload." In: *Nature* 535 (2016), pp. 457–458. DOI: 10.1038/nj7612-457a.

[87] Richard J. Landis and Gary G. Koch. "The Measurement of Observer Agreement for Categorical Data." In: *Biometrics* 33.1 (1977), pp. 159–174. DOI: 10.2307/2529310.

[88] Vincent Lariviére and Cassidy Sugimoto. "The Journal Impact Factor: A Brief History, Critique, and Discussion of Adverse Effects." In: *Springer Handbook of Science and Technology Indicators*. 2019, pp. 3–24. DOI: 10.1007/978-3-030-02511-3_1.

[89] Carole J. Lee, Cassidy R. Sugimoto, Guo Zhang, and Blaise Cronin. "Bias in Peer Review." In: *Journal of the American Society for Information Science and Technology* 64.1 (2012), pp. 2–17. DOI: 10.1002/asi.22784.

[90] Daniel Leung, Rob Law, Deniz Kucukusta, and Basak Denizci Guillet. "How to Review Journal Manuscripts: A Lesson Learnt from the World's Excellent Reviewers." In: *Tourism Management Perspectives* 10 (2014), pp. 46–56. DOI: 10.1016/j.tmp.2014.01.003.

[91] Faina Linkov, Mita Lovalekar, and Ronald LaPorte. "Scientific Journals are ''Faith Based": Is there Science Behind Peer Review?" In: *Journal of the Royal Society of Medicine* 99.12 (2006), pp. 596–598. DOI: 10.1258/jrsm.99.12.596.

[92] Giuseppe Lippi, Philippe Gillery, Karl J. Lackner, Bohuslav Melichar, Deborah A. Payne, Peter Schlattmann, Jill R. Tate, and Mario Plebani. "Scientific Publishing in the "Predatory" Era." In: *Clinical Chemistry and Laboratory Medicine (CCLM)* 56.5 (2018), pp. 683–684. DOI: 10.1515/cclm-2017-1079.

[93] Pasquale Lisena, Albert Meroño-Peñuela, Tobias Kuhn, and Raphaël Troncy. "Easy Web API Development with SPARQL Transformer." In: *Proceedings of The 18th International Semantic Web Conference*. ISWC'19. 2019, pp. 454–470. DOI: 10.1007/978-3-030-30796-7_28.

[94] Stephen Lock. *A Difficult Balance: Editorial Peer Review in Medicine*. Nuffield Trust, 1985. URL: www.nuffieldtrust.org.uk/sites/default/files/2017-01/a-difficult-balance-web-final.pdf.

[95] Sumit Madan, Justyna Szostak, Ravikumar Komandur Elayavilli, Richard Tzong-Han Tsai, Mehdi Ali, Longhua Qian, Majid Rastegar-Mojarad, Julia Hoeng, and Juliane Fluck. "The Extraction of Complex Relationships and Their Conversion to Biological Expression Language (BEL) Overview of the BioCreative VI (2017) BEL Track." In: *Database: The Journal of Biological Databases and Curation* (2019). DOI: 10.1093/database/baz084.

[96] Chris Markman and Constantine Zavras. "BitTorrent and Libraries: Cooperative Data Publishing, Management and Discovery." In: *D-Lib Magazine* 20.3-4 (2014). DOI: 10.1045/march2014-markman.

[97] David Mazières and Frans M. Kaashoek. "Escaping the Evils of Centralized Control with Self-Certifying Pathnames." In: *Proceedings of the 8th ACM SIGOPS European workshop on Support for composing distributed applications*. ACM SIGOPS'98. 1998, pp. 118–125. DOI: 10.1145/319195.319213.

[98] Bruce McGregor. "Facets and Hierarchies in Scientific Search." In: *Journal of Electronic Publishing* 11.2 (2008). DOI: 10.3998/3336451.0011.205.

[99] Marcia K. McNutt et al. "Transparency in Authors' Contributions and Responsibilities to Promote Integrity in Scientific Publication." In: *Proceedings of the National Academy of Sciences of the United States of America* 115.11 (2018), pp. 2557–2560. DOI: 10.1073/pnas.1715374115.

[100] Albert Meroño-Peñuela and Rinke Hoekstra. "grlc Makes GitHub Taste Like Linked Data APIs." In: *Proceedings of the 13th European Semantic Web Conference*. ESWC'16. 2016, pp. 342–353. DOI: 10.1007/978-3-319-47602-5_48.

[101] Jill P. Mesirov. "Accessible Reproducible Research." In: *International Journal on Semantic Web and Information* (2010), pp. 415–416. DOI: 10.1126/science.1179653.

[102] Silvia Mirri, Silvio Peroni, Paola Salomoni, Fabio Vitali, and Vincenzo Rubano. "Towards Accessible Graphs in HTML-based Scientific Articles." In: *Proceedings of the 14th IEEE Annual Consumer Communications & Networking Conference*. CCNC'17. 2017, pp. 1067–1072. DOI: 10.1109/CCNC.2017.7983287.

[103] Barend Mons. "Which Gene did You Mean?" In: *BMC Bioinformatics* 6.1 (2005). DOI: 10.1186/1471-2105-6-142.

[104] Tareq Al-Moslmi, Marc Gallofré Ocaña, Andreas L. Opdahl, and Csaba Veres. "Named Entity Extraction for Knowledge Graphs: A Literature Overview." In: *IEEE Access* 8 (2020), pp. 32862–32881. DOI: 10.1109/ACCESS.2020.2973928.

[105] H.M. Müller, K.M. Van Auken, Y. Li, and P.W. Sternberg. "Textpresso Central: A Customizable Platform for Searching, Text Mining, Viewing, and Curating Biomedical Literature." In: *BMC Bioinformatics* 19.94 (2018). DOI: 10.1186/s12859-018-2103-8.

[106] Tobias Opthof. "Sense and Nonsense about the Impact Factor." In: *Cardiovascular Research* 33.1 (1997), pp. 1–7. DOI: 10.1016/S0008-6363(96)00215-5.

[107] Vassilis Papakonstantinou, Irini Fundulaki, and Giorgos Flouris. "Assessing Linked Data Versioning Systems: The Semantic Publishing Versioning Benchmark." In: *Proceedings of the 12th International Workshop on Scalable Semantic Web Knowledge Base Systems*. SSWS'18. 2018. URL: https://ceur-ws.org/Vol-2179/SSWS2018_paper4.pdf.

[108] Lyubomir Penev, Donat Agosti, Teodor Georgiev, Viktor Senderov, Guido Sautter, Terry Catapano, and Pavel Stoev. "The Open Biodiversity Knowledge Management (eco-)System: Tools and Services for Extraction, Mobilization, Handling and Reuse of Data from the Published Literature." In: *Biodiversity Information Science and Standards*. Vol. 2. 2018. DOI: 10.3897/biss.2.25748.

[109] Lyubomir Penev, Terry Catapano, Donat Agosti, Teodor Georgiev, Guido Sautter, and Pavel Stoev. "Implementation of TaxPub, an NLM DTD Extension for Domain-Specific Markup in Taxonomy, from the Experience of a Biodiversity Publisher."

In: *Proceedings of the Journal Article Tag Suite Conference*. JATS-Con'12. 2012. URL: www.ncbi.nlm.nih.gov/books/NBK100351.

[110] Lyubomir Penev, Mariya Dimitrova, Viktor Senderov, Georgi Zhelezov, Teodor Georgiev, Pavel Stoev, and Kiril Ivanov Simov. "OpenBiodiv: A Knowledge Graph for Literature-Extracted Linked Open Data in Biodiversity Science." In: *Publications* 7.2 (2019). DOI: 10.3390/publications7020038.

[111] Martha O. Perez-Arriaga. "Automated Development of Semantic Data Models Using Scientific Publications." PhD thesis. University of New Mexico, 2018. URL: https://core.ac.uk/download/pdf/160274576.pdf.

[112] Silvio Peroni. "Semantic Publishing: Issues, Solutions and New Trends in Scholarly Publishing within the Semantic Web Era." PhD thesis. Universita di Bologna, 2012. DOI: 10.6092/UNIBO/AMSDOTTORATO/4766.

[113] Silvio Peroni. "The Digital Publishing Revolution." In: *Semantic Web Technologies and Legal Scholarly Publishing*. Vol. 15. Law, Governance and Technology Series. 2014, pp. 7–43. DOI: 10.1007/978-3-319-04777-5_2.

[114] Silvio Peroni. "The Semantic Publishing and Referencing Ontologies." In: *Semantic Web Technologies and Legal Scholarly Publishing*. Vol. 15. Law, Governance and Technology Series. 2014, pp. 121–193. DOI: 10.1007/978-3-319-04777-5_5.

[115] Silvio Peroni. "Automating Semantic Publishing." In: *Data Science* 1.1-2 (2017), pp. 155–173. DOI: 10.3233/DS-170012.

[116] Silvio Peroni, Enrico Motta, and Mathieu d'Aquin. "Identifying Key Concepts in an Ontology, through the Integration of Cognitive Principles with Statistical and Topological Measures." In: *Proceedings of the 3rd Asian Semantic Web Conference*. Vol. 5367. ASWC'08. 2008, pp. 242–256. DOI: 10.1007/978-3-540-89704-0_17.

[117] Silvio Peroni, Francesco Osborne, Angelo Di Iorio, Andrea Giovanni Nuzzolese, Francesco Poggi, Fabio Vitali, and Enrico Motta. "Research Articles in Simplified HTML: A Web-First Format for HTML-Based Scholarly Articles." In: *PeerJ Computer Science* 3 (2017). DOI: 10.7717/peerj-cs.132.

[118] Silvio Peroni and David Shotton. "FaBiO and CiTO: Ontologies for Describing Bibliographic Resources and Citations." In: *Journal of Web Semantics* 17 (2012), pp. 33–43. DOI: 10.1016/j.websem.2012.08.001.

[119] Silvio Peroni and David Shotton. "The SPAR Ontologies." In: *Proceedings of the 17th International Semantic Web Conference*. Vol. 128. ISWC'18. 2018, pp. 119–136. DOI: 10.1007/978-3-030-00668-6_8.

[120] Silvio Peroni, David Shotton, and Fabio Vitali. "Scholarly Publishing and Linked Data: Describing Roles, Statuses, Temporal and Contextual Extents." In: *Proceedings of the 8th International Conference on Semantic Systems*. I-SEMANTICS'12. 2012, pp. 9–16. DOI: 10.1145/2362499.2362502.

[121] Silvio Peroni, Francesca Tomasi, Fabio Vitali, and Jacopo Zingoni. "Semantic lenses as exploration method for scholarly articles." In: *Italian Research Conference on Digital Libraries*. Vol. 385. IRCDL'13. 2013, pp. 118–129. DOI: 10.1007/978-3-642-54347-0_13.

[122] Jason Priem. "Beyond the Paper." In: *Nature* 495 (2013), pp. 437–440. DOI: 10.1038/495437a.

