

RESEARCH CENTRE

**Inria Centre  
at Université Côte d'Azur**

IN PARTNERSHIP WITH:

CNRS, Université Côte d'Azur

2023

ACTIVITY REPORT

Project-Team

WIMMICS

**Web-Instrumented Man-Machine  
Interactions, Communities and Semantics**

IN COLLABORATION WITH: Laboratoire informatique, signaux systèmes  
de Sophia Antipolis (I3S)

**DOMAIN**

**Perception, Cognition and Interaction**

**THEME**

**Data and Knowledge Representation and  
Processing**

*Inria*

# Contents

<b>Project-Team WIMMICS</b>	<b>1</b>
<b>1 Team members, visitors, external collaborators</b>	<b>4</b>
<b>2 Overall objectives</b>	<b>6</b>
2.1 Context and Objectives . . . . .	6
2.2 Research Topics . . . . .	7
<b>3 Research program</b>	<b>8</b>
3.1 Users Modeling and Designing Interaction on the Web and with AI systems . . . . .	8
3.2 Communities and Social Media Interactions and Content Analysis on the Web and Linked Data . . . . .	8
3.3 Vocabularies, Semantic Web and Linked Data Based Knowledge Representation and Extraction of Knowledge Graphs on the Web . . . . .	8
3.4 Artificial Intelligence Processing: Learning, Analyzing and Reasoning on Heterogeneous Knowledge Graphs . . . . .	9
<b>4 Application domains</b>	<b>9</b>
4.1 Social Semantic Web . . . . .	9
4.2 Linked Data on the Web and on Intranets . . . . .	9
4.3 Assisting Web-based Epistemic Communities . . . . .	10
4.4 Linked Data for a Web of Diversity . . . . .	10
4.5 Artificial Web Intelligence . . . . .	10
4.6 Human-Data Interaction (HDI) on the Web . . . . .	11
4.7 Web-augmented interactions with the world . . . . .	11
4.8 Analysis of scientific co-authorship . . . . .	11
<b>5 Highlights of the year</b>	<b>12</b>
5.1 Awards . . . . .	12
<b>6 New software, platforms, open data</b>	<b>13</b>
6.1 New software . . . . .	13
6.1.1 ACTA . . . . .	13
6.1.2 ARViz . . . . .	13
6.1.3 Attune . . . . .	14
6.1.4 CORESE-Core . . . . .	14
6.1.5 CORESE-GUI . . . . .	14
6.1.6 CORESE-Server . . . . .	15
6.1.7 CORESE-Command . . . . .	15
6.1.8 CREEP semantic technology . . . . .	16
6.1.9 CROBORA . . . . .	16
6.1.10 DBpedia . . . . .	16
6.1.11 Fuzzy labelling argumentation module . . . . .	17
6.1.12 GUsT-3D . . . . .	17
6.1.13 IndeGx . . . . .	17
6.1.14 ISSA-pipeline . . . . .	18
6.1.15 ISSA Visualization Web Application . . . . .	18
6.1.16 KartoGraphI . . . . .	18
6.1.17 Licentia . . . . .	18
6.1.18 SPARQL Micro-services . . . . .	19
6.1.19 Metadatamatic . . . . .	20
6.1.20 MGExplorer . . . . .	20
6.1.21 Morph-xR2RML . . . . .	20
6.1.22 Muvin . . . . .	21

6.1.23	wam-studio	21
6.1.24	WebAudio tube guitar amp sims CLEAN, DISTO and METAL MACHINES	22
6.2	Open data	22
6.2.1	CyberAgressionAdo	22
6.2.2	ElecDeb60to20	23
6.2.3	ISSA Agritrop Dataset	23
6.2.4	WeKG-MF	23
6.2.5	WheatGenomicsSLKG	24
6.2.6	Pharmacogenomics datasets for Ontology Matching	24
6.2.7	DBpedia.fr : French chapter of the DBpedia knowledge graph dataset	24
<b>7</b>	<b>New results</b>	<b>25</b>
7.1	User Modeling and Designing Interaction	25
7.1.1	Spatio-Temporal Data Visualization	25
7.1.2	Tracking and Exploring Audiovisual Archives about the European Union	25
7.1.3	Interaction with extended reality	25
7.1.4	Incremental visual exploration of linked data	26
7.1.5	Development of an extensible User Profile model	27
7.1.6	Interactive WebAudio applications	27
7.1.7	KartoGraphI	28
7.2	Communities and Social Interactions Analysis	28
7.2.1	Autonomous agents in a social and ubiquitous Web	28
7.2.2	Online Hate Detection in Conversational Data	28
7.2.3	Hate Speech Target Community Detection	29
7.2.4	An In-depth Analysis of Implicit and Subtle Hate Speech Messages	29
7.2.5	Generating Adversarial Examples for Implicit Hate Speech Detection	30
7.2.6	Bridging Implicit and Explicit Hate Speech Embedding Representations	30
7.2.7	Fallacious Argument Classification in Political Debates	30
7.2.8	DISPUTool 2.0: A Modular Architecture for Multi-Layer Argumentative Analysis of Political Debates	31
7.2.9	Argument-based Detection and Classification of Fallacies in Political Debates	31
7.2.10	Augmenting participation, co-creation, trust and transparency in Deliberative Democracy at all scales through Argument Mining	31
7.2.11	Unveiling fake news through argumentative evidence	32
7.3	Vocabularies, Semantic Web and Linked Data Based Knowledge Representation and Artificial Intelligence Formalisms on the Web	32
7.3.1	Semantic Web for Biodiversity	32
7.3.2	Semantic Web for Life Sciences	33
7.3.3	Evolutionary agent-based evaluation of the sustainability of different knowledge sharing strategies in open multi-agent systems	33
7.3.4	W3C Data activity and AC Rep	33
7.3.5	AutomaZoo: Automatic annotation of an ancient Zoological Corpus	34
7.3.6	Contributions to Agronomical knowledge graphs	34
7.3.7	WheatGenomicsSLKG: a Knowledge graph for wheat genomics studies	35
7.3.8	Ontology engineering: tooling and methodology	35
7.4	Analyzing and Reasoning on Heterogeneous Semantic Graphs	36
7.4.1	Corese Semantic Web Factory	36
7.4.2	SHACL Extension	36
7.4.3	Automatic Assessment of Text Difficulty	37
7.4.4	On the Automatic Assessment of Natural Language Expert Explanations in Medicine	37
7.4.5	Explanatory Argumentation in Natural Language for Correct and Incorrect Medical Diagnoses	37
7.4.6	ACTA Module Upgrade	38
7.4.7	Qualitative evaluation of arguments in persuasive essays	38
7.4.8	RDF Mining	38

7.4.9	Capturing Geospatial Knowledge from Real-Estate Advertisements	39
7.4.10	ISSA: semantic indexing of scientific articles and advanced services	39
7.4.11	IndeGx	40
7.4.12	Metadatamatic	40
7.4.13	Learning Pattern-Based Extractors from Natural Language and Knowledge Graphs: Applying Large Language Models to Wikipedia and Linked Open Data	40
7.4.14	AI agent to convert natural language questions into SPARQL queries	41
7.4.15	Hybridizing machine learning and knowledge graphs: vision, model evaluation, and knowledge injection	42
7.4.16	Semantic Overlay Network for Linked Data Access	42
<b>8</b>	<b>Bilateral contracts and grants with industry</b>	<b>42</b>
8.1	Bilateral contracts with industry	42
<b>9</b>	<b>Partnerships and cooperations</b>	<b>44</b>
9.1	International initiatives	44
9.1.1	Associate Teams in the framework of an Inria International Lab or in the framework of an Inria International Program	44
9.1.2	Participation in other International Programs	45
9.2	International research visitors	45
9.2.1	Visits of international scientists	45
9.2.2	Visits to international teams	46
9.3	European initiatives	46
9.3.1	H2020 projects	46
9.3.2	Other european programs/initiatives	46
9.4	National initiatives	47
9.5	Regional initiatives	50
<b>10</b>	<b>Dissemination</b>	<b>51</b>
10.1	Promoting scientific activities	51
10.1.1	Scientific events: organisation	51
10.1.2	Scientific events: selection	51
10.1.3	Journal	52
10.1.4	Invited talks	53
10.1.5	Leadership within the scientific community	54
10.1.6	Scientific expertise	54
10.1.7	Research administration	54
10.1.8	Promoting Open Science practices	54
10.2	Teaching - Supervision - Juries	54
10.2.1	Teaching	54
10.2.2	Supervision	57
10.2.3	Juries	59
10.3	Popularization	60
10.3.1	Articles and contents	60
10.3.2	Education	60
10.3.3	Interventions	60
<b>11</b>	<b>Scientific production</b>	<b>60</b>
11.1	Major publications	60
11.2	Publications of the year	62
11.3	Other	70
11.4	Cited publications	70

## Project-Team WIMMICS

*Creation of the Project-Team: 2013 July 01*

### Keywords

#### Computer sciences and digital sciences

- A1.2.9. – Social Networks
- A1.3.1. – Web
- A3.1.1. – Modeling, representation
- A3.1.2. – Data management, quering and storage
- A3.1.3. – Distributed data
- A3.1.4. – Uncertain data
- A3.1.6. – Query optimization
- A3.1.7. – Open data
- A3.1.10. – Heterogeneous data
- A3.1.11. – Structured data
- A3.2. – Knowledge
- A3.2.1. – Knowledge bases
- A3.2.2. – Knowledge extraction, cleaning
- A3.2.3. – Inference
- A3.2.4. – Semantic Web
- A3.2.5. – Ontologies
- A3.2.6. – Linked data
- A3.3.2. – Data mining
- A3.4. – Machine learning and statistics
- A3.4.1. – Supervised learning
- A3.4.6. – Neural networks
- A3.4.8. – Deep learning
- A3.5. – Social networks
- A3.5.1. – Analysis of large graphs
- A3.5.2. – Recommendation systems
- A5.1. – Human-Computer Interaction
- A5.1.1. – Engineering of interactive systems
- A5.1.2. – Evaluation of interactive systems
- A5.1.9. – User and perceptual studies
- A5.2. – Data visualization
- A5.7.2. – Music
- A5.8. – Natural language processing

- A7.1.3. – Graph algorithms
- A7.2.2. – Automated Theorem Proving
- A8.2.2. – Evolutionary algorithms
- A9.1. – Knowledge
- A9.2. – Machine learning
- A9.4. – Natural language processing
- A9.6. – Decision support
- A9.7. – AI algorithmics
- A9.8. – Reasoning
- A9.9. – Distributed AI, Multi-agent
- A9.10. – Hybrid approaches for AI

### **Other research topics and application domains**

- B1.2.2. – Cognitive science
- B2. – Health
- B5.1. – Factory of the future
- B5.6. – Robotic systems
- B5.8. – Learning and training
- B6.3.1. – Web
- B6.3.2. – Network protocols
- B6.3.4. – Social Networks
- B6.4. – Internet of things
- B6.5. – Information systems
- B8.5. – Smart society
- B8.5.1. – Participative democracy
- B9. – Society and Knowledge
- B9.1. – Education
- B9.1.1. – E-learning, MOOC
- B9.1.2. – Serious games
- B9.2. – Art
- B9.2.1. – Music, sound
- B9.3. – Medias
- B9.5.1. – Computer science
- B9.5.6. – Data science
- B9.6. – Humanities
- B9.6.1. – Psychology
- B9.6.2. – Juridical science
- B9.6.5. – Sociology
- B9.6.7. – Geography
- B9.6.8. – Linguistics

B9.6.9. – Political sciences

B9.6.10. – Digital humanities

B9.7. – Knowledge dissemination

B9.7.1. – Open access

B9.7.2. – Open data

B9.9. – Ethics

B9.10. – Privacy

# 1 Team members, visitors, external collaborators

## Research Scientists

- Fabien Gandon [Team leader, INRIA, Senior Researcher, HDR]
- Victor David [INRIA, ISFP, from Oct 2023]
- Serena Villata [CNRS, Senior Researcher, HDR]

## Faculty Members

- Marco Alba Winckler [UNIV COTE D'AZUR, Professor, HDR]
- Michel Buffa [UNIV COTE D'AZUR, Professor, HDR]
- Elena Cabrio [UNIV COTE D'AZUR, Professor, HDR]
- Pierre-Antoine Champin [UNIV LYON, Associate Professor, Delegation, HDR]
- Catherine Faron [UNIV COTE D'AZUR, Professor, HDR]
- Aline Menin [UNIV COTE D'AZUR, Associate Professor]
- Pierre Monnin [UNIV COTE D'AZUR, Associate Professor, from Sep 2023, EFELIA AI Fellow]
- Amaya Nogales Gomez [UNIV COTE D'AZUR, Associate Professor, until Jul 2023, 3AI Fellow]
- Anaïs Ollagnier [UNIV COTE D'AZUR, Associate Professor, from Feb 2023, EFELIA AI Fellow]
- Andrea Tettamanzi [UNIV COTE D'AZUR, Professor, HDR]