[123] Núria Queralt-Rosinach, Tobias Kuhn, Christine Chichester, Michel Dumontier, Ferran Sanz, and Laura I. Furlong. "Publishing DisGeNET as Nanopublications." In: *Semantic Web* 7.5 (2016), pp. 519–528. DOI: 10.3233/SW-150189.

[124] Untung Rahardja, Ninda Lutfiani, and Hega Lutfilah Juniar. "Scientific Publication Management Transformation in Disruption Era." In: *APTISI Transactions on Management (ATM)* 3.2 (2019), pp. 109–118. DOI: 10.33050/atm.v3i2.1008.

[125] Habeeb Ibrahim Abdula Razack, Sam T. Mathew, Fathinul Fikri Ahmada Saad, and Saleh A. Alqahtani. "Artificial Intelligence-Assisted Tools for Redefining the Communication Landscape of the Scholarly World." In: *Science Editing* 8.2 (2021), pp. 134–144. DOI: 10.6087/kcse.244.

[126] Filipe N. Ribeiro, Matheus Araújo, Pollyanna Gonçalves, Marcos André Gonçalves, and Fabrício Benevenuto. "SentiBench - A Benchmark Comparison of State-of-the-Practice Sentiment Analysis Methods." In: *EPJ Data Science* 5.23 (2016). DOI: 10.1140/epjds/s13688-016-0085-1.

[127] Peter Gordon Roetzel. "Information Overload in the Information Age: A Review of the Literature from Business Administration, Business Psychology, and Related Disciplines with a Bibliometric Approach and Framework Development." In: *Business Research* 12.2 (2019), pp. 479–522. DOI: 10.1007/s40685-018-0069-z.

[128] Afshin Sadeghi, Sarven Capadisli, Johannes Wilm, Christoph Lange, and Philipp Mayr. "Opening and Reusing Transparent Peer Reviews with Automatic Article Annotation." In: *Publications* (2019). DOI: 10.3390/publications7010013.

[129] Somnath Saha, Sanjay Saint, and Dimitri A. Christakis. "Impact Factor: A Valid Measure of Journal Quality?" In: *Journal of the Medical Library Association (JMLA)* 91.1 (2003). URL: https://pubmed.ncbi.nlm.nih.gov/12572533.

[130] Idafen Santana-Perez and María Poveda-Villalon. *FAIR* Reviews Ontology (FR)*. URL: http://purl.org/spar/fr.

[131] Bahar Sateli and René Witte. "From Papers to Triples: An Open Source Workflow for Semantic Publishing Experiments." In: *International Workshop on Semantics, Analytics, Visualization. Enhancing Scholarly Data*. Vol. 9792. SAVE-SD '16. 2016, pp. 39–44. DOI: 10.1007/978-3-319-53637-8_5.

[132] Jodi Schneider, Paolo Ciccarese, Tim Clark, and Richard David Boyce. "Using the Micropublications Ontology and the Open Annotation Data Model to Represent Evidence within a Drug-Drug Interaction Knowledge Base." In: *Proceedings of the 4th International Conference on Linked Science*. LISC'14. 2014, pp. 60–70. DOI: 10.5555/2878584.2878590.

[133] Robert Schulz et al. "Is the Future of Peer Review Automated?" In: *BMC Research Notes* 15.1 (2022), p. 203. DOI: 10.1186/s13104-022-06080-6.

[134] Per O. Seglen. "Why the Impact Factor of Journals Should Not be Used for Evaluating Research." In: *British Medical Journal (BMJ)* 314 (1997), pp. 498–502. DOI: 10.1136/bmj.314.7079.497.

[135] Viktor Senderov and Lyubomir Penev. "The Open Biodiversity Knowledge Management System in Scholarly Publishing." In: *Research Ideas and Outcomes* 2 (2016). DOI: 10.3897/rio.2.e7757.

[136] Pedro Sernadela, Pedro Lopes, David Campos, Sérgio Matos, and José Luís" Oliveira. "A Semantic Layer for Unifying and Exploring Biomedical Document Curation Results." In: *Proceedings of the 3rd International Conference on Bioinformatics and Biomedical Engineering*. IWBBIO'15. 2015, pp. 8–17. DOI: 10.1007/978-3-319-16483-0_2.

[137] Pedro Sernadela and José Luís Oliveira. "A Semantic-Based Workflow for Biomedical Literature Annotation." In: *Database: The Journal of Biological Databases and Curation* (2017). DOI: 10.1093/database/bax088.

[138] Nigel Shadbolt, Tim Berners-Lee, and Wendy Hall. "The Semantic Web Revisited." In: *IEEE Intelligent Systems* (2006). DOI: 10.1109/MIS.2006.62.

[139] Yifan Shao, Haoru Li, Jinghang Gu, Longhua Qian, and Guodong Zhou. "Extraction of Causal Relations Based on SBEL and BERT Model." In: *Database: The Journal of Biological Databases and Curation* 2021 (2021). DOI: 10.1093/database/baab005.

[140] David Shotton. "Semantic Publishing: The Coming Revolution in Scientific Journal Publishing." In: *Learned Publishing* 22.2 (2009), pp. 85–94. DOI: 10.1087/2009202.

[141] David Shotton. "The Five Stars of Online Journal Articles - A Framework for Article Evaluation." In: *D-Lib Magazine* 18.1/2 (2012), pp. 457–458. DOI: 10.1045/january2012-shotton.

[142] David Shotton, Katie Portwin, Graham Klyne, and Alistair Miles. "Adventures in Semantic Publishing: Exemplar Semantic Enhancements of a Research Article." In: *PLoS Computational Biology* 5.4 (2009). DOI: 10.1371/journal.pcbi.1000361.

[143] Mohamed Saleem Abdul Shukkoor, Kalpana Raja, and Mohamad Taufik Hidayat Baharuldin. "A Text Mining Protocol for Predicting Drug-Drug Interaction and Adverse Drug Reactions from PubMed Articles." In: *Methods in molecular biology* 2496 (2022), pp. 237–258. DOI: 10.1007/978-1-0716-2305-3_13.

[144] Leslie F. Sikos. "Knowledge Representation with Semantic Web Standards." In: *Description Logics in Multimedia Reasoning*. 2017, pp. 11–49. DOI: 10.1007/978-3-319-54066-5_2.

[145] Ted Slater. "Recent Advances in Modeling Languages for Pathway Maps and Computable Biological Networks." In: *Drug Discovery Today* 19.2 (2014), pp. 193–198. DOI: 10.1016/j.drudis.2013.12.011.

[146] Ted Slater and Diane H. Song. "Saved by the BEL: Ringing in a Common Language for the Life Sciences." In: *Drug Discovery World (DDW)* (2012), pp. 75–80. URL: www.ddw-online.com/media/32/saved-by-the-bel---ringing-in-a-common-language-for-the-life-sciences.pdf.

[147] Richard Smith. "Problems With Peer Review and Alternatives." In: *British Medical Journal (BMJ)* 296.6624 (1988), pp. 774–777. URL: www.jstor.org/stable/29530055.

[148] Richard Smith. "Classical Peer Review: An Empty Gun." In: *Breast cancer research* 12 (2010). DOI: 10.1186/bcr2742.

[149] Richard Socher, Alex Perelygin, Jean Y. Wu, Jason Chuang, Christopher D. Manning, Andrew Y. Ng, and Christopher Potts. "Recursive Deep Models for Semantic Compositionality over a Sentiment Treebank." In: *Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing*. EMNP'13. 2013. URL: https://aclanthology.org/D13-1170/.

[150] David Soergel, Adam Saunders, and Andrew McCallum. "Open Scholarship and Peer Review: a Time for Experimentation." In: *Proceedings of the 30th International Conference on Machine Learning*. ICML'13. 2013. URL: https://openreview.net/pdf?id=xf0zSBd2iufMg.

[151] Bodo M. Stern and Erin K. O'Shea. "A proposal for the future of scientific publishing in the life sciences." In: *PLOS Biology* 17.2 (2019), pp. 1–10. DOI: 10.1371/journal.pbio.3000116.

[152] Gina S. Sucato and Cynthia Holland-Hall. "Reviewing Manuscripts: A Systematic Approach." In: *Journal of Pediatric and Adolescent Gynecology* 31.5 (2018), pp. 441–445. DOI: 10.1016/j.jpag.2018.06.005.

[153] Maite Taboada, Julian Brooke, Milan Tofiloski, Kimberly Voll, and Manfred Stede. "Lexicon-based methods for sentiment analysis." In: *Computational Linguistics* 37.2 (2011), pp. 267–307. DOI: 10.1162/COLI_a_00049.

[154] Jaana Takis, AQM Saiful Islam, Christoph Lange, and Sören Auer. "Crowdsourced Semantic Annotation of Scientific Publications and Tabular Data in PDF." In: *Proceedings of the 11th International Conference on Semantic Systems*. SEMANTICS'15. 2015, pp. 1–8. DOI: 10.1145/2814864.2814887.

[155] Antonio Tenorio-Fornés, Viktor Jacynycz, David Llop-Vila, Antonio A. Sánchez-Ruiz-Granados, and Samer Hassan. "Towards a Decentralized Process for Scientific Publication and Peer Review using Blockchain and IPFS." In: *Proceedings of the 52nd Hawaii International Conference on System Sciences*. HICSS'19. 2019. URL: http://hdl.handle.net/10125/59901.

[156] Ilaria Tiddi, Daniel Balliet, and Annette ten Teije. "Fostering Scientific Meta-analyses with Knowledge Graphs: A Case-Study." In: *Proceedings of the 17th Extended Semantic Web Conference*. ESWC'20. 2020, pp. 287–303. DOI: 10.1007/978-3-030-49461-2_17.

[157] Shahadat Uddin, Arif Khan, and Louise A. Baur. "A Framework to Explore the Knowledge Structure of Multidisciplinary Research Fields." In: *PLoS ONE* 10.4 (2015). DOI: 10.1371/journal.pone.0123537.

[158] Brian Uzzi, Satyam Mukherjee, Michael Stringer, and Ben Jones. "Atypical Combinations and Scientific Impact." In: *Science* 342.6157 (2013), pp. 468–472. DOI: 10.1126/science.1240474.

[159] Sahar Vahdati, Said Fathalla, Sören Auer, Christoph Lange, and Maria-Esther Vidal. "Semantic Representation of Scientific Publications." In: *23rd International Conference on Theory and Practice of Digital Libraries*. TPDL'19. 2019, pp. 375–379. DOI: 10.1007/978-3-030-30760-8_37.

[160] Lars Vogt. "Levels and Building Blocks—Toward a Domain Granularity Framework for the Life Sciences." In: *Journal of Biomedical Semantics* 10 (2019). DOI: 10.1186/s13326-019-0196-2.

[161] Anita de Waard and Jodi Schneider. "Formalising Uncertainty: An Ontology of Reasoning, Certainty and Attribution (ORCA)." In: *Proceedings of the Joint Workshop on Semantic Technologies Applied to Biomedical Informatics and Individualized Medicine*. Vol. 930.

SATBI+SWIM'12. 2012, pp. 10–17. URL: https://ceur-ws.org/Vol-930/p2.pdf.

[162] Peiling Wang, Manasa Rath, Michael Deike, and Wu Qiang. "Open Peer Review: An Innovation in Scientific Publishing." In: *Journal of Data and Information Science* 1.4 (2016). DOI: 10.9776/16315.

[163] Mark Ware and Michael Mabe. *The STM Report: An Overview of Scientific and Scholarly Journal Publishing*. Fourth Edition. International Association of Scientific, Technical and Medical Publishers. 2015. URL: https://digitalcommons.unl.edu/scholcom/9.

[164] David Westergaard, Hans Henrik Stærfeldt, Christian Tønsberg, Lars Juhl Jensen, and Søren Brunak. "A Comprehensive and Quantitative Comparison of Text-Mining in 15 Million Full-text Articles Versus Their Corresponding Abstracts." In: *PLoS Computational Biology* 14.2 (2018). DOI: 10.1371/journal.pcbi.1005962.

[165] Mark D. Wilkinson et al. "The FAIR Guiding Principles for Scientific Data Management and Stewardship." In: *Scientific Data* 3.1 (2016). DOI: 10.1038/sdata.2016.18.

[166] Yan Xu, Lili Mou, Ge Li, Yunchuan Chen, Hao Peng, and Zhi Jin. "Classifying Relations via Long Short Term Memory Networks along Shortest Dependency Paths." In: *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing*. EMNLP'15. 2015, pp. 1785–1794. DOI: 10.18653/v1/D15-1206.

[167] Vikas Yadav and Steven Bethard. "A Survey on Recent Advances in Named Entity Recognition from Deep Learning Models." In: *Proceedings of the 27th International Conference on Computational Linguistics*. COLING'18. 2018, pp. 2145–2158. URL: https://aclanthology.org/C18-1182.

[168] Weizhe Yuan, Pengfei Liu, and Graham Neubig. "Can We Automate Scientific Reviewing?" In: *Journal of Artificial Intelligence Research* 75 (2022). DOI: 10.1613/jair.1.12862.

[169] Amrapali Zaveri, Anisa Rula, Andrea Maurino, Ricardo Pietrobon, and Jens Lehmann Sören Auer. "Quality Assessment for Linked Data: A Survey." In: *Semantic Web* 7.1 (2015). DOI: 10.3233/SW-150175.

[170] Daojian Zeng, Kang Liu, Siwei Lai, Guangyou Zhou, and Jun Zhao. "Relation Classification via Convolutional Deep Neural Network." In: *Proceedings of the 25th International Conference on Computational Linguistics*. COLING'14. 2014, pp. 2335–2344. URL: https://aclanthology.org/C14-1220.

[171] Jeremy Zucker et al. "Leveraging Structured Biological Knowledge for Counterfactual Inference: A Case Study of Viral Pathogenesis." In: *IEEE Transactions on Big Data* 7.1 (2021), pp. 25–37. DOI: 10.1109/TBDATA.2021.3050680.

SUMMARY

Scientific publishing is the main way in which novel research ideas, evidence, data, and scientific results are communicated, shared and assessed. Currently, this process of scientific publishing lacks, most of the time, transparency and machine-interpretable representations. As such, scientific publication is still done in scientific articles, basically long, coarse-grained text with complicated structures in natural language that are optimized for human readers and not for automated systems. Moreover, peer reviewing continues to be the main method of quality assessment in science, despite the many and serious flaws emphasized by the scientific community. These include the lack of transparency, accuracy and efficiency. And, as science is rapidly changing and moving more towards a digital environment, with scientific contributions increasing in volume and complexity each day, scientific publishing that still follows this old paradigm of publishing seems to be at odds with the current scientific progress.

We address some of these issues that make it hard for researchers in the Digital Era to disseminate, share and assess their work in an automated fashion. We focus on making the publication process more organized, structured and transparent. Mainly, we investigate how approaches from the Semantic Web can be applied in the digital environment of scientific publishing in a way that semantics can be embedded with data from the start, with technologies such as RDF and OWL that use common data formats and exchange protocols to formally represent data. Moreover, we use newly developed approaches like nanopublications together with other formalisms based on formal logic to make scientific knowledge accessible from the start, to both human and machines alike. Thus, we explore how we can further use machines to publish, aggregate, summarize, and automate more of the scientific publishing process.

In particular, each chapter focuses on one of the main activities involved in scientific publishing, activities that are mainly performed by reviewers, editors, readers (both human and machines) and authors. Also, we consider the different aspects of scientific articles and reviews together with their findings, while the contributions of each chapter are theoretical or practical in nature with respect to the topic addressed by each chapter. Basically, we propose a fine-grained model for reviewing using semantic comments in Chapter 2, then we show that it is possible to further use and apply a general scheme for the representation and publication of interlinked publication elements in Chapter 3, while in Chapter 4 we focus on a formal semantic representation of the content of scientific claims. Finally, Chapter 5 combines all the models and

applications presented in the previous chapters in a field study with "formalization papers" which proves the overall technical and practical feasibility of our approach, where not only the publication of scientific contributions would be possible in a structured manner, but also all the publication processes in between.

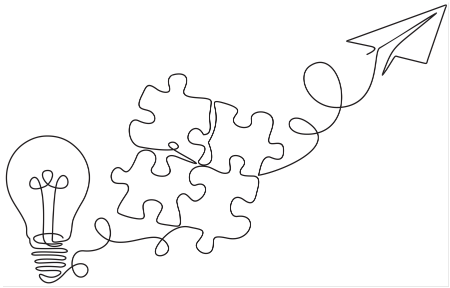
In Chapter 2 we applied some of the principles of the Web and the Semantic Web in the field of scientific publishing, concentrating on the reviewing process. Our goal was to make these reviewing processes more efficient and more accurate by using a finer-grained model of reviewing. For this, we created a novel model titled Linkflows that is able to support a fine-grained, more detailed and semantically rich view on the reviewing process. In order to evaluate this model, we created a manually curated dataset from several recent open peer review computer science journals and conferences and applied this model post-publication for the selected articles and reviews. We asked the actual peer reviewers to rate some of their review comments based on the dimensions of the Linkflows model (hence establishing a ground-truth). Then, the model experts rated the complete review snippets for the articles, where answers from the actual peer reviewers were received, to also check the inter-annotator agreement for the dimensions of the Linkflows reviewing model and compare them with the ground-truth. Next, we performed a user experiment where we compare experts with the actual peer reviewer (ground-truth) and model experts answers. Our results suggest not only that this new, fine-grained model of reviewing can be easily understood and applied by model experts and peers, but also that it can provide a way to capture in a formal and machine-readable manner, without discrepancies, the reviewers' intentions from the start, making the reviewing process more accurate and efficient.

In Chapter 3 we proposed a different way of publishing that again makes use of the principles of the Web and the Semantic Web in the publishing field. For this, we used a finer-grained way to represent and structure a scientific article and its corresponding reviews (using the Linkflows model developed in Chapter 2) and we made use of semantically enriched models and ontologies for representing the article and review structures. This approach allowed us to use nanopublications as a unifying model to represent semantically the elements of scientific publications, their assessments together with the involved processes and actors, creating an interlinked network of the elements of publications and their reviews. In our evaluation, we see that this representation helps us answer interesting competency questions from the different perspectives of the editors. Moreover, this representation enabled us to develop a prototype interface to support the editors in their work, and the results of our editor study indicate that the interface developed was useful in answering the editors' competency questions. Overall, the research done in this chapter shows that a unified and semantic publication model based on nanopublications can make scientific communication more effective and user-friendly.

In Chapter 4 we continued in our endeavour to make scientific knowledge accessible to automated systems as well, by focusing on whether a fully formal representation of the content of the core claims or scientific findings of research work is possible. We analyzed the high-level scientific claims from a sample of scientific articles across disciplines and despite noticing that their semantics are more complex than the simple application of formalisms like RDF and OWL, eliciting a clear semantic pattern is possible. We called this semantic pattern "the super-pattern" and showed that its instantiation leads to a strictly defined statement in high-order logic. In our evaluation, we applied this super-pattern to an enlarged sample of scientific claims and performed a study with knowledge representation experts. Our results showed that it is possible to express with a high degree of consistency the content of high-level scientific claims by writing formalizations of such claims that make use of the super-pattern ontology we developed. Moreover, the formalization study we conducted with knowledge experts proves that even if prerequisite knowledge is necessary to achieve this, this task is one that can be done in a reliable manner. This paves the way for machines to be more involved in the scientific endeavour and perform automated consistency checking, question answering, aggregation and much more.

In Chapter 5 we propose a first step in the direction of making scientific articles more machine-readable and interpretable by proposing a new way of publishing scientific results in which scientific knowledge can be expressed with formal semantics from the start. Here, we combine all the models and applications developed in previous chapters and we set out to perform a field study in which high-level scientific claims represented in formal logic are published in a special issue of an existing journal. By using the concept and technology of nanopublications to publish such formalization papers, a novel formalization of a previously published scientific claim, we were able to represent not only such high-level claims as publications, but also the entire publication process, including the reviews, responses and decisions. Our evaluation demonstrated the feasibility of our approach, showing high levels of interest and confidence from the participating authors. The results seem to indicate that it is possible to improve in a significant way the effectiveness and efficiency of the scientific publication, bringing us closer to a world where machines can help us interpret, organize and understand scientific knowledge better in a reliable and transparent manner.

Altogether, our research shows that it is possible to bring scientific publishing closer to automated systems and make the elements of publications and their assessments, along with their corresponding processes more machine-interpretable, especially if semantics are captured from the start. The complementary models proposed can be used in the publishing practice and can be adequately used in tools that support the publishing activities and workflows. And, regardless of the challenges encountered along these processes, after reasonable guidance, users are able to use these tools and models with ease despite their novelty, demonstrating that overall our ideas work. This can be a first step in the direction of involving machines more in the research community, where a new way of publishing that considers the Semantic Web principles can help in creating a society where computers can assist the Open Science research community become more efficient and make scientific contributions FAIR: easier to find, more accessible, interoperable and reusable.



SAMENVATTING

Wetenschappelijk publiceren is de belangrijkste manier waarop nieuwe onderzoeksideeën, bewijzen, gegevens en wetenschappelijke resultaten worden gecommuniceerd, gedeeld en beoordeeld. Momenteel ontbreekt het dit proces van wetenschappelijke publicatie meestal aan transparantie en machine-interpreteerbare weergaven. Wetenschappelijke publicaties worden nog steeds gedaan in de vorm van wetenschappelijke artikelen, in feite lange, grofkorrelige teksten met ingewikkelde structuren in natuurlijke taal die geoptimaliseerd zijn voor menselijke lezers en niet voor geautomatiseerde systemen. Bovendien blijft peer review de belangrijkste methode voor kwaliteitsbeoordeling in de wetenschap, ondanks de vele en ernstige gebreken die door de wetenschappelijke gemeenschap worden benadrukt. Deze omvatten het gebrek aan transparantie, nauwkeurigheid en efficiëntie. En aangezien de wetenschap snel verandert en zich meer en meer in de richting van een digitale omgeving beweegt, met wetenschappelijke bijdragen die elke dag in volume en complexiteit toenemen, lijkt wetenschappelijk publiceren dat nog steeds dit oude publicatieparadigma volgt, haaks te staan op de huidige wetenschappelijke vooruitgang.