## Post-Doctoral Fellows

- Ankica Barisic [UNIV COTE D'AZUR, until Oct 2023]
- Cristian Cardellino [UNIV COTE D'AZUR, from Sep 2023]
- Nadia Yacoubi Ayadi [CNRS, until Aug 2023]

## PhD Students

- Ali Ballout [UNIV COTE D'AZUR]
- Lucie Cadorel [KINAXIA, CIFRE, (now SEPTEO)]
- Dupuy-Rony Charles [DORIANE, CIFRE]
- Remi Felin [UNIV COTE D'AZUR]
- Pierpaolo Goffredo [CNRS]
- Mina Ilhan [UNIV COTE D'AZUR, GREDEG]
- Santiago Marro [CNRS, until Nov 2023]
- Benjamin Molinet [UNIV COTE D'AZUR]
- Nicolas Ocampo [UNIV COTE D'AZUR]
- Clement Quere [UNIV COTE D'AZUR]
- Shihong Ren [UNIV JEAN MONNET]



- Celian Ringwald [INRIA]
- Ekaterina Sviridova [UNIV COTE D'AZUR, from Aug 2023]
- Maroua Tikat [UNIV COTE D'AZUR]
- Xiaou Wang [CNRS]

### **Technical Staff**

- Arnaud Barbe [INRIA, Engineer, from Aug 2023]
- Anna Bobasheva [INRIA, Engineer, from Dec 2023]
- Remi Ceres [INRIA, Engineer]
- Mariana Eugenia Chaves Espinoza [UNIV COTE D'AZUR, Engineer, from Apr 2023]
- Theo Alkibiades Collias [UNIV COTE D'AZUR, Engineer, from Feb 2023]
- Antoine De Smidt [CNRS, Engineer]
- Molka Dhouib [INRIA, Engineer, from Aug 2023]
- Maxime Lecoq [INRIA, Engineer]
- Christopher Leturc [INRIA, Engineer, until Aug 2023]
- Pierre Maillot [INRIA, Engineer]
- Franck Michel [CNRS, Engineer]
- Iliana Petrova [INRIA, Engineer, until Jun 2023]
- Nicolas Robert [INRIA, Engineer, from Sep 2023]
- Ekaterina Sviridova [UNIV COTE D'AZUR, Engineer, until Aug 2023]

### **Interns and Apprentices**

- Nicolas Audoux [INRIA, Intern, from Jun 2023 until Sep 2023]
- Hugo Carton [INRIA, Apprentice, from Nov 2023]
- Deborah Dore [UNIV SAPIENZA, Intern, from Sep 2023]
- Erwan Hain [UNIV COTE D'AZUR, Intern, from Apr 2023 until Jun 2023]
- Raphael Julien [UNIV COTE D'AZUR, Intern, from Mar 2023 until Aug 2023]
- Ekaterina Kolos [CNRS, Intern]
- Thomas Mac Vicar [YNOV CAMPUS, Intern, from May 2023 until Aug 2023]
- Quentin Plet [INRIA, Intern, from Mar 2023 until Aug 2023]
- Rim Richa [UNIV COTE D'AZUR, Intern, until Jun 2023]
- Martina Rossini [CNRS, Intern, until Feb 2023]
- Antoine Vidal Mazuy [INRIA, Apprentice, until Aug 2023]
- Manuel Vimercati [UNIV MILANO, Intern]

### Administrative Assistants

- Delphine Robache [INRIA]
- Lionel Tavanti [UNIV COTE D'AZUR, I3S]

### Visiting Scientists

- Greta Damo [UNIV COTE D'AZUR, from Sep 2023]
- Roberto Demaria [UNI TORINO, from Jul 2023 until Sep 2023]
- Cristhian Figueroa [UNIV DEL CAUCA, from Jun 2023 until Jul 2023, Assistant Professor]
- Yusuke Higuchi [JPO (Japan Patent Office), until Aug 2023]
- Dario Malchiodi [UNIV MILANO, from Oct 2023]
- Smaranda Muresan [Columbia University, from May 2023 until Jun 2023, Research Scientist]
- Matteo Palmonari [UNIV MILANO BICOCCA, from Jun 2023 until Jul 2023, Professor]

### External Collaborators

- Hanna Abi Akl [DSTI, from Feb 2023]
- Helena Bonaldi [UNIV TRENTE, from Nov 2023]
- Andrei Ciortea [UNIV ST GALLEN, Assistant Professor]
- Olivier Corby [INRIA, Retired Researcher]
- Alain Giboin [INRIA, Retired Researcher]
- Freddy Lecue [JP MORGAN, AI Research Director]
- Oscar Rodriguez Rocha [TEACH ON MARS]
- Stefan Sarkadi [KINGS COLLEGE LONDON, Research Fellow]

## 2 Overall objectives

### 2.1 Context and Objectives

The World Wide Web has transformed into a virtual realm where individuals and software interact in diverse communities. The Web has the potential to become the collaborative space for both natural and artificial intelligence, thereby posing the challenge of supporting these global interactions. The large-scale, mixed interactions inherent in this scenario present a plethora of issues that must be addressed through multidisciplinary approaches<sup>[110]</sup>.

One particular problem is to reconcile the formal semantics of computer science (such as logics, ontologies, typing systems, protocols, etc.) on which the Web architecture is built, with the soft semantics of human interactions (such as posts, tags, status, relationships, etc.) that form the foundation of Web content. This requires a holistic approach that considers both the technical and social aspects of the Web, in order to ensure that the interactions between computational and natural intelligence are seamless and meaningful.

*Wimmics proposes a range of models and methods to bridge the gap between formal semantics and social semantics on the World Wide Web<sup>[111]</sup>, in order to address some of the challenges associated with constructing a universal space that connects various forms of intelligence.*

From a formal modeling point of view, one of the consequences of the evolutions of the Web is that the initial graph of linked pages has been joined by a growing number of other graphs. This initial graph

is now mixed with sociograms capturing the social network structure, workflows specifying the decision paths to be followed, browsing logs capturing the trails of our navigation, service compositions specifying distributed processing, open data linking distant datasets, etc. Moreover, these graphs are not available in a single central repository but distributed over many different sources. Some sub-graphs are small and local (e.g. a user's profile on a device), some are huge and hosted on clusters (e.g. Wikipedia), some are largely stable (e.g. thesaurus of Latin), some change several times per second (e.g. social network statuses), etc. Moreover, each type of network of the Web is not an isolated island. Networks interact with each other: the networks of communities influence the message flows, their subjects and types, the semantic links between terms interact with the links between sites and vice-versa, etc.

Not only do we need means to represent and analyze each kind of graphs, we also do need means to combine them and to perform multi-criteria analysis on their combination. Wimmics contributes to these challenges by: (1) proposing multidisciplinary approaches to analyze and model the many aspects of these intertwined information systems, their communities of users and their interactions; (2) formalizing and reasoning on these models using graphs-based knowledge representation from the semantic Web [1] to propose new analysis tools and indicators, and to support new functionalities and better management. In a nutshell, the first research direction looks at models of systems, users, communities and interactions while the second research direction considers formalisms and algorithms to represent them and reason on their representations.

## 2.2 Research Topics

The research objectives of Wimmics can be grouped according to four topics that we identify in reconciling social and formal semantics on the Web:

**Topic 1 - users modeling and designing interaction on the Web and with knowledge graphs:** The general research question addressed by this objective is "*How do we improve our interactions with a semantic and social Web more and more complex and dense?*". Wimmics focuses on specific sub-questions: "How can we capture and model the users' characteristics?" "How can we represent and reason with the users' profiles?" "How can we adapt the system behaviors as a result?" "How can we design new interaction means?" "How can we evaluate the quality of the interaction designed?". This topic includes a long-term research direction in Wimmics on information visualization of semantic graphs on the Web. The general research question addressed in this last objective is "*How to represent the inner and complex relationships between data obtained from large and multivariate knowledge graph?*". Wimmics focuses on several sub-questions: "Which visualization techniques are suitable (from a user point of view) to support the exploration and the analysis of large graphs?" "How to identify the new knowledge created by users during the exploration of knowledge graph?" "How to formally describe the dynamic transformations allowing to convert raw data extracted from the Web into meaningful visual representations?" "How to guide the analysis of graphs that might contain data with diverse levels of accuracy, precision and interestingness to the users?"

**Topic 2 - communities and social interactions and content analysis on the Web:** The general question addressed in this second objective is "*How can we manage the collective activity on social media?*". Wimmics focuses on the following sub-questions: "How do we analyze the social interaction practices and the structures in which these practices take place?" "How do we capture the social interactions and structures?" "How can we formalize the models of these social constructs?" "How can we analyze and reason on these models of the social activity?"

**Topic 3 - vocabularies, semantic Web and linked data based knowledge extraction and representation with knowledge graphs on the Web:** The general question addressed in this third objective is "*What are the needed schemas and extensions of the semantic Web formalisms for our models?*". Wimmics focuses on several sub-questions: "What kinds of formalism are the best suited for the models of the previous section?" "What are the limitations and possible extensions of existing formalisms?" "What are the missing schemas, ontologies, vocabularies?" "What are the links and possible combinations between existing formalisms?" We also address the question of knowledge extraction and especially AI and NLP methods to extract knowledge from text. In a nutshell, an important part of this objective is to formalize the models identified in the previous objectives as typed graphs and to populate them in order for software to exploit these knowledge graphs in their processing (in the next objective).

**Topic 4 - artificial intelligence processing: learning, analyzing and reasoning on heterogeneous**

**semantic graphs on the Web:** The general research question addressed in this objective is “*What are the algorithms required to analyze and reason on the heterogeneous graphs we obtained?*”. Wimmics focuses on several sub-questions: “How do we analyze graphs of different types and their interactions?” “How do we support different graph life-cycles, calculations and characteristics in a coherent and understandable way?” “What kind of algorithms can support the different tasks of our users?”.

### 3 Research program

#### 3.1 Users Modeling and Designing Interaction on the Web and with AI systems

Wimmics focuses on interactions of ordinary users with ontology-based knowledge systems, with a preference for semantic Web formalisms and Web 2.0 applications. We specialize in interaction design and evaluation methods to Web application tasks such as searching, browsing, contributing or protecting data. The team is especially interested in using semantics in assisting the interactions. We propose knowledge graph representations and algorithms to support interaction adaptation, for instance for context-awareness or intelligent interactions with machine. We propose and evaluate Web-based visualization techniques for linked data, querying, reasoning, explaining and justifying. Wimmics also integrates natural language processing approaches to support natural language based interactions. We rely on cognitive studies to build models of the system, the user and the interactions between users through the system, in order to support and improve these interactions. We extend the user modeling technique known as *Personas* where user models are represented as specific, individual humans. *Personas* are derived from significant behavior patterns (i.e., sets of behavioral variables) elicited from interviews with and observations of users (and sometimes customers) of the future product. Our user models specialize *Personas* approaches to include aspects appropriate to Web applications. Wimmics also extends user models to capture very different aspects (e.g. emotional states).

#### 3.2 Communities and Social Media Interactions and Content Analysis on the Web and Linked Data

The domain of social network analysis is a whole research domain in itself and Wimmics targets what can be done with typed graphs, knowledge representations and social models. We also focus on the specificity of social Web and semantic Web applications and in bridging and combining the different social Web data structures and semantic Web formalisms. Beyond the individual user models, we rely on social studies to build models of the communities, their vocabularies, activities and protocols in order to identify where and when formal semantics is useful. We propose models of collectives of users and of their collaborative functioning extending the collaboration personas and methods to assess the quality of coordination interactions and the quality of coordination artifacts. We extend and compare community detection algorithms to identify and label communities of interest with the topics they share. We propose mixed representations containing social semantic representations (e.g. folksonomies) and formal semantic representations (e.g. ontologies) and propose operations that allow us to couple them and exchange knowledge between them. Moving to social interaction we develop models and algorithms to mine and integrate different yet linked aspects of social media contributions (opinions, arguments and emotions) relying in particular on natural language processing and argumentation theory. To complement the study of communities we rely on multi-agent systems to simulate and study social behaviors. Finally we also rely on Web 2.0 principles to provide and evaluate social Web applications.

#### 3.3 Vocabularies, Semantic Web and Linked Data Based Knowledge Representation and Extraction of Knowledge Graphs on the Web

For all the models we identified in the previous sections, we rely on and evaluate knowledge representation methodologies and theories, in particular ontology-based modeling. We also propose models and formalisms to capture and merge representations of different levels of semantics (e.g. formal ontologies and social folksonomies). The important point is to allow us to capture those structures precisely and flexibly and yet create as many links as possible between these different objects. We propose vocabularies

and semantic Web formalizations for all the aspects that we model and we consider and study extensions of these formalisms when needed. The results have all in common to pursue the representation and publication of our models as linked data. We also contribute to the extraction, transformation and linking of existing resources (informal models, databases, texts, etc.) to publish knowledge graphs on the Semantic Web and as Linked Data. Examples of aspects we formalize include: user profiles, social relations, linguistic knowledge, bio-medical data, business processes, derivation rules, temporal descriptions, explanations, presentation conditions, access rights, uncertainty, emotional states, licenses, learning resources, etc. At a more conceptual level we also work on modeling the Web architecture with philosophical tools so as to give a realistic account of identity and reference and to better understand the whole context of our research and its conceptual cornerstones.