Wij pakken enkele van deze problemen aan die het voor onderzoekers in het digitale tijdperk moeilijk maken om hun werk op een geautomatiseerde manier te verspreiden, te delen en te beoordelen. Wij richten ons op het organiseren, structureren en transparanter maken van het publicatieproces. We onderzoeken vooral hoe benaderingen van het Semantisch Web kunnen worden toegepast in de digitale omgeving van wetenschappelijke publicaties, zodat semantiek vanaf het begin kan worden ingebed in gegevens, met technologieën zoals RDF en OWL die gemeenschappelijke gegevensformaten en uitwisselingsprotocollen gebruiken om gegevens formeel te vertegenwoordigen. Bovendien gebruiken we nieuw ontwikkelde benaderingen zoals nanopublicaties samen met andere formalismen gebaseerd op formele logica om wetenschappelijke kennis vanaf het begin toegankelijk te maken, zowel voor mensen als voor machines. Zo onderzoeken we hoe we machines verder kunnen gebruiken om meer van het wetenschappelijke publicatieproces te publiceren, te aggregeren, samen te vatten en te automatiseren.

In het bijzonder richt elk hoofdstuk zich op een van de belangrijkste activiteiten bij het publiceren van wetenschappelijke artikelen, activiteiten die voornamelijk worden uitgevoerd door recensenten, redacteuren, lezers (zowel mensen als machines) en auteurs. Ook bekijken we de verschillende aspecten van wetenschappelijke artikelen en recensies samen met hun bevindingen, terwijl de bijdragen van elk hoofdstuk theoretisch of praktisch van aard zijn met betrekking tot het onderwerp dat in elk hoofdstuk wordt behandeld. In principe stellen we in hoofdstuk 2 een fijnkorrelig model voor om te reviewen met behulp van semantisch commentaar, vervolgens laten we zien dat het mogelijk is om een algemeen schema voor de representatie en publicatie van onderling verbonden publicatie-elementen te gebruiken en toe te passen in hoofdstuk 3, terwijl we ons in hoofdstuk 4 richten op een formele semantische representatie van de inhoud van wetenschappelijke claims. Tenslotte combineert hoofdstuk 5 alle modellen en toepassingen die in de voorgaande hoofdstukken zijn gepresenteerd in een veldstudie met "formalisatiepapieren" die de algemene technische en praktische haalbaarheid van onze aanpak bewijst, waarbij niet alleen de publicatie van wetenschappelijke bijdragen op een gestructureerde manier mogelijk zou zijn, maar ook alle publicatieprocessen daartussenin.

In hoofdstuk 2 pasten we enkele principes van het Web en het Semantische Web toe op het gebied van wetenschappelijke publicaties, waarbij we ons concentreerden op het beoordelingsproces. Ons doel was om deze reviewprocessen efficiënter en nauwkeuriger te maken door een fijnmaziger reviewmodel te gebruiken. Hiervoor hebben we een nieuw model gecreëerd, Linkflows genaamd, dat een fijnkorrelig, meer gedetailleerd en semantisch rijk beeld van het reviewproces kan ondersteunen. Om dit model te evalueren, creëerden we een handmatig gecureerde dataset van verschillende recente open peer review tijdschriften en conferenties over computerwetenschappen en pasten we dit model na publicatie toe op de geselecteerde artikelen en recensies. We vroegen de daadwerkelijke peer reviewers om een aantal van hun reviewcommentaren te beoordelen op basis van de dimensies van het Linkflows-model (om zo een ground-truth vast te stellen). Vervolgens beoordeelden de modelexperts de volledige recensiefragmenten voor de artikelen, waarop antwoorden van de echte peer reviewers werden ontvangen, om ook de inter-annotatorovereenkomst voor de dimensies van het Linkflows-recensiemodel te controleren en ze te vergelijken met de ground-truth. Vervolgens hebben we een gebruikersexperiment uitgevoerd waarbij we de antwoorden van experts vergelijken met de werkelijke peer reviewer (ground-truth) en de antwoorden van modelexperts. Onze resultaten suggereren niet alleen dat dit nieuwe, fijnkorrelige model van reviewen gemakkelijk kan worden begrepen en toegepast door modelexperts en peers, maar ook dat het een manier kan zijn om de bedoelingen van de reviewers vanaf het begin vast te leggen op een formele en machineleesbare manier, zonder discrepanties, waardoor het reviewproces nauwkeuriger en efficiënter wordt.

In hoofdstuk 3 stelden we een andere manier van publiceren voor die opnieuw gebruik maakt van de principes van het Web en het Semantische Web op het gebied van publiceren. Hiervoor gebruikten we een fijnkorreligere manier om een wetenschappelijk artikel en de bijbehorende recensies weer te geven en te structureren (met behulp van het Linkflows-model dat ontwikkeld is in hoofdstuk 2) en we maakten gebruik van semantisch verrijkte modellen en ontologieën om de artikel- en recensiestructuren weer te geven. Met deze aanpak konden we nanopublicaties gebruiken als een verenigend model om de elementen van wetenschappelijke publicaties, hun beoordelingen en de betrokken processen en actoren semantisch weer te geven, waardoor een onderling verbonden netwerk van de elementen van publicaties en hun beoordelingen ontstond. In onze evaluatie zien we dat deze representatie ons helpt bij het beantwoorden van interessante competentievragen vanuit de verschillende perspectieven van de redacteuren. Bovendien konden we dankzij deze weergave een prototype van de interface ontwikkelen om de redacteuren in hun werk te ondersteunen, en de resultaten van onze studie onder redacteuren geven aan dat de ontwikkelde interface nuttig was om de competentievragen van de redacteuren te beantwoorden. Over het algemeen toont het onderzoek in dit hoofdstuk aan dat een uniform en semantisch publicatiemodel op basis van nanopublicaties wetenschappelijke communicatie effectiever en gebruiksvriendelijker kan maken.

In hoofdstuk 4 gingen we verder met ons streven om wetenschappelijke kennis ook toegankelijk te maken voor geautomatiseerde systemen, door ons te richten op de vraag of een volledig formele representatie van de inhoud van de kernclaims of wetenschappelijke bevindingen van onderzoekswerk mogelijk is. We analyseerden de wetenschappelijke beweringen op hoog niveau uit een steekproef van wetenschappelijke artikelen uit verschillende disciplines en ondanks het feit dat hun semantiek complexer is dan de eenvoudige toepassing van formalismen zoals RDF en OWL, is het mogelijk om een duidelijk semantisch patroon te ontlokken. Wij noemden dit semantische patroon "het superpatroon" en toonden aan dat de instantiatie ervan leidt tot een strikt gedefinieerde uitspraak in hoge-orde logica. In onze evaluatie hebben we dit superpatroon toegepast op een uitgebreide steekproef van wetenschappelijke beweringen en hebben we een onderzoek uitgevoerd met experts op het gebied van kennisrepresentatie. Onze resultaten toonden aan dat het mogelijk is om de inhoud van wetenschappelijke beweringen op hoog niveau met een hoge mate van consistentie uit te drukken door formalisaties van dergelijke beweringen te schrijven die gebruikmaken van de ontologie van het superpatroon die we hebben ontwikkeld. Bovendien bewijst het formalisatieonderzoek dat we met kennisexperts hebben uitgevoerd dat, zelfs als hiervoor vereiste kennis nodig is, deze taak op een betrouwbare manier kan worden uitgevoerd. Dit maakt de weg vrij voor machines om meer betrokken te raken bij wetenschappelijke inspanningen en geautomatiseerde consistentiecontroles, vraagbeantwoording, aggregatie en nog veel meer uit te voeren.

In 5 stellen we een eerste stap voor in de richting van het machineleesbaar en interpreteerbaar maken van wetenschappelijke artikelen door een nieuwe manier voor te stellen voor het publiceren van wetenschappelijke resultaten waarbij wetenschappelijke kennis vanaf het begin kan worden uitgedrukt met formele semantiek. In dit hoofdstuk combineren we alle modellen en toepassingen die in eerdere hoofdstukken zijn ontwikkeld en voeren we een veldstudie uit waarin wetenschappelijke beweringen op hoog niveau, weergegeven in formele logica, worden gepubliceerd in een speciale uitgave van een bestaand tijdschrift. Door gebruik te maken van het concept en de technologie van nanopublicaties voor het publiceren van dergelijke formalisatiepublicaties, een nieuwe formalisatie van een eerder gepubliceerde wetenschappelijke bewering, konden we niet alleen dergelijke claims op hoog niveau als publicaties weergeven, maar ook het hele publicatieproces, inclusief de beoordelingen, reacties en beslissingen. Onze evaluatie toonde de haalbaarheid van onze aanpak aan, met een hoge mate van interesse en vertrouwen van de deelnemende auteurs. De resultaten lijken erop te wijzen dat het mogelijk is om de effectiviteit en efficiëntie van wetenschappelijke publicaties aanzienlijk te verbeteren, waardoor we dichter bij een wereld komen waarin machines ons kunnen helpen om wetenschappelijke kennis op een betrouwbare en transparante manier te interpreteren, organiseren en begrijpen.

Al met al laat ons onderzoek zien dat het mogelijk is om wetenschappelijk publiceren dichter bij geautomatiseerde systemen te brengen en de elementen van publicaties en hun beoordelingen, samen met de bijbehorende processen, beter machinaal interpreteerbaar te maken, vooral als de semantiek vanaf het begin wordt vastgelegd. De voorgestelde aanvullende modellen kunnen in de uitgeefpraktijk worden gebruikt en kunnen adequaat worden toegepast in tools die de publicatieactiviteiten en workflows ondersteunen. En, ongeacht de uitdagingen die men tijdens deze processen tegenkomt, na redelijke begeleiding zijn gebruikers in staat om deze tools en modellen met gemak te gebruiken, ondanks hun nieuwheid, wat aantoont dat onze ideeën over het algemeen werken. Dit kan een eerste stap zijn in de richting van het meer betrekken van machines bij de onderzoeksgemeenschap, waar een nieuwe manier van publiceren die rekening houdt met de principes van het Semantisch Web kan helpen bij het creëren van een maatschappij waarin computers de Open Science onderzoeksgemeenschap kunnen helpen efficiënter te worden en wetenschappelijke bijdragen FAIR te maken: gemakkelijker te vinden, toegankelijker, interoperabel en herbruikbaar.

OPMERKING: De Nederlandse vertaling van deze samenvatting is automatisch uit het Engels vertaald door DeepL Translate. Voor klachten, suggesties of opmerkingen over deze vertaling kunt u contact opnemen met de vertaler AI DeepL, die zichzelf "The world's most accurate translator" noemt. We hopen dat het deze test heeft doorstaan, in ieder geval met deze Nederlandse vertaling ☺.