### **3.4 Artificial Intelligence Processing: Learning, Analyzing and Reasoning on Heterogeneous Knowledge Graphs**

One of the characteristics of Wimmics is to rely on graph formalisms unified in an abstract graph model and operators unified in an abstract graph machine to formalize and process semantic Web data, Web resources, services metadata and social Web data. In particular Corese, the core software of Wimmics, maintains and implements that abstraction. We propose algorithms to process the mixed representations of the previous section. In particular we are interested in allowing cross-enrichment between them and in exploiting the life cycle and specificity of each one to foster the life-cycles of the others. Our results all have in common to pursue analyzing and reasoning on heterogeneous knowledge graphs issued from social and semantic Web applications. Many approaches emphasize the logical aspect of the problem especially because logics are close to computer languages. We defend that the graph nature of Linked Data on the Web and the large variety of types of links that compose them call for typed graphs models. We believe the relational dimension is of paramount importance in these representations and we propose to consider all these representations as fragments of a typed graph formalism directly built above the Semantic Web formalisms. Our choice of a graph based programming approach for the semantic and social Web and of a focus on one graph based formalism is also an efficient way to support interoperability, genericity, uniformity and reuse.

## **4 Application domains**

### **4.1 Social Semantic Web**

A number of evolutions have changed the face of information systems in the past decade but the advent of the Web is unquestionably a major one and it is here to stay. From an initial wide-spread perception of a public documentary system, the Web as an object turned into a social virtual space and, as a technology, grew as an application design paradigm (services, data formats, query languages, scripting, interfaces, reasoning, etc.). The universal deployment and support of its standards led the Web to take over nearly all of our information systems. As the Web continues to evolve, our information systems are evolving with it.

Today in organizations, not only almost every internal information system is a Web application, but these applications more and more often interact with external Web applications. The complexity and coupling of these Web-based information systems call for specification methods and engineering tools. From capturing the needs of users to deploying a usable solution, there are many steps involving computer science specialists and non-specialists.

We defend the idea of relying on Semantic Web formalisms to capture and reason on the models of these information systems supporting the design, evolution, interoperability and reuse of the models and their data as well as the workflows and the processing.

### **4.2 Linked Data on the Web and on Intranets**

With billions of triples online (see Linked Open Data initiative), the Semantic Web is providing and linking open data at a growing pace and publishing and interlinking the semantics of their schemas. Information systems can now tap into and contribute to this Web of data, pulling and integrating data on demand.

Many organisations also started to use this approach on their intranets leading to what is called linked enterprise data.

A first application domain for us is the publication and linking of data and their schemas through Web architectures. Our results provide software platforms to publish and query data and their schemas, to enrich these data in particular by reasoning on their schemas, to control their access and licenses, to assist the workflows that exploit them, to support the use of distributed datasets, to assist the browsing and visualization of data, etc.

Examples of collaboration and applied projects include: Corese, DBpedia.fr, DekaLog, D2KAB, MonaLIA.

### 4.3 Assisting Web-based Epistemic Communities

In parallel with linked open data on the Web, social Web applications also spread virally (e.g. Facebook growing toward 1.5 billion users) first giving the Web back its status of a social read-write media and then putting it back on track to its full potential of a virtual place where to act, react and interact. In addition, many organizations are now considering deploying social Web applications internally to foster community building, expert cartography, business intelligence, technological watch and knowledge sharing in general.

By reasoning on the Linked Data and the semantics of the schemas used to represent social structures and Web resources, we provide applications supporting communities of practice and interest and fostering their interactions in many different contexts (e-learning, business intelligence, technical watch, etc.).

We use typed graphs to capture and mix: social networks with the kinds of relationships and the descriptions of the persons; compositions of Web services with types of inputs and outputs; links between documents with their genre and topics; hierarchies of classes, thesauri, ontologies and folksonomies; recorded traces and suggested navigation courses; submitted queries and detected frequent patterns; timelines and workflows; etc.

Our results assist epistemic communities in their daily activities such as biologists exchanging results, business intelligence and technological watch networks informing companies, engineers interacting on a project, conference attendees, students following the same course, tourists visiting a region, mobile experts on the field, etc. Examples of collaboration and applied projects: ISSA, TeachOnMars, CREEP, ATTENTION, CIGALA.

### 4.4 Linked Data for a Web of Diversity

We intend to build on our results on explanations (provenance, traceability, justifications) and to continue our work on opinions and arguments mining toward the global analysis of controversies and online debates. One result would be to provide new search results encompassing the diversity of viewpoints and providing indicators supporting opinion and decision making and ultimately a Web of trust. Trust indicators may require collaborations with teams specialized in data certification, cryptography, signature, security services and protocols, etc. This will raise the specific problem of interaction design for security and privacy. In addition, from the point of view of the content, this requires to foster the publication and coexistence of heterogeneous data with different points of views and conceptualizations of the world. We intend to pursue the extension of formalisms to allow different representations of the world to co-exist and be linked and we will pay special attention to the cultural domain and the digital humanities. Examples of collaboration and applied projects: ACTA, DISPUTOOL.

### 4.5 Artificial Web Intelligence

We intend to build on our experience in artificial intelligence (knowledge representation, reasoning) and distributed artificial intelligence (multi-agent systems - MAS) to enrich formalisms and propose alternative types of reasoning (graph-based operations, reasoning with uncertainty, inductive reasoning, non-monotonic, etc.) and alternative architectures for linked data with adequate changes and extensions required by the open nature of the Web. There is a clear renewed interest in AI for the Web in general and for Web intelligence in particular. Moreover, distributed AI and MAS provide both new architectures and

new simulation platforms for the Web. At the macro level, the evolution accelerated with HTML5 toward Web pages as full applications and direct Page2Page communication between browser clearly is a new area for MAS and P2P architectures. Interesting scenarios include the support of a strong decentralization of the Web and its resilience to degraded technical conditions (downscaling the Web), allowing pages to connect in a decentralized way, forming a neutral space, and possibly going offline and online again in erratic ways. At the micro level, one can imagine the place **RDF (Resource Description Framework)** and **SPARQL (SPARQL Protocol and RDF Query Language)** could take as data model and programming model in the virtual machines of these new Web pages and, of course, in the Web servers. RDF is also used to serialize and encapsulate other languages and becomes a pivot language in linking very different applications and aspects of applications. Example of collaboration and applied projects: HyperAgents, DekaLog, AI4EU, AI4Media.

#### **4.6 Human-Data Interaction (HDI) on the Web**

We need more interaction design tools and methods for linked data access and contribution. We intend to extend our work on exploratory search coupling it with visual analytics to assist sense making. It could be a continuation of the Gephi extension that we built targeting more support for non experts to access and analyze data on a topic or an issue of their choice. More generally speaking SPARQL is inappropriate for common users and we need to support a larger variety of interaction means with linked data. We also believe linked data and natural language processing (NLP) have to be strongly integrated to support natural language based interactions. Linked Open Data (LOD) for NLP, NLP for LOD and Natural Dialog Processing for querying, extracting and asserting data on the Web is a priority to democratize its use. Micro accesses and micro contributions are important to ensure public participation and also call for customized interfaces and thus for methods and tools to generate these interfaces. In addition, the user profiles are being enriched now with new data about the user such as his/her current mental and physical state, the emotion he/she just expressed or his/her cognitive performances. Taking into account this information to improve the interactions, change the behavior of the system and adapt the interface is a promising direction. And these human-data interaction means should also be available for “small data”, helping the user to manage his/her personal information and to link it to public or collective one, maintaining his/her personal and private perspective as a personal Web of data. Finally, the continuous knowledge extractions, updates and flows add the additional problem of representing, storing, querying and interacting with dynamic data. Examples of collaboration and applied projects: WASABI, MuvIn, LDViz.

#### **4.7 Web-augmented interactions with the world**

The Web continues to augment our perception and interaction with reality. In particular, Linked Open Data enable new augmented reality applications by providing data sources on almost any topic. The current enthusiasm for the Web of Things, where every object has a corresponding Web resource, requires evolutions of our vision and use of the Web architecture. This vision requires new techniques as the ones mentioned above to support local search and contextual access to local resources but also new methods and tools to design Web-based human devices interactions, accessibility, etc. These new usages are placing new requirements on the Web Architecture in general and on the semantic Web models and algorithms in particular to handle new types of linked data. They should support implicit requests considering the user context as a permanent query. They should also simplify our interactions with devices around us jointly using our personal preferences and public common knowledge to focus the interaction on the vital minimum that cannot be derived in another way. For instance, the access to the Web of data for a robot can completely change the quality of the interactions it can offer. Again, these interactions and the data they require raise problems of security and privacy. Examples of collaboration and applied projects: ALOOF, AZKAR, MoreWAIS.

#### **4.8 Analysis of scientific co-authorship**

Over the last decades, scientific research has matured and diversified. In all areas of knowledge, we observe an increasing number of scientific publications, a rapid development of ever more specialized

conferences and journals, and the creation of dynamic collaborative networks that cross borders and evolve over time. In this context, analyzing scientific publications and the resulting inner co-authorship networks is a major issue for the sustainability of scientific research. To illustrate this, let us consider what happens in the context of the COVID-19 pandemics, when the whole scientific community engaged numerous fields of research to contribute in a common effort to study, understand and fight the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). In order to support the scientific community, many datasets covering the publications about coronaviruses and related diseases have been compiled. In a short time, the number of publications available (over 200,000+ and still increasing) suggests that it is impossible for any researcher to examine every publication and extract the relevant information.

By reasoning on the Linked Data and Web semantic schemas, we investigate methods and tools to assist users on finding relevant publications to answer their research questions. Hereafter we present some example of typical domain questions and how we can contribute to the matter.

- How to find relevant publication in huge datasets? We investigate the use of association rules as a suitable solution to identify relevant scientific publications. By extracting association rules that determine the co-occurrence between terms in a text, it is possible to create clusters of scientific publications that follow a certain pattern; users can focus the search on clusters that contain the terms of interests rather than search the whole dataset.
- How to explain the contents of scientific publications? By reasoning on the Linked Data and Web semantic schemas, we investigate methods for the creation and exploration of argument graphs that describe association and development of ideas in scientific papers.
- How to understand the impact of co-authorship (collaboration of one or more authors) in the development of scientific knowledge? For that, we proposed visualization techniques that allows the description of co-authorship networks describing the clusters of collaborations that evolve over time. Co-authorship networks can inform both collaboration between authors and institutions.

Currently, the analysis of co-publications has been performed over two majors datasets: Hal open archive, the Covid-on-the-Web datasets, and the Agritrop (CIRAD's open dataset).

## 5 Highlights of the year

We have three new young researchers in the team : Victor David (ISFP Inria), Anaïs Ollagnier ([EFELIA Fellow](#)) and Pierre Monnin ([EFELIA Fellow](#))

We have renewed the following 3IA Chairs: Elena Cabrio, Fabien Gandon and Serena Villata.

### 5.1 Awards

The article "Probabilistic Information in SHACL Validation Reports" [61] by Rémi Felin, Catherine Faron and Andrea G. B. Tettamanzi receives the exæquo best research paper award at the European Conference for Semantic Web (ESWC 2023) and the best spotlight paper award at Ingénierie des Connaissances (IC 2023).

The article "An Integrated Framework for Understanding Multimodal Embodied Experiences in Interactive Virtual Reality" [79] by Florent Robert, Hui-Yin Wu, Lucile Sassatelli, Stephen Ramanoël, Auriane Gros, and Marco Winckler, receives the best paper award at the ACM International Conference on Interactive Media Experiences (ACM IMX 2023).

The paper "Driver Model for Take-Over-Request in Autonomous Vehicles" [54] by Ankica Barisic, Pierre Sigrist, Sylvain Oliver, Aurélien Sciarra, Marco Winckler, receives the best paper award at the Workshop HAAPIE 2023: Human Aspects in Adaptive and Personalized Interactive Environments, held in conjunction with the 31st ACM Conference on User Modeling, Adaptation and Personalization, UMAP 2023, Limassol, Cyprus, June 26-29, 2023.



## 6 New software, platforms, open data

### 6.1 New software

#### 6.1.1 ACTA

**Name:** A Tool for Argumentative Clinical Trial Analysis

**Keywords:** Artificial intelligence, Natural language processing, Argument mining

**Functional Description:** Argumentative analysis of textual documents of various nature (e.g., persuasive essays, online discussion blogs, scientific articles) allows to detect the main argumentative components (i.e., premises and claims) present in the text and to predict whether these components are connected to each other by argumentative relations (e.g., support and attack), leading to the identification of (possibly complex) argumentative structures. Given the importance of argument-based decision making in medicine, ACTA is a tool for automating the argumentative analysis of clinical trials. The tool is designed to support doctors and clinicians in identifying the document(s) of interest about a certain disease, and in analyzing the main argumentative content and PICO elements.

**URL:** <http://ns.inria.fr/acta/>

**Contact:** Serena Villata

#### 6.1.2 ARViz

**Name:** Association Rules Visualization

**Keyword:** Information visualization

**Scientific Description:** ARViz supports the exploration of data from named entities knowledge graphs based on the joint use of association rule mining and visualization techniques. The former is a widely used data mining method to discover interesting correlations, frequent patterns, associations, or casual structures among transactions in a variety of contexts. An association rule is an implication of the form  $X \rightarrow Y$ , where  $X$  is an antecedent itemset and  $Y$  is a consequent itemset, indicating that transactions containing items in set  $X$  tend to contain items in set  $Y$ . Although the approach helps reduce and focus the exploration of large datasets, analysts are still confronted with the inspection of hundreds of rules in order to grasp valuable knowledge. Moreover, when extracting association rules from named entities (NE) knowledge graphs, the items are NEs that form antecedent  $\rightarrow$  consequent links, which the user should be able to cross to recover information. In this context, information visualization can help analysts visually identify interesting rules that are worthy of further investigation, while providing suitable visual representation to communicate the relationships between itemsets and association rules.

**Functional Description:** ARViz supports the exploration of thematic attributes describing association rules (e.g. confidence, interestingness, and symmetry) through a set of interactive, synchronized, and complementary visualisation techniques (i.e. a chord diagram, an association graph, and a scatter plot). Furthermore, the interface allows the user to recover the scientific publications related to rules of interest.

**Release Contributions:** Visualization of association rules within the scientific literature of COVID-19.

**URL:** <http://covid19.i3s.unice.fr:8080/arviz/>

**Publication:** hal-03292140

**Contact:** Marco Alba Winckler

**Participants:** Aline Menin, Lucie Cadorel, Andrea Tettamanzi, Alain Giboin, Fabien Gandon, Marco Alba Winckler

### 6.1.3 Attune

**Name:** Attune - A Web-Based Digital Audio Workstation to Empower Cochlear Implant Users

**Keywords:** Web Application, Audio signal processing, Plug-in

**Functional Description:** Attune is online software based on the Wam-Studio open-source digital audio workstation, adapted to help cochlear implant users perceive music more clearly. During multi-track listening, simple settings such as "clarity", "power", "attenuation" can be used. In reality, these settings control many parameters of sound processing plugins, which operate behind the scenes. The mapping between these plugins and the settings offered to users is carried out by researchers, using a dedicated graphical interface and a powerful macro management system.

**Contact:** Michel Buffa

**Partner:** CCRMA Lab, Stanford

### 6.1.4 CORESE-Core

**Name:** COnceptual REsource Search Engine - Core

**Keywords:** Semantic Web, RDE, RDFS, SPARQL, OWL, SHACL, Automated Reasoning, Validation, Interoperability, Linked Data, Knowledge Graphs, Knowledge Bases, Knowledge representation, Querying, Ontologies

**Scientific Description:** CORESE-Core is a library used in research to apply and evaluate Semantic Web standards and the algorithms they require. It is also the basis for proposing and prototyping extensions to these standards and their processing.

**Functional Description:** CORESE-Core is a library that implements and extends the Semantic Web standards established by the W3C, such as RDE, RDFS, SPARQL1.1 Query & Update, OWL RL, SHACL, and others.

This library offers a wide range of features for creating, manipulating, parsing, serializing, querying, reasoning and validating RDF data.

In addition, it offers advanced extensions such as STTL, SPARQL Rule and LDScript, which extend the functionality and processing capabilities of the data.

NB: CORESE-Core is a library derived from the earlier CORESE software.

**Release Contributions:** <https://github.com/Wimmics/corese/blob/master/CHANGELOG.md>

**URL:** <http://project.inria.fr/corese>

**Contact:** Remi Ceres

**Participants:** Remi Ceres, Fabien Gandon

### 6.1.5 CORESE-GUI

**Name:** COnceptual REsource Search Engine - Graphical User Interface

**Keywords:** GUI (Graphical User Interface), User Interfaces, Knowledge Bases, Knowledge Graphs, Knowledge graph, Knowledge representation, Ontologies, Linked Data, Validation, Automated Reasoning, SHACL, OWL, SPARQL, RDFS, RDE, Querying, Applications

**Scientific Description:** CORESE-GUI is a graphical user interface developed to interact with the CORESE-Core library. It provides users, especially those less experienced in programming, with an intuitive and visual access to the functionalities of CORESE-Core. This interface includes tools for visualizing semantic data, editing SPARQL queries, and monitoring data processing results. CORESE-GUI also serves as a platform for experimenting with new extensions and processing methods in the field of semantic web, thereby making these technologies more accessible to researchers and practitioners.

**Functional Description:** This desktop application allows the user to call up CORESE-Core features for creating, manipulating, parsing, serializing, querying, reasoning and validating RDF data.

The application enables direct use of Semantic Web languages standardized by the W3C, such as RDF and its syntaxes, RDFS, SPARQL1.1 Query & Update, OWL RL, SHACL, and others.

**URL:** <http://project.inria.fr/corese>

**Contact:** Remi Ceres

**Participants:** Remi Ceres, Fabien Gandon

#### 6.1.6 CORESE-Server

**Name:** COnceptual REsource Search Engine - Server

**Keywords:** Server, Linked Data, Semantic Web, Ontologies, Knowledge Graphs, Knowledge Bases, RDF, RDFS, SPARQL, SHACL, Querying, Validation, Automated Reasoning

**Scientific Description:** This server version allows remote applications to access CORESE-Core functionalities for creating, manipulating, analyzing, serializing, querying, reasoning, and validating RDF data. The server facilitates remote use of W3C-standardized Semantic Web languages, such as RDF and its syntaxes, RDFS, SPARQL1.1 Query & Update, OWL RL, SHACL, and more.

**Functional Description:** This server version enables a remote application to call CORESE-Core's functions for creating, manipulating, analyzing, serializing, querying, reasoning and validating RDF data.

The server enables remote use of Semantic Web languages standardized by the W3C, such as RDF and its syntaxes, RDFS, SPARQL1.1 Query & Update, OWL RL, SHACL, and others.

**URL:** <https://project.inria.fr/corese/>

**Contact:** Remi Ceres

**Participants:** Remi Ceres, Fabien Gandon

#### 6.1.7 CORESE-Command

**Name:** COnceptual REsource Search Engine - Command Line

**Keywords:** Command, RDF, RDFS, SPARQL, SHACL, Knowledge acquisition

**Scientific Description:** This command-line version of CORESE enables users to incorporate CORESE-Core functionalities into scripts, workflows, and consoles for creating, manipulating, analyzing, serializing, querying, reasoning, and validating RDF data. It allows direct use of W3C-standardized Semantic Web languages, such as RDF and its syntaxes, RDFS, SPARQL1.1 Query & Update, OWL RL, SHACL, and more.

**Functional Description:** This command-line version enables users to call CORESE-Core's functionality in scripts, workflows and console mode for the creation, manipulation, analysis, serialization, querying, reasoning and validation of RDF data.

The command enables direct use of W3C-standardized Semantic Web languages, such as RDF and its syntaxes, RDFS, SPARQL1.1 Query & Update, OWL RL, SHACL, and others.

**URL:** <https://project.inria.fr/corese/>

**Contact:** Remi Ceres

**Participants:** Remi Ceres, Fabien Gandon

### 6.1.8 CREEP semantic technology

**Keywords:** Natural language processing, Machine learning, Artificial intelligence

**Scientific Description:** The software provides a modular architecture specifically tailored at the classification of cyberbullying and offensive content on social media platforms. The system can use a variety of features (ngrams, different word embeddings, etc) and all the network parameters (number of hidden layers, dropout, etc) can be altered by using a configuration file.

**Functional Description:** The software uses machine learning techniques to classify cyberbullying instances in social media interactions.

**Release Contributions:** Attention mechanism, Hyperparameters for emoji in config file, Predictions output, Streamlined labeling of arbitrary files

**Publications:** [hal-01906096v1](#), [hal-01920266v1](#)

**Contact:** Michele Corazza

**Participants:** Michele Corazza, Elena Cabrio, Serena Villata

### 6.1.9 CROBORA

**Name:** Crossing borders Archives. The circulation of images of Europe.

**Keywords:** Audiovisual, Data visualization

**Functional Description:** This platform gives access to 36 000 stock shots reused in the evening news of six national channels in France and Italy (France 2, Arte, TF1, Rai Uno, Rai Due and Canale 5) and the YouTube accounts of the European institutions between 2001 and 2021. The platform gives access four types of data: screenshots (one for each stock shot), metadata (for each stock shot), videos (news), and original metadata for each video. The platform integrates three visualization tools (Treemaps, Muviz, and ArViz) presenting patterns and relationships between records. The tool is available at <https://crobora.huma-num.fr/crobora?tab=0>

**Contact:** Marco Alba Winckler

### 6.1.10 DBpedia

**Name:** DBpedia

**Keywords:** RDF, SPARQL

**Functional Description:** DBpedia is an international crowd-sourced community effort to extract structured information from Wikipedia and make this information available on the semantic Web as linked open data. The DBpedia triple stores then allow anyone to solve sophisticated queries against Wikipedia extracted data, and to link the different data sets on these data. The French chapter of DBpedia was created and deployed by Wimmics and is now an online running platform providing data to several projects such as: QAKIS, Izipedia, zone47, Sépage, HdA Lab., JocondeLab, etc.

**Release Contributions:** The new release is based on updated Wikipedia dumps and the inclusion of the DBpedia history extraction of the pages.

**URL:** <http://wiki.dbpedia.org/>

**Contact:** Fabien Gandon

**Participants:** Fabien Gandon, Elmahdi Korfed

### 6.1.11 Fuzzy labelling argumentation module

**Name:** Fuzzy labelling algorithm for abstract argumentation

**Keywords:** Artificial intelligence, Multi-agent, Knowledge representation, Algorithm

**Functional Description:** The goal of the algorithm is to compute the fuzzy acceptability degree of a set of arguments in an abstract argumentation framework. The acceptability degree is computed from the trustworthiness associated with the sources of the arguments.

**Contact:** Serena Villata

**Participant:** Serena Villata

### 6.1.12 GUsT-3D

**Name:** Guided User Tasks Unity plugin for 3D virtual reality environments

**Keywords:** 3D, Virtual reality, Interactive Scenarios, Ontologies, User study

**Functional Description:** We present the GUsT-3D framework for designing Guided User Tasks in embodied VR experiences, i.e., tasks that require the user to carry out a series of interactions guided by the constraints of the 3D scene. GUsT-3D is implemented as a set of tools that support a 4-step workflow to :

(1) annotate entities in the scene with names, navigation, and interaction possibilities, (2) define user tasks with interactive and timing constraints, (3) manage scene changes, task progress, and user behavior logging in real-time, and (4) conduct post-scenario analysis through spatio-temporal queries on user logs, and visualizing scene entity relations through a scene graph.

The software also includes a set of tools for processing gaze tracking data, including: - clean and synchronise the data - calculate fixations with I-VT, I-DT, IDTVR, IS5T, Remodnav, and IDVT algorithms - visualize the data (points of regard and fixations) in both real time and collectively

**URL:** <https://project.inria.fr/creattive3d/gust-3d/>

**Contact:** Hui-Yin Wu

**Participants:** Hui-Yin Wu, Marco Alba Winckler, Lucile Sassatelli, Florent Robert

**Partner:** I3S

### 6.1.13 IndeGx

**Keywords:** Semantic Web, Indexation, Metadata

**Functional Description:** IndeGx is a framework for the creation of an index of a set of SPARQL endpoints. The framework relies only on available semantic web technologies and the index appears as an RDF database. The index is primarily composed of the self-description available in the endpoint. This original description is verified and expanded by the framework, using SPARQL queries.

**URL:** <https://github.com/Wimmics/dekalog>

**Contact:** Pierre Maillot

**Participants:** Fabien Gandon, Catherine Faron, Olivier Corby, Franck Michel

#### 6.1.14 ISSA-pipeline

**Name:** Processing pipeline of the ISSA project

**Keywords:** Indexation, Semantic Web, Natural language processing, Knowledge graph, Open Access, Open data, LOD - Linked open data

**Functional Description:** See the description at <https://github.com/issa-project/issa-pipeline/tree/main/pipeline>

**Release Contributions:** <https://github.com/issa-project/issa-pipeline/releases/tag/2.0>

**URL:** <https://github.com/issa-project/issa-pipeline>

**Contact:** Franck Michel

**Partners:** CIRAD, IMT Mines Alès

#### 6.1.15 ISSA Visualization Web Application

**Keywords:** Open Access, Data visualization, Knowledge graph, NLP

**Functional Description:** The ISSA project focuses on the semantic indexing of scientific publications in an open archive. The ISSA Visualization Web Application is a React and node.js based web application meant to search articles from the ISSA knowledge base using the rich semantics of the reference vocabularies, and provide a visualization of their metadata. This application consists of a frontend and a backend hosted on separate repositories: <https://github.com/issa-project/web-visualization/> <https://github.com/issa-project/web-backend>

**URL:** <https://github.com/issa-project/web-visualization/>

**Contact:** Franck Michel

**Partners:** CIRAD, IMT Mines Alès

#### 6.1.16 KartoGraphi

**Keywords:** Semantic Web, LOD - Linked open data

**Functional Description:** Website displaying a screenshot of the state of the Linked Data web according to the description retrieved by the IndeGx software

**URL:** <http://prod-dekalog.inria.fr/>

**Publication:** hal-03652865

**Contact:** Pierre Maillot

#### 6.1.17 Licentia

**Keywords:** Right, License

**Scientific Description:** In order to ensure the high quality of the data published on the Web of Data, part of the self-description of the data should consist in the licensing terms which specify the admitted use and re-use of the data by third parties. This issue is relevant both for data publication as underlined in the “Linked Data Cookbook” where it is required to specify an appropriate license for the data, and for the open data publication as expressing the constraints on the reuse of the data would encourage the publication of more open data. The main problem is that data producers and publishers often do not have extensive knowledge about the existing licenses, and the legal terminology used to express the terms of data use and reuse. To address this open issue, we present Licentia, a suite of services to support data producers and publishers in data licensing by means of a user-friendly interface that masks to the user the complexity of the legal reasoning process.