REZUMAT

Publicarea științifică este principala modalitate prin care sunt comunicate, împărtășite și evaluate ideile noi de cercetare, dovezile, datele și rezultatele științifice. În prezent, acestui proces de publicare științifică îi lipsesc, de cele mai multe ori, transparența și reprezentările interpretabile de către mașini. Ca atare, publicarea științifică se face încă în articole științifice, în principiu texte lungi, cu text grosier, cu structuri complicate în limbaj natural, optimizate pentru cititorii umani și nu pentru sistemele automate. În plus, evaluarea inter pares continuă să fie principala metodă de evaluare a calității în știință, în ciuda numeroaselor și gravele defecte subliniate de comunitatea științifică. Printre acestea se numără lipsa de transparență, acuratețe și eficiență. Și, în condițiile în care știința se schimbă rapid și se îndreaptă tot mai mult către un mediu digital, iar contribuțiile științifice cresc în volum și complexitate în fiecare zi, publicarea științifică care încă urmează această veche paradigmă de publicare pare să fie în contradicție cu progresul științific actual.

Abordăm unele dintre aceste probleme care îngreunează diseminarea, partajarea și evaluarea automată a activității cercetătorilor din era digitală. Ne concentrăm pe transformarea procesului de publicare în unul mai organizat, mai structurat și mai transparent. În principal, investigăm modul în care abordările din Web-ul semantic pot fi aplicate în mediul digital al publicării științifice, astfel încât semantica să poată fi încorporată de la început în date, cu ajutorul unor tehnologii precum RDF și OWL, care utilizează formate de date și protocoale de schimb comune pentru a reprezenta în mod normal datele. Mai mult, folosim abordări nou dezvoltate, cum ar fi nanopublicațiile, împreună cu alte formalisme bazate pe logica formală, pentru a face cunoștințele științifice accesibile de la început, atât pentru oameni, cât și pentru mașini. Astfel, explorăm modul în care putem utiliza în continuare mașinile pentru a publica, agrega, rezuma și automatiza o mai mare parte din procesul de publicare științifică.

În special, fiecare capitol se concentrează asupra uneia dintre principalele activități implicate în publicarea științifică, activități care sunt efectuate în principal de recenzori, editori, cititori (atât umani, cât și mașini) și autori. De asemenea, luăm în considerare diferitele aspecte ale articolelor științifice și ale recenziilor, împreună cu rezultatele acestora, în timp ce contribuțiile fiecărui capitol sunt de natură teoretică sau practică în ceea ce privește subiectul abordat în fiecare capitol. Practic, în capitolul 2 propunem un model cu granulație fină pentru recenzii folosind comentarii semantice, apoi arătăm că este posibilă utilizarea și aplicarea ulterioară a unei scheme generale pentru reprezentarea și publicarea elementelor de publicare interconectate în capitolul 3, în timp ce în capitolul 4 ne concentrăm pe o reprezentare semantică formală a conținutului afirmațiilor științifice. În cele din urmă, capitolul 5 combină toate modelele și aplicațiile prezentate în capitolele anterioare într-un studiu de teren cu "lucrări de formalizare" care demonstrează fezabilitatea tehnică și practică generală a abordării noastre, în care nu numai publicarea contribuțiilor științifice ar fi posibilă într-o manieră structurată, ci și toate procesele de publicare intermediare.

În capitolul 2 am aplicat unele dintre principiile web și ale web-ului semantic în domeniul publicării științifice, concentrându-ne asupra procesului de revizuire. Scopul nostru a fost de a face aceste procese de revizuire mai eficiente și mai precise prin utilizarea unui model de revizuire cu o mai mare finețe. În acest scop, am creat un model nou intitulat Linkflows, care este capabil să susțină o viziune fină, mai detaliată și mai bogată din punct de vedere semantic asupra procesului de revizuire. Pentru a evalua acest model, am creat un set de date curatoriat manual din mai multe reviste și conferințe recente de informatică cu evaluare inter pares deschisă și am aplicat acest model după publicare pentru articolele și recenziile selectate. Le-am cerut recenzorilor reali să evalueze unele dintre comentariile lor pe baza dimensiunilor modelului Linkflows (stabilind astfel un adevăr de bază). Apoi, experții modelului au evaluat fragmente complete de recenzii pentru articolele la care au fost primite răspunsuri de la recenzenții reali, pentru a verifica, de asemenea, acordul dintre recenzenți în ceea ce privește dimensiunile modelului de recenzare Linkflows și pentru a le compara cu adevărul de bază. În continuare, am efectuat un experiment cu utilizatorii, în care am comparat răspunsurile experților cu cele ale recenzorilor reali (ground-truth) și ale experților modelului. Rezultatele noastre sugerează nu numai că acest nou model de revizuire cu granulație fină poate fi ușor de înțeles și aplicat de către experții model și de către colegi, dar și că poate oferi o modalitate de a capta într-un mod formal și lizibil de către mașină, fără discrepanțe, intențiile recenzorilor de la început, făcând procesul de revizuire mai precis și mai eficient.

În capitolul 3 am propus un mod diferit de publicare care utilizează din nou principiile Web și ale Web-ului semantic în domeniul publicării. În acest scop, am utilizat o modalitate mai fină de reprezentare și structurare a unui articol științific și a recenziilor corespunzătoare (utilizând modelul Linkflows dezvoltat în capitolul 2) și am folosit modele și ontologii îmbogățite semantic pentru a reprezenta structurile articolelor și recenziilor. Această abordare ne-a permis să folosim nanopublicațiile ca model unificator pentru a reprezenta din punct de vedere semantic elementele publicațiilor științifice, evaluările acestora împreună cu procesele și actorii implicați, creând o rețea interconectată a elementelor publicațiilor și a recenziilor acestora. În evaluarea noastră, observăm că această reprezentare ne ajută să răspundem la întrebări interesante privind competențele din diferitele perspective ale editorilor. În plus, această reprezentare ne-a permis să dezvoltăm un prototip de interfață pentru a sprijini editorii în activitatea lor, iar rezultatele studiului nostru asupra editorilor indică faptul că interfața dezvoltată a fost utilă pentru a răspunde la întrebările editorilor privind competențele. În general, cercetările efectuate în acest capitol arată că un model de publicare unificat și semantic bazat pe nanopublicații poate face comunicarea științifică mai eficientă și mai ușor de utilizat.

În capitolul 4 am continuat în efortul nostru de a face cunoștințele științifice accesibile și pentru sistemele automate, concentrându-ne asupra posibilității unei reprezentări complet formale a conținutului afirmațiilor de bază sau a rezultatelor științifice ale lucrărilor de cercetare. Am analizat afirmațiile științifice de nivel înalt dintr-un eșantion de articole științifice din mai multe discipline și, în ciuda faptului că am observat că semantica lor este mai complexă decât simpla aplicare a formalismelor precum RDF și OWL, este posibilă obținerea unui model semantic clar. Am numit acest model semantic "super-model" și am arătat că instanțierea sa conduce la o afirmație strict definită în logica de ordin înalt. În cadrul evaluării noastre, am aplicat acest super-model unui eșantion extins de afirmații științifice și am realizat un studiu cu experți în reprezentarea cunoștințelor. Rezultatele noastre au arătat că este posibilă exprimarea cu un grad ridicat de coerență a conținutului afirmațiilor științifice de nivel înalt prin scrierea formalizărilor unor astfel de afirmații care utilizează ontologia super-modelului pe care am dezvoltat-o. Mai mult, studiul de formalizare pe care l-am realizat cu experți în cunoaștere dovedește că, chiar dacă sunt necesare cunoștințe prealabile pentru a realiza acest lucru, această sarcină este una care poate fi realizată într-un mod fiabil. Acest lucru deschide calea pentru ca mașinile să se implice mai mult în activitatea științifică și să realizeze verificarea automată a coerenței, răspunsul la întrebări, agregarea și multe altele.

În capitolul 5 propunem un prim pas în direcția de a face articolele științifice mai ușor de citit și de interpretat de către mașini, propunând un nou mod de publicare a rezultatelor științifice, în care cunoștințele științifice pot fi exprimate de la început cu ajutorul unei semantici formale. Aici, combinăm toate modelele și aplicațiile dezvoltate în capitolele anterioare și ne propunem să realizăm un studiu de teren în care afirmațiile științifice de nivel înalt reprezentate în logica formală sunt publicate într-un număr special al unei reviste existente. Prin utilizarea conceptului și tehnologiei de nanopublicații pentru a publica astfel de lucrări de formalizare, o formalizare nouă a unei afirmații științifice publicate anterior, am reușit să reprezentăm nu numai astfel de afirmații de nivel înalt ca publicații, ci și întregul proces de publicare, inclusiv revizuirile, răspunsurile și deciziile. Evaluarea noastră a demonstrat fezabilitatea abordării noastre, arătând niveluri ridicate de interes și încredere din partea autorilor participanți. Rezultatele par să indice faptul că este posibil să se îmbunătățească în mod semnificativ eficacitatea și eficiența publicării științifice, apropiindu-ne de o lume în care mașinile ne pot ajuta să interpretăm, să organizăm și să înțelegem mai bine cunoștințele științifice într-un mod fiabil și transparent.

În ansamblu, cercetarea noastră arată că este posibil să aducem publicarea științifică mai aproape de sistemele automate și să facem ca elementele publicațiilor și evaluările acestora, împreună cu procesele corespunzătoare, să fie mai ușor de interpretat de către mașini, mai ales dacă semantica este captată de la început. Modelele complementare propuse pot fi folosite în practica editorială și pot fi utilizate în mod adecvat în instrumentele care sprijină activitățile și fluxurile de lucru din domeniul publicării. Și, indiferent de provocările întâlnite de-a lungul acestor procese, după o îndrumare rezonabilă, utilizatorii sunt capabili să utilizeze aceste instrumente și modele cu ușurință, în ciuda noutății lor, demonstrând că, în general, ideile noastre funcționează. Acesta poate fi un prim pas în direcția unei mai mari implicări a mașinilor în comunitatea de cercetare, unde un nou mod de publicare care ia în considerare principiile Web-ului semantic poate contribui la crearea unei societăți în care computerele pot ajuta comunitatea de cercetare Open Science să devină mai eficientă și să facă contribuțiile științifice FAIR: mai ușor de găsit, mai accesibile, interoperabile și reutilizabile.

NOTĂ: Traducerea în română a acestui rezumat a fost efectuată în mod automat din limba engleză de către DeepL Translate. Pentru orice reclamații, sugestii sau comentarii legate de această traducere vă rugăm să luați legătura cu AI-ul traducător DeepL, care se auto-intitulează ca fiind "The world's most accurate translator", adică "Cel mai precis traducător din lume". Sperăm că a trecut acest test, cel puțin cu traducerea aceasta în română ☺.