In particular, Licentia offers two services: i) the user selects among a pre-defined list those terms of use and reuse (i.e., permissions, prohibitions, and obligations) she would assign to the data and the system returns the set of licenses meeting (some of) the selected requirements together with the machine readable licenses' specifications, and ii) the user selects a license and she can verify whether a certain action is allowed on the data released under such license. Licentia relies on the dataset of machine-readable licenses (RDF, Turtle syntax, ODRL vocabulary and Creative Commons vocabulary) available at <http://datahub.io/dataset/rdflicense>. We rely on the deontic logic presented by Governatori et al. to address the problem of verifying the compatibility of the licensing terms in order to find the license compatible with the constraints selected by the user. The need for licensing compatibility checking is high, as shown by other similar services (e.g., Licensius5 or Creative Commons Choose service6 ). However, the advantage of Licentia with respect to these services is twofold: first, in these services compatibility is pre-calculated among a pre-defined and small set of licenses, while in Licentia compatibility is computed at runtime and we consider more than 50 heterogeneous licenses, second, Licentia provides a further service that is not considered by the others, i.e., it allows to select a license from our dataset and verify whether some selected actions are compatible with such license.

**Functional Description:** Licentia is a web service application with the aim to support users in licensing data. Our goal is to provide a full suite of services to help in the process of choosing the most suitable license depending on the data to be licensed.

The core technology used in our services is powered by the SPINdle Reasoner and the use of Defeasible Deontic Logic to reason over the licenses and conditions.

The dataset of RDF licenses we use in Licentia is the RDF licenses dataset where the Creative Commons Vocabulary and Open Digital Rights Language (ODRL) Ontology are used to express the licenses.

**URL:** <http://licentia.inria.fr/>

**Contact:** Serena Villata

**Participant:** Cristian Cardellino

#### 6.1.18 SPARQL Micro-services

**Name:** SPARQL micro-services

**Keywords:** Web API, SPARQL, Microservices, LOD - Linked open data, Data integration

**Functional Description:** The approach leverages the micro-service architectural principles to define the SPARQL Micro-Service architecture, aimed at querying Web APIs using SPARQL. A SPARQL micro-service is a lightweight SPARQL endpoint that typically provides access to a small, resource-centric graph. Furthermore, this architecture can be used to dynamically assign dereferenceable URIs to Web API resources that do not have URIs beforehand, thus literally “bringing” Web APIs into the Web of Data. The implementation supports a large scope of JSON-based Web APIs, may they be RESTful or not.

**URL:** <https://github.com/frmichel/sparql-micro-service>

**Publications:** [hal-02060966](#), [hal-01722792](#), [hal-01947589](#), [hal-02168164](#)

**Author:** Franck Michel

**Contact:** Franck Michel

### 6.1.19 Metadatamatic

**Keywords:** RDF, Semantic Web, Metadata

**Functional Description:** Website offering a form to generate in RDF the description of an RDF base.

**URL:** <https://wimmics.github.io/voidmatic/>

**Contact:** Pierre Maillot

**Participants:** Fabien Gandon, Franck Michel, Olivier Corby, Catherine Faron

### 6.1.20 MGExplorer

**Name:** Multivariate Graph Explorer

**Keyword:** Information visualization

**Scientific Description:** MGExplorer (Multidimensional Graph Explorer) allows users to explore different perspectives to a dataset by modifying the input graph topology, choosing visualization techniques, arranging the visualization space in meaningful ways to the ongoing analysis and retracing their analytical actions. The tool combines multiple visualization techniques and visual querying while representing provenance information as segments connecting views, which each supports selection operations that help define subsets of the current dataset to be explored by a different view. The adopted exploratory process is based on the concept of chained views to support the incremental exploration of large, multidimensional datasets. Our goal is to provide visual representation of provenance information to enable users to retrace their analytical actions and to discover alternative exploratory paths without losing information on previous analyses.

**Functional Description:** MGExplorer is an information visualization tool suite that integrates many information visualization techniques aimed at supporting the exploration of multivariate graphs. MGExplorer allows users to choose and combine the information visualization techniques creating a graph that describes the exploratory path of dataset. It is an application based on the D3.JS library, which is executable in a web browser. The use of MGExplorer requires a customization to connect the dashboard to a SPARQL endpoint. MGExplorer has been customized to facilitate the search of scientific articles related to covid.

**Release Contributions:** Visualization of data extracted from linked data datasets.

**URL:** <http://covid19.i3s.unice.fr:8080/>

**Publications:** [hal-03292172](#), [hal-03404572](#), [hal-03404580](#)

**Contact:** Marco Alba Winckler

**Participants:** Aline Menin, Marco Alba Winckler, Olivier Corby

**Partner:** Universidade Federal do Rio Grande do Sul

### 6.1.21 Morph-xR2RML

**Name:** Morph-xR2RML

**Keywords:** RDF, Semantic Web, LOD - Linked open data, MongoDB, SPARQL

**Functional Description:** The xR2RML mapping language that enables the description of mappings from relational or non relational databases to RDF. It is an extension of R2RML and RML.

Morph-xR2RML is an implementation of the xR2RML mapping language, targeted to translate data from the MongoDB database, as well as relational databases (MySQL, PostgreSQL, MonetDB). Two running modes are available: (1) the graph materialization mode creates all possible RDF triples at



once, (2) the query rewriting mode translates a SPARQL 1.0 query into a target database query and returns a SPARQL answer. It can run as a SPARQL endpoint or as a stand-alone application.

Morph-xR2RML was developed by the I3S laboratory as an extension of the Morph-RDB project which is an implementation of R2RML.

**URL:** <https://github.com/frmichel/morph-xr2rml/>

**Publications:** [hal-01207828](#), [hal-01330146](#), [hal-01280951](#)

**Author:** Franck Michel

**Contact:** Franck Michel

### 6.1.22 Muvin

**Name:** Multimodal Visualization of Networks

**Keywords:** Data visualization, Music, LOD - Linked open data

**Functional Description:** Muvin supports the exploration of a two-layer network describing the collaborations between artists and the discography of an artist, defined by the albums and songs released by the artist over time. It implements an incremental approach, allowing the user to dynamically import data from a SPARQL endpoint to the exploration flow. Furthermore, this approach seeks to improve user perception by associating audio to the visualization, in a way that the users can listen to the songs visually represented in their screen.

**URL:** <http://dataviz.i3s.unice.fr/muvin/>

**Contact:** Aline Menin

### 6.1.23 wam-studio

**Keywords:** Web Application, Web API, Audio signal processing

**Functional Description:** WAM Studio is an open source online Digital Audio Workstation (DAW) that takes advantage of a number of APIs and standard W3C technologies, such as Web Audio, WebAssembly, Web Components, Web Midi, Media Devices and more. WAM Studio is also based on the Web Audio Modules (WAM) standard, which was designed to facilitate the development of interoperable audio plug-ins (effects, virtual instruments, virtual piano keyboards as controllers, etc.) - a kind of "VSTs for the Web". DAWs are feature-rich software programs, and therefore particularly complex to develop in terms of design, implementation, performance and ergonomics. Today, the majority of online DAWs are commercial, while the only open source examples lack functionality (no plug-in support, for example) and don't take advantage of the recent possibilities of web browsers (such as WebAssembly). WAM Studio was designed as a technology demonstrator to promote the possibilities offered by recent innovations proposed by the W3C. Developing it was a challenge, as we had to take into account the limitations of sandboxed and constrained environments such as Web browsers, and compensate for latency when we can't know what hardware is being used, etc.). An online demo and a GitHub repository for the source code are available (<https://wam-studio.i3s.univ-cotedazur.fr/>).

**URL:** <https://wam-bank.i3s.univ-cotedazur.fr/>

**Contact:** Michel Buffa

### 6.1.24 WebAudio tube guitar amp sims CLEAN, DISTO and METAL MACHINES

**Name:** Tube guitar amplifier simulators for Web Browser : CLEAN MACHINE, DISTO MACHINE and METAL MACHINE

**Keyword:** Tube guitar amplifier simulator for web browser

**Scientific Description:** This software is one of the only ones of its kind to work in a web browser. It uses "white box" simulation techniques combined with perceptual approximation methods to provide a quality of guitar playing in hand comparable to the best existing software in the native world.

**Functional Description:** Software programs for creating real-time simulations of tube guitar amplifiers that behave most faithfully like real hardware amplifiers, and run in a web browser. In addition, the generated simulations can run within web-based digital audio workstations as plug-ins. The "CLEAN MACHINE" version specializes in the simulation of acoustic guitars when playing electric guitars. The DISTO machine specializes in classic rock tube amp simulations, and METAL MACHINE targets metal amp simulations. These programs are one of the results of the ANR WASABI project.

**Release Contributions:** First stable version, delivered and integrated into the ampedstudio.com software. Two versions have been delivered: a limited free version and a commercial one.

**Publications:** [hal-01721463](#), [hal-01893681](#), [hal-02337828](#), [hal-03087768](#), [hal-01721483](#), [hal-01735478](#), [hal-02366725](#), [hal-02557901](#), [hal-01589330](#), [hal-03087763](#), [hal-01893660](#), [hal-01589229](#)

**Contact:** Michel Buffa

**Participant:** Michel Buffa

**Partner:** Amp Track Ltd, Finland

## 6.2 Open data

### 6.2.1 CyberAgressionAdo

**Participants:** Anaïs Ollagnier, Serena Villata, Elena Cabrio.

**Name:** CyberAgressionAdo-v1

**Description:** The CyberAgressionAdo-v1 dataset comprises 19 instances of aggressive multiparty chats in French, gathered through a role-playing game conducted in high schools. This dataset is built upon scenarios that emulate cyber aggression situations prevalent among teenagers, addressing sensitive topics like ethnic origin, religion, or skin color. The recorded conversations have undergone meticulous annotation, taking into account various facets, including participant roles, the occurrence of hate speech, the nature of verbal abuse within the messages, and the identification of humor devices like sarcasm or irony in utterances.

The guidelines and data are provided on a [dedicated repository](#).

**Release Contributions:** CyberAgressionAdo-V2 uses a multi-label, fine-grained tagset marking the discursive role of exchanged messages as well as the context in which they occur – for instance, attack (ATK), defend (DFN), counterspeech (CNS), abet/instigate (AIN), gaslight (GSL), etc.

The guidelines and data are provided on a [dedicated repository](#).

**Publications:** [113]

**Contact:** Anaïs Ollagnier

### 6.2.2 ElecDeb60to20

**Participants:** Serena Villata, Elena Cabrio, Pierpaolo Goffredo.

**Name:** ElecDeb60to20

**Description:** The ElecDeb60to20 dataset is built from the official transcripts of the televised presidential debates in the US from 1960 until 2020, from the website of the Commission on Presidential Debates (CPD). These political debates are manually annotated with argumentative components (claim, premise) and relations (support, attack). In addition, it also includes the annotation of fallacious arguments, based on the following 6 classes of fallacies: ad hominem, appeal to authority, appeal to emotion, false cause, slogan, slippery slope.

The guidelines and data are provided on a [dedicated repository](#).

**Publications:** [62], [63]

**Contact:** Serena Villata

### 6.2.3 ISSA Agritrop Dataset

**Participants:** Anna Bobasheva, Catherine Faron, Franck Michel.

**Name:** ISSA Agritrop Dataset

**Description:** The ISSA Agritrop Dataset was produced in the context of the [ISSA 2 project](#). It provides a semantic index of the articles of the [Agritrop scientific archive](#). It is built by extracting thematic descriptors and named entities from the articles' text and linking them with resources from DBpedia, Wikidata, the AGROVOC thesaurus and GeoNames.

The data model and generation pipeline are provided on a [dedicated repository](#) (DOI: 10.5281/zenodo.10376913).

**Release Contributions:** version 2.0 extends the dataset with the RDF representation of additional data, and the result of processing 13.000+ full-text published articles, as well as 10.000+ abstracts from other documents that were not processed before.

**DOI:** [10.5281/zenodo.10381606](#)

**Publications:** [44], [114]

**Contact:** Franck Michel

### 6.2.4 WeKG-MF

**Participants:** Nadia Yacoubi Ayadi, Catherine Faron, Franck Michel.

**Name:** Weather Knowledge Graph of Météo France Meteorological Observations

**Description:** WeKG-MF represents the meteorological observations made by 62 Météo-France weather stations located in different regions in metropolitan France and overseas departments, from 2012 to 2022. This work is supported by the French National Research Agency under grant ANR-18-CE23-0017 (project [D2KAB](#)). The raw data was obtained from Météo France's website. The data model and generation pipeline are provided on a [dedicated repository](#).

**Release Contributions:**

**DOI:** [10.5281/zenodo.10420888](#)

**Publications:** [83], [115], [116]

**Contact:** Franck Michel

### 6.2.5 WheatGenomicsSLKG

**Participants:** Nadia Yacoubi Ayadi, Catherine Faron, Franck Michel.

**Name:** Wheat Genomics Scientific Literature Knowledge Graph

**Description:** The Wheat Genomics Scientific Literature Knowledge Graph is a FAIR knowledge graph that exploits the Semantic Web technologies to integrate information about Named Entities (NE) extracted automatically from a corpus of PubMed scientific papers on wheat genetics and genomics. This work is supported by the French National Research Agency under grant ANR-18-CE23-0017 (project [D2KAB](#)). The data model and generation pipeline are provided on a [dedicated repository](#).

**Release Contributions:**

**DOI:** [10.5281/zenodo.10420888](https://doi.org/10.5281/zenodo.10420888)

**Publications:**

**Contact:** Franck Michel

### 6.2.6 Pharmacogenomics datasets for Ontology Matching

**Name:** Pharmacogenomics datasets for Ontology Matching

**Participants:** Pierre Monnin.

**Description:** These datasets constitute benchmarks to evaluate Ontology Matching algorithms on a complex structure-based instance matching task from the domain of pharmacogenomics. Pharmacogenomics involves  $n$ -ary tuples representing so-called “pharmacogenomic relationships” and their components of three distinct types: drugs, genetic factors, and phenotypes. The goal resides in matching such tuples. These datasets were extracted from the [PGxLOD](#) knowledge graph.

**Release Contributions:**

**DOI:** [10.5281/zenodo.8419361](https://doi.org/10.5281/zenodo.8419361)

**Publications:**

**Contact:** Pierre Monnin

### 6.2.7 DBpedia.fr : French chapter of the DBpedia knowledge graph dataset

**Name:** DBpedia.fr

**Participants:** Fabien Gandon, Franck Michel, Célian Ringwald.

The [DBpedia.fr](#) project ensures the creation and maintenance of a French chapter of the DBpedia knowledge base a crowd-sourced community effort to extract structured content from the information created in various Wikimedia projects.

No new release was done this year but we carried out continuous monitoring and support to ensure a high-availability service.

Statistics indicate very high usage rate: the server processed 1.8+ billion queries over the year. This represents a 3.86 million daily average and 32.5 million daily max.

**URL:** <http://dbpedia.fr>

**Contact:** Célian Ringwald, Franck Michel, Fabien Gandon

## 7 New results

### 7.1 User Modeling and Designing Interaction

#### 7.1.1 Spatio-Temporal Data Visualization

**Participants:** Aline Menin, Marco Winckler, Franck Michel.

Knowledge graph visualization is a challenge on itself specially due the presence of multidimensional data: as the number data dimensions increases, a single view displaying all the information at once is no longer suitable due to cognitive overload and visual cluttering [112]. A frequent type of multidimensional data in many datasets concerns spatio-temporal data. When designing visualizations to represent spatio-temporal data, one must respect the order of time units and the geographical location of data records, which hinders the design process. In this context, we investigate the use of glyph-based data representations which might be suitable to convey multiple attributes in a small and compact visual representation. We have chosen glyphs because (i) they can be used independently and constructively to depict attributes of a data record or the composition of a set of data records; (ii) they can be spatially connected to convey topological relationships between data records; and (iii) they are visual signs that can make use of visual features of other visual signs such as icons, indices and symbols [108]. Whilst most of visualization techniques use static representations of glyphs, we propose the use of polymorphic glyphs (i.e. glyphs that change its shape according to the surrounding semantics and/or user actions) to assist the exploration of hierarchical temporal data [70]. As a proof-of-concept, we implemented our approach in the form of a web-based visualization tool, called PHiTGlyph (Polymorphic and Hierarchical Temporal Glyph), that uses interaction to activate the polymorphous aspect of glyphs. We have applied the approach to explore hierarchical data over space and time. In particular, we demonstrate the suitability of our approach by exploring the ISSA KG, which describes scholarly articles in the field of agriculture. The tool is available [here](#).

#### 7.1.2 Tracking and Exploring Audiovisual Archives about the European Union

**Participants:** Aline Menin, Marco Winckler, Shiming Shen, Matteo Treleani, Jean-Marie Dormoy.

As part of the project [Crossing Borders Archives \(CROBORA\)](#) we investigate the use of visual queries and visualization techniques as a mean to explore large collections of audiovisual documents within television and web archives. One of the main goals of the project CROBORA is to identify the use of sequences of images/videos that are repeatedly used to illustrate a discourse within a corpus of thousands of videos). In fact, any audiovisual sequence reused within different contexts exists conceptually as the repetition of one single visual unit, but also from the point of view of the metadata tagging its occurrences, each item is a distinct document. Our contribution is mainly focused on the support to the retrieval and analysis of stock shots collected on TV news broadcasted from the period 2001 - 2020 in major national channels in France and in Italy (TF1, France 2, and ARTE on the French side, and Rai 1, Rai 2, and Canale 5 on the Italian one). The resulting dataset contains over 27,000 stock shots that were manually annotated to describe their usage, i.e. either representing an event, a celebrity, a place or used as illustration [42]. The volume and multidimensionality of the corpus makes it impossible for a human to explore it without technological help. We thus investigate the usage of visualization techniques to support the domain experts on analyzing the data and finding the relationship between stock shots and their usage throughout time. We have developed a tool that supports stock shots search and visualization through multiple perspectives: hierarchically, temporal distribution and association through their co-occurrence in the same television archive [43]. The tool is available [here](#).

#### 7.1.3 Interaction with extended reality

**Participants:** Aline Menin, Clément Quéré, Florent Robert, Hui-Yin Wu, Marco Winckler.

Virtual reality (VR) and Augmented Reality (AR) offer extraordinary opportunities to create immersive or mixed-reality experiences allowing users to explore relations (patterns, trends, clusters) in information distributed in a 3D space. The feeling of immersion created by such a technology is expected to foster the user cognition thus helping to make better decisions that require an analysis of spatial information. A major challenge of designing these 3D experiences and user tasks, however, lies in bridging the inter-relational gaps of perception between the designer, the user, the information communicated, and the 3D scene. In the context of the PhD thesis of Florent Robert we have started a series of studies aiming to understand how the design of user of 3D scenes affects the user perception and how such perception affects decision-making processes [79]. Our ultimate goal is to understand how the many components of the user interaction (including user attention) might affect the embodied experience virtual immersive environments. Moreover, we want to explore how metadata information (attached to virtual objects) might help users to navigate in the immersive environment and how such information affects users' decisions. We have proposed a tool, called GUsT-3D framework, for designing Guided User Tasks in embodied VR experiences, i.e., tasks that require the user to carry out a series of interactions guided by the constraints of the 3D scene. Whilst in the PhD thesis of Florent Robert, we mainly explore 3D scenes that are a recreation of parts of the real-world, in the PhD thesis of Clément Quéré, we use VR/AR technologies to explore visualization of more abstract volumes of spatio-temporal datasets. In cooperation with the **IMREDD** we are developing an immersive environment that allows to visualize the multidimensional sensors' data (ex. traffic, weather forecast, pollution...) that describe the evolution of smart cities' states. Another ongoing work concerns the study of augmented reality tools that might be used to annotate both virtual and real-life objects, thus increasing the semantics of objects annotated and reducing the gap between digital and real-life objects.

#### 7.1.4 Incremental visual exploration of linked data

**Participants:** Marco Winckler, Aline Menin, Olivier Corby, Catherine Faron, Alain Giboin.

Information visualization techniques are useful to discover patterns and causal relationships within LOD datasets. However, since the discovery process is often exploratory (i.e., users have no predefined goal and do not expect a particular outcome), when users find something interesting, they should be able to (i) retrace their exploratory path to explain how results have been found, and (ii) branch out the exploratory path to compare data observed in different views or found in different datasets. Furthermore, as most of LOD datasets are very specialized, users often need to explore multiple datasets to obtain the knowledge required to support decision-making processes. Thus, the design of visualization tools is confronted with two main challenges: the visualization system should provide multiple views to enable the exploration of different or complementary perspectives to the data; and the system should support the combination of diverse data sources during the exploration process. To our knowledge, the existing tools before our work, are limited to visualizing a single dataset at a time and, often, use static and preprocessed data. Thus, we proposed the concept of follow-up queries to allow users to create queries on demand during the exploratory process while connecting multiple LOD datasets with chained views. Our approach relies on an exploration process supported by the use of predefined SPARQL queries that the user can select on-the-fly to retrieve data from different SPARQL endpoints. It enables users to enrich the ongoing analysis by bringing external and complementary data to the exploration process, while also supporting the visual analysis and comparison of different subsets of data (from the same or different SPARQL endpoints) and, thus, the incremental exploration of the LOD cloud. The resulting publication [92] presents a generic visualization approach to assist the analysis of multiple LOD datasets based on the concepts of chained views and follow-up queries. We demonstrate the feasibility of our approach via four use case scenarios and a formative evaluation where we explore scholarly data described by RDF graphs publicly available through SPARQL endpoints. These scenarios demonstrate how the tool supports

(i) composing, running, and visualizing the results of a query; (ii) subsetting the data and exploring it via different visualization techniques; (iii) instantiating a follow-up query to retrieve external data; and (iv) querying a different database and compare datasets. The usability and usefulness of the proposed approach is confirmed by results obtained with a series of semi-structured interviews. The results are encouraging while showing the relevance of the approach to explore big linked data. This work resulted in a visualization tool, called LDViz, publicly accessible at [dataviz.i3s.unice.fr/ldviz](http://dataviz.i3s.unice.fr/ldviz). The source code is also open and published as [10.5281/zenodo.6511782](https://doi.org/10.5281/zenodo.6511782). A large study concerning the scalability of the tools has demonstrated over 420 public end points [92].

### 7.1.5 Development of an extensible User Profile model

**Participants:** Ankica Barišić, Marco Winckler.

During the post-doc of Ankica Barišić (from January to December 2023), we have investigated the evolving landscape of user modeling in response to technological advances and increasing complexities in interaction systems. The aim was to develop an extensible User Profile model that could be seamlessly used for the development of complex interactive systems. The objectives include acquiring knowledge of various user modeling techniques, distinguishing between short-term and long-term user models, and developing adaptive models across different phases of system development and operation. We investigated the use of ontologies as a mean to ensure interoperability and facilitate data federation in representing user profiles. We systematically analyzed existing ontologies for user profiles, such as GUMO, UPO, OntobUMf, HPO, PO, Grapple ontology, and User Modeling Meta-Ontology. In order to overcome the lack of flexibility to extend the attribute of existing user profile ontologies, we have developed the User Profile Meta-Ontology (UPMO) which is rooted in Knowledge Graphs (RDF) and MDE principles [90]. The UPMO played a pivotal role in facilitating the sharing and reuse of domain knowledge while streamlining the analysis and extraction of domain-specific information. We also have investigated methods for obtaining user profiles, encompassing qualitative and quantitative research methods; we have focused on three prevalent approaches: role-based, persona development, and user profiling. The UPMO was showcased as a foundational element guiding developers through various phases of system development, promoting consistency, interoperability, and efficiency. Our approach was applied in automobile domain, leading to a scientific contribution in terms of a driver model tailored for Take-Over-Request (TOR) scenarios in autonomous vehicles [54].

### 7.1.6 Interactive WebAudio applications

**Participants:** Michel Buffa, Shihong Ren.

During the WASABI ANR research project (2017-2020), we built a 2M song database made of metadata collected from the Web of Data and from the analysis of song lyrics of the audio files provided by Deezer. This dataset is still exploited by current projects inside the team (in particular by the PhD of Maroua Tickat). Other initiatives closely related to the WASABI datasets include several Web Audio interactive applications and frameworks. The Web Audio Modules 2.0, a WebAudio plugin standard for developing high performance plugins in the browser published for the first time in 2021, has encountered a large adoption by researchers and developers [57, 105]. The open source Wam-Studio Digital Audio Workstation developed by Michel Buffa and Antoine Vidal-Mazuy led to a scientific collaboration with the MERI team from the CCRMA Laboratory in Stanford [88, 56, 55]. The team is also participating to the ANR DOTS project [47], that aims to produce distributed music performances using Web Audio based infrastructure.