SIKS DISSERTATION SERIES

2016

2016-1 Syed Saiden Abbas (RUN), Recognition of Shapes by Humans and Machines

2016-2 Michiel Christiaan Meulendijk (UU), Optimizing medication reviews through decision support: prescribing a better pill to swallow

2016-3 Maya Sappelli (RUN), Knowledge Work in Context: User Centered Knowledge Worker Support

2016-4 Laurens Rietveld (VU), Publishing and Consuming Linked Data

2016-5 Evgeny Sherkhonov (UVA), Expanded Acyclic Queries: Containment and an Application in Explaining Missing Answers

2016-6 Michel Wilson (TUD), Robust scheduling in an uncertain environment

2016-7 Jeroen de Man (VU), Measuring and modeling negative emotions for virtual training

2016-8 Matje van de Camp (TiU), A Link to the Past: Constructing Historical Social Networks from Unstructured Data

2016-9 Archana Nottamkandath (VU), Trusting Crowdsourced Information on Cultural Artefacts

2016-10 George Karafotias (VUA), Parameter Control for Evolutionary Algorithms

2016-11 Anne Schuth (UVA), Search Engines that Learn from Their Users

2016-12 Max Knobbout (UU), Logics for Modelling and Verifying Normative Multi-Agent Systems

2016-13 Nana Baah Gyan (VU), The Web, Speech Technologies and Rural Development in West Africa - An ICT4D Approach

2016-14 Ravi Khadka (UU), Revisiting Legacy Software System Modernization

2016-15 Steffen Michels (RUN), Hybrid Probabilistic Logics - Theoretical Aspects, Algorithms and Experiments

2016-16 Guangliang Li (UVA), Socially Intelligent Autonomous Agents that Learn from Human Reward

2016-17 Berend Weel (VU), Towards Embodied Evolution of Robot Organisms

2016-18 Albert Merono Penuela (VU), Refining Statistical Data on the Web

2016-19 Julia Efremova (TUE), Mining Social Structures from Genealogical Data

2016-20 Daan Odijk (UVA), Context & Semantics in News & Web Search

2016-21 Alejandro Moreno Colleri (UT), From Traditional to Interactive Playspaces: Automatic Analysis of Player Behavior in the Interactive Tag Playground

2016-22 Grace Lewis (VU), Software Architecture Strategies for Cyber-Foraging Systems

2016-23 Fei Cai (UVA), Query Auto Completion in Information Retrieval

2016-24 Brend Wanders (UT), Repurposing and Probabilistic Integration of Data; An Iterative and data model independent approach

2016-25 Julia Kiseleva (TUE), Using Contextual Information to Understand Searching and Browsing Behavior

2016-26 Dilhan Thilakarathne (VU), In or Out of Control: Exploring Computational Models to Study the Role of Human Awareness and Control in Behavioural Choices, with Applications in Aviation and Energy Management Domains

2016-27 Wen Li (TUD), Understanding Geo-spatial Information on Social Media

2016-28 Mingxin Zhang (TUD), Large-scale Agent-based Social Simulation - A study on epidemic prediction and control

2016-29 Nicolas Honing (TUD), Peak reduction in decentralised electricity systems -Markets and prices for flexible planning

2016-30 Ruud Mattheij (UvT), The Eyes Have It

2016-31 Mohammad Khelghati (UT), Deep web content monitoring

2016-32 Eelco Vriezekolk (UT), Assessing Telecommunication Service Availability Risks for Crisis Organisations

2016-33 Peter Bloem (UVA), Single Sample Statistics, exercises in learning from just one example

2016-34 Dennis Schunselaar (TUE), Configurable Process Trees: Elicitation, Analysis, and Enactment

2016-35 Zhaochun Ren (UVA), Monitoring Social Media: Summarization, Classification and Recommendation

2016-36 Daphne Karreman (UT), Beyond R2D2: The design of nonverbal interaction behavior optimized for robot-specific morphologies

2016-37 Giovanni Sileno (UvA), Aligning Law and Action - a conceptual and computational inquiry

2016-38 Andrea Minuto (UT), Materials that Matter - Smart Materials meet Art & Interaction Design

2016-39 Merijn Bruijnes (UT), Believable Suspect Agents; Response and Interpersonal Style Selection for an Artificial Suspect

2016-40 Christian Detweiler (TUD), Accounting for Values in Design

2016-41 Thomas King (TUD), Governing Governance: A Formal Framework for Analysing Institutional Design and Enactment Governance
2016-42 Spyros Martzoukos (UVA), Combinatorial and Compositional Aspects of Bilingual Aligned Corpora
2016-43 Saskia Koldijk (RUN), Context-Aware Support for Stress Self-Management: From Theory to Practice
2016-44 Thibault Sellam (UVA), Automatic Assistants for Database Exploration
2016-45 Bram van de Laar (UT), Experiencing Brain-Computer Interface Control
2016-46 Jorge Gallego Perez (UT), Robots to Make you Happy
2016-47 Christina Weber (UL), Real-time foresight - Preparedness for dynamic innovation networks
2016-48 Tanja Buttler (TUD), Collecting Lessons Learned
2016-49 Gleb Polevoy (TUD), Participation and Interaction in Projects. A Game-Theoretic Analysis
2016-50 Yan Wang (UVT), The Bridge of Dreams: Towards a Method for Operational Performance Alignment in IT-enabled Service Supply Chains

2017

2017-1 Jan-Jaap Oerlemans (UL), Investigating Cybercrime
2017-2 Sjoerd Timmer (UU), Designing and Understanding Forensic Bayesian Networks using Argumentation
2017-3 Daniel Harold Telgen (UU), Grid Manufacturing; A Cyber-Physical Approach with Autonomous Products and Reconfigurable Manufacturing Machines
2017-4 Mrunal Gawade (CWI), Multi-Core Parallelism in a Column-Store
2017-5 Mahdieh Shadi (UVA), Collaboration Behavior
2017-6 Damir Vandic (EUR), Intelligent Information Systems for Web Product Search
2017-7 Roel Bertens (UU), Insight in Information: from Abstract to Anomaly
2017-8 Rob Konijn (VU), Detecting Interesting Differences: Data Mining in Health Insurance Data using Outlier Detection and Subgroup Discovery
2017-9 Dong Nguyen (UT), Text as Social and Cultural Data: A Computational Perspective on Variation in Text
2017-10 Robby van Delden (UT), (Steering), Interactive Play Behavior
2017-11 Florian Kunneman (RUN), Modelling patterns of time and emotion in Twitter #anticipointment
2017-12 Sander Leemans (TUE), Robust Process Mining with Guarantees
2017-13 Gijs Huisman (UT), Social Touch Technology - Extending the reach of social touch through haptic technology
2017-14 Shoshannah Tekofsky (UvT), You Are Who You Play You Are: Modelling Player Traits from Video Game Behavior
2017-15 Peter Berck, Radboud University (RUN), Memory-Based Text Correction
2017-16 Aleksandr Chuklin (UVA), Understanding and Modeling Users of Modern Search Engines
2017-17 Daniel Dimov (UL), Crowdsourced Online Dispute Resolution
2017-18 Ridho Reinanda (UVA), Entity Associations for Search
2017-19 Jeroen Vuurens (TUD), Proximity of Terms, Texts and Semantic Vectors in Information Retrieval
2017-20 Mohammadbashir Sedighi (TUD), Fostering Engagement in Knowledge Sharing: The Role of Perceived Benefits, Costs and Visibility
2017-21 Jeroen Linssen (UT), Meta Matters in Interactive Storytelling and Serious Gaming (A Play on Worlds),
2017-22 Sara Magliacane (VU), Logics for causal inference under uncertainty
2017-23 David Graus (UVA), Entities of Interest - Discovery in Digital Traces
2017-24 Chang Wang (TUD), Use of Affordances for Efficient Robot Learning
2017-25 Veruska Zamborlini (VU), Knowledge Representation for Clinical Guidelines, with applications to Multimorbidity Analysis and Literature Search
2017-26 Merel Jung (UT), Socially intelligent robots that understand and respond to human touch
2017-27 Michiel Joosse (UT), Investigating Positioning and Gaze Behaviors of Social Robots: People's Preferences, Perceptions and Behaviors
2017-28 John Klein (VU), Architecture Practices for Complex Contexts
2017-29 Adel Alhuraibi (UVT), From IT-BusinessStrategic Alignment to Performance: A Moderated Mediation Model of Social Innovation, and Enterprise Governance of IT
2017-30 Wilma Latuny (UVT), The Power of Facial Expressions
2017-31 Ben Ruijl (UL), Advances in computational methods for QFT calculations
2017-32 Thaer Samar (RUN), Access to and Retrievability of Content in Web Archives
2017-33 Brigit van Loggem (OU), Towards a Design Rationale for Software Documentation: A Model of Computer-Mediated Activity
2017-34 Maren Scheffel (OUN), The Evaluation Framework for Learning Analytics
2017-35 Martine de Vos (VU), Interpreting natural science spreadsheets
2017-36 Yuanhao Guo (UL), Shape Analysis for Phenotype Characterisation from High-throughput Imaging

2017-37 Alejandro Montes Garcia (TUE), WiBAF: A Within Browser Adaptation Framework that Enables Control over Privacy
2017-38 Alex Kayal (TUD), Normative Social Applications
2017-39 Sara Ahmadi (RUN), Exploiting properties of the human auditory system and compressive sensing methods to increase noise robustness in ASR
2017-40 Altaf Hussain Abro (VUA), Steer your Mind: Computational Exploration of Human Control in Relation to Emotions, Desires and Social Support For applications in human-aware support systems"
2017-41 Adnan Manzoor (VUA), Minding a Healthy Lifestyle: An Exploration of Mental Processes and a Smart Environment to Provide Support for a Healthy Lifestyle
2017-42 Elena Sokolova (RUN), Causal discovery from mixed and missing data with applications on ADHD datasets
2017-43 Maaike de Boer (RUN), Maaike de Boer (RUN),
2017-44 Garm Lucassen (UU), Understanding User Stories - Computational Linguistics in Agile Requirements Engineering
2017-45 Bas Testerink (UU), Decentralized Runtime Norm Enforcement
2017-46 Jan Schneider (OU), Sensor-based Learning Support
2017-47 Yie Yang (TUD), Crowd Knowledge Creation Acceleration
2017-48 Angel Suarez (OU), Collaborative inquiry-based learning