We also developed new methods for real-time tube guitar amplifier simulations that run in the browser [33, 91]. Some of these results are still unique in the world as in 2023, and have been acclaimed by several awards in international conferences. The guitar amp simulations are now commercialized by the CNRS SATT service and are available in the online collaborative Digital Audio Workstation *ampedstudio*. Some

other tools we designed are linked to the WASABI knowledge base, that allow, for example, songs to be played along with sounds similar to those used by artists. An ongoing PhD proposes a visual language for music composers to create instruments and effects linked to the WASABI corpus content and a research collaboration with the Shanghai conservatory of Music is also being pursued about the generation of music on the Web with real-time Brain Wave [78].

### 7.1.7 KartoGraphI: Drawing a Map of Linked Data

**Participants:** Pierre Maillot, Olivier Corby, Catherine Faron, Fabien Gandon, Franck Michel.

A large number of semantic Web knowledge bases have been developed and published on the Web. To help the user identify the knowledge bases relevant for a given problem, and estimate their usability, we propose a declarative indexing framework and an associated visualization Web application, KartoGraphI. It provides an overview of important characteristics for more than 400 knowledge bases including, for instance, dataset location, SPARQL compatibility level, shared vocabularies, etc. [38]

## 7.2 Communities and Social Interactions Analysis

### 7.2.1 Autonomous agents in a social and ubiquitous Web

**Participants:** Andrei Ciortea, Olivier Corby, Fabien Gandon, Franck Michel.

Recent W3C recommendations for the Web of Things (WoT) and the Social Web are turning hypermedia into a homogeneous information fabric that interconnects heterogeneous resources: devices, people, information resources, abstract concepts, etc. The integration of multi-agent systems with such hypermedia environments now provides a means to distribute autonomous behavior in worldwide pervasive systems.

A central problem then is to enable autonomous agents to discover heterogeneous resources in world wide and dynamic hypermedia environments. This is a problem in particular in WoT environments that rely on open standards and evolve rapidly—thus requiring agents to adapt their behavior at runtime in pursuit of their design objectives. To this end, we developed a hypermedia search engine for the WoT that allows autonomous agents to perform approximate search queries in order to retrieve relevant resources in their environment in (weak) real time. The search engine crawls dynamic WoT environments to discover and index device metadata described with the W3C WoT Thing Description, and exposes a SPARQL endpoint that agents can use for approximate search. To demonstrate the feasibility of our approach, we implemented a prototype application for the maintenance of industrial robots in worldwide manufacturing systems. The prototype demonstrates that our semantic hypermedia search engine enhances the flexibility and agility of autonomous agents in a social and ubiquitous Web [109].

More generally, we defend that Knowledge Graphs (KGs) are one of the prime approaches to support the programming of autonomous software systems at the knowledge level [34]. We are working towards defining a new class of Web-based Multi-Agent Systems (MAS) that can inherit the architectural properties of the Web (scalability, heterogeneity, evolvability, etc.), preserve the architectural properties of MAS (adaptability, openness, robustness, etc.), and are human-centric (usable, transparent, accountable, etc.) [46]. We study the idea of decentralized hypermedia applications in general [45] and consider the abstractions needed for designing multi-agent systems on the Web [48], the challenges in achieving Linked Multi-Agent Systems (MAS) on the Web [49] and specific problems such as the one of situatedness and embodiment of agents acting on the Web (namely, "Web agents") [84].

### 7.2.2 Online Hate Detection in Conversational Data



**Participants:** Elena Cabrio, Serena Villata, Anais Ollagnier.

Online Hate Detection (OHD) encompasses various sub-tasks, including the classification of hateful content (e.g., distinguishing hate and non-hate speech, identifying hate targets, or detecting types of hate). More recently, the focus has extended to the task of Participant Role Identification (PRI). In PRI, the objective is to identify different participant roles involved in cyberbullying episodes, such as bully, victim, bystander, assistant, defender, among others. While existing research has predominantly targeted social networks like Twitter and Instagram, the significant role of private instant messaging platforms in facilitating bullying, particularly among teens, has gained attention.

Despite the prevalence of bullying on private messaging platforms, few studies have delved into the task of PRI in multi-party chats due to the privacy policies of these platforms. Leveraging the CyberAgressionAdo-V1 dataset [113], collected under the UCA IDEX OTESIA project "Artificial Intelligence to prevent cyberviolence, cyberbullying, and hate speech online"<sup>1</sup>, we aim to address this gap. Our contribution involves introducing a pipeline designed to automate the identification of participant roles in cyberbullying episodes occurring within multi-party chats [75]. Additionally, we have developed a fully web-based platform for participant role detection in multi-party chats, featuring a novel role scoring function [76].

A conference paper presenting the CyberAgressionAdo-V2 dataset is currently under submission. This enhanced version of CyberAgressionAdo-V1 introduces a new tagset that incorporates markers for pragmatic-level information observed in cyberbullying situations. The refined tagset comprises six layers, with two specifically focusing on pragmatic considerations, capturing both intentions and context. These layers aim to grasp the authors' intentions and establish the context in which messages serve as responses. This paper delivers a three-fold contribution: (1) it advances the development of solutions addressing diversity in digital harassment, (2) it capitalizes on annotators' disagreements in tagset development, and (3) it delves into pattern mining to enhance the modeling of intricate communication patterns prevalent in cyberbullying.

### 7.2.3 Hate Speech Target Community Detection

**Participants:** Elena Cabrio, Serena Villata, Anais Ollagnier.

Pursuing the ongoing endeavor in reaching a fine-grained online hate speech characterization to provide appropriate solutions to curb online abusive behaviors, we proposed a full pipeline that enables to capture targeting characteristics in hatred contents (i.e., types of hate, such as race and religion) aiming at improving the understanding on how hate is conveyed on Twitter [39]. Our contribution is threefold: (1) we leverage multiple data views of a different nature to contrast different kinds of abusive behaviors expressed towards targets; (2) we develop a full pipeline relying on a multi-view clustering technique to address the task of hate speech target characterization; and (3) we propose a methodology to assess the quality of generated hate speech target communities. Relying on multiple data views built from multilingual pre-trained language models (i.e., multilingual BERT and multilingual Universal Sentence Encoder) and the Multi-view Spectral Clustering (MvSC) algorithm, the experiments conducted on a freely available multilingual dataset of tweets (i.e., the MLMA hate speech dataset) show that most of the configurations of the proposed pipeline significantly outperform state-of-the-art clustering algorithms on all the tested clustering quality metrics on both French and English.

### 7.2.4 An In-depth Analysis of Implicit and Subtle Hate Speech Messages

**Participants:** Nicolás Benjamín Ocampo, Ekaterina Sviridova, Elena Cabrio, Serena Villata.

---

<sup>1</sup>OTESIA Project Launch

The research carried out so far in detecting abusive content in social media has primarily focused on overt forms of hate speech. While explicit hate speech (HS) is more easily identifiable by recognizing hateful words, messages containing linguistically subtle and implicit forms of HS (as circumlocution, metaphors and sarcasm) constitute a real challenge for automatic systems. While the sneaky and tricky nature of subtle messages might be perceived as less hurtful with respect to the same content expressed clearly, such abuse is at least as harmful as overt abuse. In this paper, we first provide an in-depth and systematic analysis of 7 standard benchmarks for HS detection, relying on a fine-grained and linguistically-grounded definition of implicit and subtle messages. Then, we experiment with state-of-the-art neural network architectures on two supervised tasks, namely implicit HS and subtle HS message classification. We show that while such models perform satisfactory on explicit messages, they fail to detect implicit and subtle content, highlighting the fact that HS detection is not a solved problem and deserves further investigation [74].

### 7.2.5 Generating Adversarial Examples for Implicit Hate Speech Detection

**Participants:** Nicolás Benjamín Ocampo, Elena Cabrio, Serena Villata.

Research on abusive content detection on social media has primarily focused on explicit forms of hate speech (HS), that are often identifiable by recognizing hateful words and expressions. Messages containing linguistically subtle and implicit forms of hate speech still constitute an open challenge for automatic hate speech detection. In this paper, we propose a new framework for generating adversarial implicit HS short-text messages using Auto-regressive Language Models. Moreover, we propose a strategy to group the generated implicit messages in complexity levels (EASY, MEDIUM, and HARD categories) characterizing how challenging these messages are for supervised classifiers. Finally, we propose a “build it, break it, fix it”, training scheme using HARD messages showing how iteratively retraining on HARD messages substantially leverages state-of-the-art models’ performances on implicit HS benchmarks [72].

### 7.2.6 Bridging Implicit and Explicit Hate Speech Embedding Representations

**Participants:** Nicolás Benjamín Ocampo, Elena Cabrio, Serena Villata.

Research on automatic hate speech (HS) detection has mainly focused on identifying explicit forms of hateful expressions on user-generated content. Recently, a few works have started to investigate methods to address more implicit and subtle abusive content. However, despite these efforts, automated systems still struggle to correctly recognize implicit and more veiled forms of HS. As these systems heavily rely on proper textual representations for classification, it is crucial to investigate the differences in embedding implicit and explicit messages. Our contribution to address this challenging task is fourfold. First, we present a comparative analysis of transformer-based models, evaluating their performance across five datasets containing implicit HS messages. Second, we examine the embedding representations of implicit messages across different targets, gaining insight into how veiled cases are encoded. Third, we compare and link explicit and implicit hateful messages across these datasets through their targets, enforcing the relation between explicitness and implicitness and obtaining more meaningful embedding representations. Lastly, we show how these newer representation maintains high performance on HS labels, while improving classification in borderline cases [73].

### 7.2.7 Fallacious Argument Classification in Political Debates

**Participants:** Pierpaolo Goffredo, Shohreh Haddadan, Vorakit Vorakitphan, Elena Cabrio, Serena Villata.

Fallacies play a prominent role in argumentation since antiquity due to their contribution to argumentation in critical thinking education. Their role is even more crucial nowadays as contemporary argumentation technologies face challenging tasks as misleading and manipulative information detection in news articles and political discourse, and counter-narrative generation. Despite some work in this direction, the issue of classifying arguments as being fallacious largely remains a challenging and an unsolved task. Our contribution is twofold: first, we present a novel annotated resource of 31 political debates from the U.S. Presidential Campaigns, where we annotated six main categories of fallacious arguments (i.e., ad hominem, appeal to authority, appeal to emotion, false cause, slogan, slippery slope) leading to 1628 annotated fallacious arguments; second, we tackle this novel task of fallacious argument classification and we define a neural architecture based on transformers outperforming state-of-the-art results and standard baselines. Our results show the important role played by argument components and relations in this task [63].

### 7.2.8 DISPUTool 2.0: A Modular Architecture for Multi-Layer Argumentative Analysis of Political Debates

**Participants:** Pierpaolo Goffredo, Elena Cabrio, Serena Villata.

Political debates are one of the most salient moments of an election campaign, where candidates are challenged to discuss the main contemporary and historical issues in a country. These debates represent a natural ground for argumentative analysis, which has always been employed to investigate political discourse structure and strategy in philosophy and linguistics. In this paper, we present DISPUTool 2.0, an automated tool which relies on Argument Mining methods to analyze the political debates from the US presidential campaigns to extract argument components (i.e., premise and claim) and relations (i.e., support and attack), and highlight fallacious arguments. DISPUTool 2.0 allows also for the automatic analysis of a piece of a debate proposed by the user to identify and classify the arguments contained in the text. A REST API is provided to exploit the tool's functionalities [62].

The new version of the DISPUTool demo is publicly accessible at [3ia-demos.inria.fr/disputool/](https://3ia-demos.inria.fr/disputool/).

### 7.2.9 Argument-based Detection and Classification of Fallacies in Political Debates

**Participants:** Pierpaolo Goffredo, Mariana Chaves Espinoza, Elena Cabrio, Serena Villata.

Fallacies are arguments that employ faulty reasoning. Given their persuasive and seemingly valid nature, fallacious arguments are often used in political debates. Employing these misleading arguments in politics can have detrimental consequences for society, since they can lead to inaccurate conclusions and invalid inferences from the public opinion and the policymakers. Automatically detecting and classifying fallacious arguments represents therefore a crucial challenge to limit the spread of misleading or manipulative claims and promote a more informed and healthier political discourse. Our contribution to address this challenging task is twofold. First, we extend the ElecDeb60To16 dataset of U.S. presidential debates annotated with fallacious arguments, by incorporating the most recent Trump-Biden presidential debate. We include updated tokenlevel annotations, incorporating argumentative components (i.e., claims and premises), the relations between these components (i.e., support and attack), and six categories of fallacious arguments (i.e., Ad Hominem, Appeal to Authority, Appeal to Emotion, False Cause, Slippery Slope, and Slogans). Second, we perform the twofold task of fallacious argument detection and classification by defining neural network architectures based on Transformers models, combining text, argumentative features, and engineered features. Our results show the advantages of complementing transformer-generated text representations with non-textual features [63].