2018

2018-1 Han van der Aa (VUA), Comparing and Aligning Process Representations
2018-2 Felix Mannhardt (TUE), Multi-perspective Process Mining
2018-3 Steven Bosems (UT), Causal Models For Well-Being: Knowledge Modeling, Model-Driven Development of Context-Aware Applications, and Behavior Prediction
2018-4 Jordan Janeiro (TUD), Flexible Coordination Support for Diagnosis Teams in Data-Centric Engineering Tasks
2018-5 Hugo Huurdeman (UVA), Supporting the Complex Dynamics of the Information Seeking Process
2018-6 Dan Ionita (UT), Model-Driven Information Security Risk Assessment of Socio-Technical Systems
2018-7 Jieting Luo (UU), A formal account of opportunism in multi-agent systems
2018-8 Rick Smetsers (RUN), Advances in Model Learning for Software Systems
2018-9 Xu Xie (TUD), Data Assimilation in Discrete Event Simulations
2018-10 Julienka Mollee (VUA), Moving forward: supporting physical activity behavior change through intelligent technology
2018-11 Mahdi Sargolzaei (UVA), Enabling Framework for Service-oriented Collaborative Networks
2018-12 Xixi Lu (TUE), Using behavioral context in process mining
2018-13 Seyed Amin Tabatabaei Using behavioral context in process mining: Exploring the added value of computational models for increasing the use of renewable energy in the residential sector
2018-14 Bart Joosten Detecting Social Signals with Spatiotemporal Gabor Filters
2018-15 Naser Davarzani (UM), Biomarker discovery in heart failure
2018-16 Jaebok Kim (UT), Automatic recognition of engagement and emotion in a group of children
2018-17 Jianpeng Zhang (TUE), On Graph Sample Clustering
2018-18 Henriette Nakad (UL), De Notaris en Private Rechtspraak
2018-19 Minh Duc Pham (VUA), Emergent relational schemas for RDF
2018-20 Manxia Liu (RUN), Time and Bayesian Networks
2018-21 Aad Slootmaker (OU), EMERGO: a generic platform for authoring and playing scenario-based serious games
2018-22 Eric Fernandes de Mello Araújo (VUA), Contagious: Modeling the Spread of Behaviours, Perceptions and Emotions in Social Networks
2018-23 Kim Schouten (EUR), Semantics-driven Aspect-Based Sentiment Analysis
2018-24 Jered Vroon (UT), Responsive Social Positioning Behaviour for Semi-Autonomous Telepresence Robots
2018-25 Riste Gligorov (VUA), Serious Games in Audio-Visual Collections
2018-26 Roelof de Vries (UT), Theory-Based And Tailor-Made: Motivational Messages for Behavior Change Technology
2018-27 Maikel Leemans (TUE), Hierarchical Process Mining for Scalable Software Analysis
2018-28 Christian Willemse (UT), Social Touch Technologies: How they feel and how they make you feel
2018-29 Yu Gu (UVT), Emotion Recognition from Mandarin Speech
2018-30 Wouter Beek (VU), The K in semantic web stands for knowledge: scaling semantics to the web

2019

2019-1 Rob van Eijk (UL), Web privacy measurement in real-time bidding systems. A graph-based approach to RTB system classification
2019-2 Emmanuelle Beauxis- Aussalet (CWI, UU), Statistics and Visualizations for Assessing Class Size Uncertainty
2019-3 Eduardo Gonzalez Lopez de Murillas (TUE), Process Mining on Databases: Extracting Event Data from Real Life Data Sources

2019-4 Ridho Rahmadi (RUN), Finding stable causal structures from clinical data
2019-5 Sebastiaan van Zelst (TUE), Process Mining with Streaming Data
2019-6 Chris Dijkshoorn (VU), Nichesourcing for Improving Access to Linked Cultural Heritage Datasets
2019-7 Soude Fazeli (TUD), Recommender Systems in Social Learning Platforms
2019-8 Frits de Nijs (TUD), Resource-constrained Multi-agent Markov Decision Processes
2019-9 Fahimeh Alizadeh Moghaddam (UVA), Self-adaptation for energy efficiency in software systems
2019-10 Qing Chuan Ye (EUR), Multi-objective Optimization Methods for Allocation and Prediction
2019-11 Yue Zhao (TUD), Learning Analytics Technology to Understand Learner Behavioral Engagement in MOOCs
2019-12 Jacqueline Heinerman (VU), Better Together
2019-13 Guanliang Chen (TUD), MOOC Analytics: Learner Modeling and Content Generation
2019-14 Daniel Davis (TUD), Large-Scale Learning Analytics: Modeling Learner Behavior & Improving Learning Outcomes in Massive Open Online Courses
2019-15 Erwin Walraven (TUD), Planning under Uncertainty in Constrained and Partially Observable Environments
2019-16 Guangming Li (TUE), Process Mining based on Object-Centric Behavioral Constraint (OCBC), Models
2019-17 Ali Hurriyetoglu (RUN), Extracting actionable information from microtexts
2019-18 Gerard Wagenaar (UU), Artefacts in Agile Team Communication
2019-19 Vincent Koeman (TUD), Tools for Developing Cognitive Agents
2019-20 Chide Groenouwe (UU), Fostering technically augmented human collective intelligence
2019-21 Cong Liu (TUE), Software Data Analytics: Architectural Model Discovery and Design Pattern Detection
2019-22 Martin van den Berg (VU), Improving IT Decisions with Enterprise Architecture
2019-23 Qin Lin (TUD), Intelligent Control Systems: Learning, Interpreting, Verification
2019-24 Anca Dumitrache (VU), Truth in Disagreement-Crowdsourcing Labeled Data for Natural Language Processing
2019-25 Emiel van Miltenburg (VU), Pragmatic factors in (automatic), image description
2019-26 Prince Singh (UT), An Integration Platform for Synchromodal Transport
2019-27 Alessandra Antonaci (OUN), The Gamification Design Process applied to (Massive), Open Online Courses
2019-28 Esther Kuindersma (UL), Cleared for take-off:Game-based learning to prepare airline pilots for critical situations
2019-29 Daniel Formolo (VU), Using virtual agents for simulation and training of social skills in safety-critical circumstances
2019-30 Vahid Yazdanpanah (UT), Multiagent Industrial Symbiosis Systems
2019-31 Milan Jelisavcic (VU), Alive and Kicking: Baby Steps in Robotics
2019-32 Chiara Sironi (UM), Monte-Carlo Tree Search for Artificial General Intelligence in Games
2019-33 Anil Yaman (TUE), Evolution of Biologically Inspired Learning in Artificial Neural Networks
2019-34 Negar Ahmadi (TUE), EEG Microstate and Functional Brain Network Features for Classification of Epilepsy and PNES
2019-35 Lisa Facey-Shaw (OUN), Gamification with digital badges in learning programming
2019-36 Kevin Ackermans (OUN), Designing Video-Enhanced Rubrics to Master Complex Skills
2019-37 Jian Fang (TUD), Database Acceleration on FPGAs
2019-38 Akos Kadar (OUN), Learning visually grounded and multilingual representations

2020

2020-1 Armon Toubman (UL), Calculated Moves: Generating Air Combat Behaviour
2020-2 Marcos de Paula Bueno (UL), Unraveling Temporal Processes using Probabilistic Graphical Models
2020-3 Mostafa Deghani (UvA), Learning with Imperfect Supervision for Language Understanding
2020-4 Maarten van Gompel (RUN), Context as Linguistic Bridges
2020-5 Yulong Pei (TUE), On local and global structure mining
2020-6 Preethu Rose Anish (UT), Stimulation Architectural Thinking during Requirements Elicitation - An Approach and Tool Support
2020-7 Wim van der Vegt (OUN), Towards a software architecture for reusable game components
2020-8 Ali Mirsoleimani (UL), Structured Parallel Programming for Monte Carlo Tree Search
2020-9 Myriam Traub (UU), Measuring Tool Bias & Improving Data Quality for Digital Humanities Research
2020-10 Alifah Syamsiyah (TUE), In-database Preprocessing for Process Mining
2020-11 Sepideh Mesbah (TUD), Semantic-Enhanced Training Data AugmentationMethods for Long-Tail Entity Recognition Models
2020-12 Ward van Breda (VU), Predictive Modeling in E-Mental Health: Exploring Applicability in Personalised Depression Treatment

2020-13 Marco Virgolin (CWI/TUD), Design and Application of Gene-pool Optimal Mixing Evolutionary Algorithms for Genetic Programming
2020-14 Mark Raasveldt (CWI/UL), Integrating Analytics with Relational Databases
2020-15 Georgiadis Konstantinos (OU), Smart CAT: Machine Learning for Configurable Assessments in Serious Games
2020-16 Ilona Wilmont (RUN), Cognitive Aspects of Conceptual Modelling
2020-17 Daniele Di Mitri (OU), The Multimodal Tutor: Adaptive Feedback from Multimodal Experiences
2020-18 Georgios Methenitis (TUD), Agent Interactions & Mechanisms in Markets with Uncertainties: Electricity Markets in Renewable Energy Systems
2020-19 Guido van Capelleveen (UT), Industrial Symbiosis Recommender Systems
2020-20 Albert Hankel (VU), Embedding Green ICT Maturity in Organisations
2020-21 Karine da Silva Miras de Araujo (VU), Where is the robot?: Life as it could be
2020-22 Maryam Masoud Khamis (RUN), Understanding complex systems implementation through a modeling approach: the case of e-government in Zanzibar
2020-23 Rianne Conijn (UT), The Keys to Writing: A writing analytics approach to studying writing processes using keystroke logging
2020-24 Lenin da Nóbrega Medeiros (VUA/RUN), How are you feeling, human? Towards emotionally supportive chatbots
2020-25 Xin Du (TUE), The Uncertainty in Exceptional Model Mining
2020-26 Krzysztof Leszek Sadowski (UU), GAMBIT: Genetic Algorithm for Model-Based mixed-Integer opTimization
2020-27 Ekaterina Muravyeva (TUD), Personal data and informed consent in an educational context
2020-28 Bibeg Limbu (TUD), Multimodal interaction for deliberate practice: Training complex skills with augmented reality
2020-29 Ioan Gabriel Bucur (RUN), Being Bayesian about Causal Inference
2020-30 Bob Zadok Blok (UL), Creatief, Creatiever, Creatiefst
2020-31 Gongjin Lan (VU), Learning Better: From Baby to Better
2020-32 Jason Rhuggenaath (TUE), Revenue management in online markets: pricing and online advertising
2020-33 Rick Gilsing (TUE), Supporting service-dominant business model evaluation in the context of business model innovation
2020-34 Anna Bon (MU), Intervention or Collaboration? Redesigning Information and Communication Technologies for Development
2020-35 Siamak Farshidi (UU), Multi-Criteria Decision-Making in Software Production

2021

2021-1 Francisco Xavier Dos Santos Fonseca (TUD), Location-based Games for Social Interaction in Public Space
2021-2 Rijk Mercuur (TUD), Simulating Human Routines:Integrating Social Practice Theory in Agent-Based Models
2021-3 Seyyed Hadi Hashemi (UVA), Modeling Users Interacting with Smart Devices
2021-4 Ioana Jivet (OU),), The Dashboard That Loved Me: Designing adaptive learning analytics for self-regulated learning
2021-5 Davide Dell'Anna (UU), Data-Driven Supervision of Autonomous Systems
2021-6 Daniel Davison (UT), "Hey robot, what do you think?" How children learn with a social robot
2021-7 Armel Lefebvre (UU), Research data management for open science
2021-8 Nardie Fanchamps (OU), The Influence of Sense-Reason-Act Programming on Computational Thinking
2021-9 Cristina Zaga (UT), The Design of Robothings. Non-Anthropomorphic and Non-Verbal Robots to Promote Children's Collaboration Through Play
2021-10 Quinten Meertens (UvA), Misclassification Bias in Statistical Learning
2021-11 Anne van Rossum (UL), Nonparametric Bayesian Methods in Robotic Vision
2021-12 Lei Pi (UL), External Knowledge Absorption in Chinese SMEs
2021-13 Bob R. Schadenberg (UT), Robots for Autistic Children: Understanding and Facilitating Predictability for Engagement in Learning
2021-14 Negin Samaeemofrad (UL), Business Incubators: The Impact of Their Support
2021-15 Onat Ege Adali (TU/e), Transformation of Value Propositions into Resource Re-Configurations through the Business Services Paradigm
2021-16 Esam A. H. Ghaleb (UM), Bimodal emotion recognition from audio-visual cues
2021-17 Dario Dotti (UM), Human Behavior Understanding from motion and bodily cues using deep neural networks
2021-18 Remi Wieten (UU), Bridging the Gap Between Informal Sense-Making Tools and Formal Systems - Facilitating the Construction of Bayesian Networks and Argumentation Frameworks
2021-19 Roberto Verdecchia (VU), Architectural Technical Debt: Identification and Management
2021-20 Masoud Mansoury (TU/e), Understanding and Mitigating Multi-Sided Exposure Bias in Recommender Systems
2021-21 Pedro Thiago Timbó Holanda (CWI), Progressive Indexes

2021-22 Sihang Qiu (TUD), Conversational Crowdsourcing
2021-23 Hugo Manuel Proença (LIACS), Robust rules for prediction and description
2021-24 Kaijie Zhu (TUE), On Efficient Temporal Subgraph Query Processing
2021-25 Eoin Martino Grua (VUA), The Future of E-Health is Mobile: Combining AI and Self-Adaptation to Create Adaptive E-Health Mobile Applications
2021-26 Benno Kruit (CWI & VU), Reading the Grid: Extending Knowledge Bases from Human-readable Tables
2021-27 Jelte van Waterschoot (UT), Personalized and Personal Conversations: Designing Agents Who Want to Connect With You
2021-28 Christoph Selig (UL), Understanding the Heterogeneity of Corporate Entrepreneurship Programs

2022

2022-1 Judith van Stegeren (UT), Flavor text generation for role-playing video games
2022-2 Paulo da Costa (TU/e), Data-driven Prognostics and Logistics Optimisation: A Deep Learning Journey
2022-3 Ali el Hassouni (VUA), A Model A Day Keeps The Doctor Away: Reinforcement Learning For Personalized Healthcare
2022-4 Ünal Aksu (UU), A Cross-Organizational Process Mining Framework
2022-5 Shiwei Liu (TU/e), Sparse Neural Network Training with In-Time Over-Parameterization
2022-6 Reza Refaei Afshar (TU/e), Machine Learning for Ad Publishers in Real Time Bidding
2022-7 Sambit Praharaj (OU), Measuring the Unmeasurable? Towards Automatic Co-located Collaboration Analytics
2022-8 Maikel L. van Eck (TU/e), Process Mining for Smart Product Design
2022-9 Oana Andreea Inel (VUA), Understanding Events: A Diversity-driven Human-Machine Approach
2022-10 Felipe Moraes Gomes (TUD), Examining the Effectiveness of Collaborative Search Engines
2022-11 Mirjam de Haas (TiU), Staying engaged in child-robot interaction, a quantitative approach to studying preschoolers' engagement with robots and tasks during second-language tutoring
2022-12 Guanyi Chen (UU), Computational Generation of Chinese Noun Phrases
2022-13 Xander Wilcke (VUA), Machine Learning on Multimodal Knowledge Graphs: Opportunities, Challenges, and Methods for Learning on Real-World Heterogeneous and Spatially-Oriented Knowledge
2022-14 Michiel Overeem (UU), Evolution of Low-Code Platforms
2022-15 Jelmer Jan Koorn (UU), Work in Process: Unearthing Meaning using Process Mining
2022-16 Pieter Gijsbers (TU/e), Systems for AutoML Research
2022-17 Laura van der Lubbe (VUA), Empowering vulnerable people with serious games and gamification
2022-18 Paris Mavromoustakos Blom (TiU), Player Affect Modelling and Video Game Personalisation
2022-19 Bilge Yigit Ozkan (UU), Cybersecurity Maturity Assessment and Standardisation
2022-20 Fakhra Jabeen (VUA), Dark Side of the Digital World - Computational Analysis of Negative Human Behaviors on Social Media
2022-21 Seethu Mariyam Christopher (UM), Intelligent Toys for Physical and Cognitive Assessments
2022-22 Alexandra Sierra Rativa (TiU), Virtual Character Design and its potential to foster Empathy, Immersion, and Collaboration Skills in Video Games and Virtual Reality Simulations
2022-23 Ilir Kola (TUD), Enabling Social Situation Awareness in Support Agents
2022-24 Samaneh Heidari (UU), Agents with Social Norms and Values - A framework for agent based social simulations with social norms and personal values
2022-25 Anna L.D. Latour (LU), Optimal decision-making under constraints and uncertainty
2022-26 Anne Dirkson (LU), Knowledge Discovery from Patient Forums: Gaining novel medical insights from patient experiences
2022-27 Christos Athanasiadis (UM), Emotion-aware cross-modal domain adaptation in video sequences
2022-28 Onuralp Ulusoy (UU), Privacy in Collaborative Systems
2022-29 Jan Kolkmeier (UT-EEMCS), From Head Transform to Mind Transplant: Social Interactions in Mixed Reality
2022-30 Dean De Leo (CWI), Analysis of Dynamic Graphs on Sparse Arrays
2022-31 Konstantinos Traganos (TU/e), Tackling Complexity in Smart Manufacturing with Advanced Manufacturing Process Management
2022-32 Cezara Pastrav (UU), Social simulation for socio-ecological systems
2022-33 Brinn Hekkelman (CWI/TUD), Fair Mechanisms for Smart Grid Congestion Management
2022-34 Nimat Ullah (VUA), Mind Your Behaviour: Computational Modelling of Emotion & Desire Regulation for Behaviour Change
2022-35 Mike E.U. Ligthart (VU), Shaping the Child-Robot Relationship: Interaction Design Patterns for a Sustainable Interaction

2023

2023-1 Bojan Simoski (VUA), Untangling the Puzzle of Digital Health Interventions
2023-2 Mariana Rachel Dias da Silva (TiU), Grounded or in flight? What our bodies can tell us about the whereabouts of our thoughts

2023-3 Shabnam Najafian (TU Delft), User Modeling for Privacy-preserving Explanations in Group Recommendations

2023-4 Gineke Wiggers (Leiden University), The Relevance of Impact: bibliometric-enhanced legal information retrieval

2023-5 P.A. (Anton), Bouter (CWI), Optimal Mixing Evolutionary Algorithms for Large-Scale Real-Valued Optimization Including Real-World Medical Applications

2023-6 António Pereira Barata (Leiden University), Reliable and Fair Machine Learning for Risk Assessment

2023-7 Tianjin Huang (TU/e), The Roles of Adversarial Examples on Trustworthiness of Deep Learning

2023-8 Lu Yin (TU/e), Knowledge Elicitation using Psychometric Learning

2023-9 Xu Wang (VUA), Scientific Dataset Recommendation with Semantic Techniques

2023-10 Dennis J.N.J. Soemers (UM), Learning State-Action Features for General Game Playing

2023-11 Fawad Taj (VUA), Towards Motivating Machines: Computational Modeling of the Mechanism of Actions for Effective Digital Health Behavior Change Applications

2023-12 Tessel Bogaard (VUA), Using Metadata to Understand Search Behavior in Digital Libraries

2023-13 Injy Sarhan (UU), Open Information Extraction for Knowledge Representation

2023-14 Selma Čaušević (TU Delft), Energy resilience through self-organization

2023-15 Alvaro Henrique Chaim Correia (TU/e), Insights on Learning Tractable Probabilistic Graphical Models

2023-16 Peter Blomsma (TiU), Building Embodied Conversational Agents: Observations on human nonverbal behaviour as a resource for the development of artificial characters

2023-17 Meike Nauta (UT), Explainable AI and Interpretable Computer Vision – From Oversight to Insight

2023-18 Gustavo Penha (TU Delft), Designing and Diagnosing Models for Conversational Search and Recommendation

2023-19 George Aalbers (TiU), Digital Traces of the Mind: Using Smartphones to Capture Signals of Well-Being in Individuals

2023-20 Arkadiy Dushatskiy (TU Delft), Expensive Optimization with Model-Based Evolutionary Algorithms applied to Medical Image Segmentation using Deep Learning

2023-21 Gerrit Jan de Bruin (Leiden University), Network Analysis Methods for Smart Inspection in the Transport Domain

2023-22 Alireza Shojaifar (UU), Volitional Cybersecurity

2023-23 Theo Theunissen (UU), Documentation in Continuous Software Development

2023-24 Agathe Balayn (TU Delft), Practices Towards Hazardous Failure Diagnosis in Machine Learning

2023-25 Jurian Baas (UU), Entity Resolution on Historical Knowledge Graphs

2023-26 Loek Tonnaer (TU/e), Linearly Symmetry-Based Disentangled Representations and their Out-of-Distribution Behaviour

2023-27 Ghada Sokar (TU/e), Learning Continually Under Changing Data Distributions

2023-28 Floris den Hengst (VUA), Learning to Behave: Reinforcement Learning in Human Contexts

2024

2024-1 Daphne Miedema (TU/e), On SQL Learning: Disentangling concepts in data systems education

2024-2 Emile van Krieken (VUA), Optimisation in Neurosymbolic Learning Systems

2024-3 Feri Wijayanto (RUN), Automated Model Selection for Rasch and Mediation Analysis

2024-4 Mike Huisman (Leiden University), Understanding Deep Meta-Learning

2024-5 Yiyong Gou (UM), Aerial Robotic Operations: Multi-environment Cooperative Inspection & Construction Crack Autonomous Repair

2024-6 Azqa Nadeem (TU Delft), Understanding Adversary Behavior via XAI: Leveraging Sequence Clustering to Extract Threat Intelligence

2024-7 Parisa Shayan (TiU), Modeling User Behavior in Learning Management Systems

2024-8 Xin Zhou (UvA), From Empowering to Motivating: Enhancing Policy Enforcement through Process Design and Incentive Implementation

2024-9 Giso Dal (UT), Probabilistic Inference Using Partitioned Bayesian Networks

colophon

This document was typeset using the typographical look-and-feel `classicthesis` developed by André Miede and Ivo Pletikosić. The style was inspired by Robert Bringhurst's seminal book on typography "*The Elements of Typographic Style*". `classicthesis` is available for both LaTeX and LyX:

https://bitbucket.org/amiede/classicthesis/

Final Version as of February 8, 2024 (`classicthesis` version 4.6).