### 7.2.10 Augmenting participation, co-creation, trust and transparency in Deliberative Democracy at all scales through Argument Mining

**Participants:** Cristian Cardellino, Elena Cabrio, Serena Villata.

ORBIS is a project<sup>2</sup> starting in 2023 that aims to connect citizenship and policy making institutions by providing a solution to enable the transition to a more inclusive, transparent and trustful Deliberative Democracy in Europe. The project shapes and supports new democratic models that are developed through deliberative democracy processes. In the context of this project, we are currently working on the creation and annotation of a dataset of public debates for exploring semi-supervised learning for Argumentation Mining in deliberative democracy. The proposed models will be integrated in the DISPUTool Model.

### 7.2.11 Unveiling fake news through argumentative evidence

**Participants:** Xiaou Wang, Elena Cabrio, Serena Villata.

The need for automated and effective fact-checking (FC) has become urgent with the increase in the spread of harmful content on social media. Recently, fake news classification (FNC) has evolved to incorporate justifications provided by fact-checkers to explain their decisions, improving the transparency of the online content assessment process. In this research line, in the context of the ATTENTION project, we argue that an argumentative representation of fact-checkers' justifications on why they label certain content as fake can enhance FNC systems in terms of both precision and explainability. To address this challenging task, we built LIARArg, a novel annotated linguistic resource composed of 2832 news and their justifications. LIARArg extends the 6-label FNC dataset LIARPLUS with argument structures, leading to the first FNC dataset annotated with argument components (i.e., claim and premise) and fine-grained relations, i.e., (partial) attack and (partial) support. To integrate argumentation in FNC, we propose a joint learning method combining argument mining and FNC which outperforms SOTA approaches, especially for news with intermediate truthfulness labels. We demonstrate the essential role of argument relations in enhancing the classifier performance and, more importantly, we highlight the contribution of fine-grained relations which allow an extra performance boost. We show that the argumentative representation of human justifications enables the explanation of the classifier assessment in a chain-of-thought manner, paving a promising avenue for research in explainable fact-checking.

## 7.3 Vocabularies, Semantic Web and Linked Data Based Knowledge Representation and Artificial Intelligence Formalisms on the Web

### 7.3.1 Semantic Web for Biodiversity

**Participants:** Franck Michel, Catherine Faron.

This activity addresses the challenges of exploiting knowledge representation and semantic Web technologies to enable data sharing and integration in the biodiversity area. The collaboration with the "Muséum National d'Histoire Naturelle" of Paris (MNHN) goes on along several axes.

In 2023, the MNHN has further extended its use of the SPARQL Micro-Services architecture and framework that we maintain, to help biologists in editing taxonomic information by confronting multiple, heterogeneous data sources. This architecture, as well as these biodiversity use cases, was the object of a presentation at the Knowledge Graph Conference [71], an industry-focused conference.

Furthermore, we have collaborated on the submission of a joint project on the representation of species life traits using the Plinian core vocabulary [40] (GBIF CfP Capacity Enhancement Support Programme). This project was accepted and will start in early 2024.

<sup>2</sup>ORBIS Web Site

### 7.3.2 Semantic Web for Life Sciences

**Participants:** Pierre Monnin.

Life sciences produce and consume vast amounts of scientific data. The graph-structured nature of these data naturally leads to data-driven research efforts leveraging Semantic Web and Knowledge Graph technologies. In a survey and position paper, we discussed recent developments and advances in the use of graph-based technologies in life sciences and set out a vision for how these technologies will impact these fields into the future [35].

Among such usages, knowledge graph construction and management is a well established topic. One subtask lies in matching similar or related units across datasets to identify possible overlaps. In this direction, we proposed the new track “**Pharmacogenomics**” in the international challenge “**Ontology Alignment Evaluation Initiative**”. This track focuses on the matching of pharmacogenomic knowledge units, which are  $n$ -ary tuples involving components of three distinct types (drugs, genetic factors, and phenotypes). None of the approaches participating in the 2023 campaign were able to produce alignments, which highlights the importance of the research question behind this task and the need to develop appropriate approaches or enrich existing ones [77].

### 7.3.3 Evolutionary agent-based evaluation of the sustainability of different knowledge sharing strategies in open multi-agent systems

**Participants:** Stefan Sarkadi, Fabien Gandon, Andrea Tettamanzi.

The advancement of agent technologies and their deployment in various fields of application has brought numerous benefits w.r.t. knowledge or data gathering and processing. However, one of the key challenges in deploying artificial intelligent agents in an open environment like the Web is their interoperability. Even though research and development of agent technologies on the Semantic Web has advanced significantly, artificial agents live on the Web in silos, that is in very limited domains, isolated from other systems and agents that live on the Web. In this work we setup a simulation framework and evaluation based on evolutionary agent-based modeling to empirically test how sustainable different strategies are for knowledge sharing in open multi-agent systems and to see which of these strategies could actually enable global interoperability between Web agents. The first results are showing the interest of translation-based approaches and the need for further incentives to support these [80].

### 7.3.4 W3C Data activity and AC Rep

**Participants:** Rémi Ceres, Pierre-Antoine Champin, Fabien Gandon, Franck Michel, Olivier Corby.

Semantic Web technologies are based on a set of standards developed by the **World Wide Web consortium (W3C)**. Participation in these standardization groups gives to researcher the opportunity to promote their results towards a broad audience, and to keep in touch with an international community of experts. Wimmics has a long history of being involved in W3C groups.

As W3C fellow, Pierre-Antoine Champin also works within the W3C team to support Semantic Web related working groups and promote the emergence of new ones, to ensure the necessary evolutions of these technologies. Two new groups were created in 2022, where Wimmics members are largely involved. **RDF Canonicalization and Hash** aims at providing a way to hash RDF data independently of the way they are represented. **RDF-star** is chartered to publish the new version of RDF and SPARQL, extending them with the ability to make statements about statements.

Furthermore, Wimmics has been an active member of the **Knowledge Graph Construction Community Group** since its creation in 2020. In 2023, a first major outcome of this group was the publication of the

RDF Mapping Language (RML) Ontology, a modular redesign of the first RML ontology, that leverages a decade of experience in mapping heterogeneous data to RDF [66].

The work has started towards the creation of a working group to standardize the Solid protocol. The **Solid project** was started by Tim Berners-Lee, inventor of the Web, and builds on Semantic Web standards to promote the (re-)decentralization of the Web.

Finally, Fabien Gandon remains the W3C AC Rep for Inria representing institute in all standardization processes and W3C meetings (annual W3C TPAC conference and W3C AC Meeting).

### 7.3.5 AutomaZoo: Automatic annotation of an ancient Zoological Corpus

**Participants:** Arnaud Barbe, Molka Dhouib, Catherine Faron.

This project is part of the scientific actions carried out by the Zoomathia international research network which aims to study the constitution and transmission of zoological knowledge from Antiquity to the Middle Ages. The aim of the project is to produce a corpus of textual resources semantically annotated by a graph of ancient zoological knowledge, respecting semantic web standards, interoperable and published on the open data web. In this context, our initial step involves constructing the ZooKG-Pliny knowledge Graph from a manual annotation of Pliny's *Naturalis Historia* using concepts gathered in the thesaurus TheZoo. ZooKG-Pliny is based on a semantic model that formalizes knowledge about the annotations of zoological information in texts. ZooKG-Pliny allows the integration and the interrogation of relevant knowledge in order to support epistemologists, historians and philologists in their analysis of these texts and knowledge transmission through them [85, 102].

The second task includes providing semantic annotations to the paragraphs of the ancient zoological text *Naturalis Historia* (Pliny the Elder) according to the concepts in the domain thesaurus TheZoo. For that, we consider two approaches to automatically classify paragraphs of Pliny's *Naturalis Historia* on ancient zoology into one or more macro collections of concepts (i.e. "Places", "Anthroponym", etc.) from the TheZoo thesaurus: (i) The baseline method consists into training a Support Vector Machine for each collection separately. (ii) The knowledge-based method extends the baseline by using the hierarchical information extracted from the thesaurus [81].

### 7.3.6 Contributions to Agronomical knowledge graphs

**Participants:** Nadia Yacoubi Ayadi, Catherine Faron, Franck Michel, Olivier Corby, Fabien Gandon.

The interest of the agronomy and biodiversity communities in the development of crop models coupled to weather and climate models has led to the need for datasets of meteorological observations in which data are semantically described and integrated. For this purpose, in 2022, we proposed a semantic model to represent and publish meteorological observational data as Linked Data. Our model reuses a network of existing ontologies to capture the semantics of data, it covers multiple dimensions of meteorological data including geospatial, temporal, observational, and provenance characteristics.

Based on this semantic model, we built the WeKG-MF knowledge graph considering the open weather observations published by Météo-France. In 2023, we extended this knowledge graph to include 10 years (2012-2022) of historical weather observations. In order to enable an interactive visualization of the WeKG-MF graph, we released a Web application that enables lay users to visualize weather observational data at different levels of spatio-temporal granularity, and hence, it offers multi-level 'tours' based on high-level aggregated views together with on-demand fine-grained data, and this through a unified multi-visualizations interface [83]).

Therefore, in order to meet specific needs of experts in the agriculture domain and provide them with access to established and significant climate parameters such as the total monthly amount of precipitation or the number of days where the amount of precipitation is higher than a threshold over a given period,

we investigated the use of WeKG-MF in order to generate and calculate a number of significant agro-meteorological parameters identified in literature. Thus, based on the WMO recommendations, we implemented different calculation methods of some of these parameters in SPARQL and stored query results as part of the WeKG-MF graph. Other agro-meteorological parameters are calculated on the fly. We devised a strategy to find a balance between calculating on-the-fly agro-meteorological parameters and pre-calculating and storing significant SPARQL queries results in order to manage the workload of the WeKG-MF SPARQL endpoint ([83]). We enriched the WeKG-MF graph using an external API of weather data, namely Open Météo, that provides new agro-meteorological parameters such as solar radiation and evapotranspiration. And we developed new visualizations in order to allow experts in agriculture to easily compute, visualize and compare agro-meteorological parameters across time and space.

On another side, we presented the CoffeeWKG knowledge graph for weather conditions of coffee regions in Colombia constructed with data extracted from yearbooks based on the WeKG semantic model. The CoffeeWKG knowledge graph can be [downloaded](#) and queried to ask for historical weather conditions in coffee regions. The main purpose of this graph is intended to correlate this information with coffee crop data to understand how these conditions may favor or hamper coffee production or diseases (Workshop Paper 2023).

Independently of these works, we carried out a study [32] on the alignment of two complementary knowledge graphs useful in agriculture: the thesaurus of cultivated plants in France named French Crop Usage (FCU) and the French national taxonomic repository TAXREF for fauna, flora, and fungi. Both knowledge graphs contain vernacular names of plants but those names are ambiguous. Thus, in the context of the D2KAB project, a group of agricultural experts produced some mappings from FCU crops to TAXREF. The metadata for the mappings and the mapping set were encoded with the Simple Standard for Sharing Ontological Mappings (SSSOM), a new model which, among other qualities, offers means to report on provenance of particular interest for this study. The produced mappings are available for download in Recherche Data Gouv, the federated national platform for research data in France.

### 7.3.7 WheatGenomicsSLKG: a Knowledge graph for wheat genomics studies

**Participants:** Nadia Yacoubi Ayadi, Catherine Faron, Franck Michel, Olivier Corby, Fabien Gandon.

One of the main challenges in wheat genomics is the large size and complexity of the wheat genome. Harvesting scientific literature may support experts in wheat genomics to understand hidden interactions between genomic entities, and the relationships between genotype and phenotypes, based on their co-occurrence in scientific publications. The main purpose is to bridge the gap between scientific results presented in publications and experts needs to explore literature and find answers to complex questions.

In the course of 2022, in the context of the D2KAB project, we proposed to structure and integrate semantic annotations, automatically extracted from scientific articles using NLP tools, into a knowledge graph. Thus, we rely on the Open Annotation Ontology (OA) to describe, structure and integrate named entities annotations and their occurrence contexts in texts. Domain specific vocabularies are also reused to describe bibliographic information of scientific publications. The resulting model was automatically populated using a mapping-based transformation pipeline implemented with Morph-xR2RML 6.1.21, and the SPARQL micro-services 6.1.18.

In 2023, we have produced and published an extended release of this knowledge graph (see section Open data). We have also used the semantic model as well as the pipeline to build two other knowledge graphs in the context of the D2KAB project, namely RiceGenomicsSLKG and PHB-KG.

The relevance of the semantic model was validated by implementing a set of competency questions with SPARQL queries which reflect how the KGs can be queried to retrieve co-occurrence of NE in texts. We also implemented different queries that aim to retrieve annotations from these different graphs in a federated manner. Our main purpose is to bridge the gap between scientific results presented in publications and experts needs to explore literature and find answers to complex questions ([83]).

### 7.3.8 Ontology engineering: tooling and methodology

**Participants:** Fabien Gandon, Nicolas Robert.

We contributed to the Agile and Continuous Integration for Modular Ontologies and Vocabularies (ACIMOV) ontology engineering methodology for developing ontologies and vocabularies. ACIMOV extends the SAMOD agile methodology to (1) ensure alignment to selected reference ontologies; (2) plan module development based on dependencies; (3) define ontology modules that can be specialized for specific domains; (4) empower active collaboration among ontology engineers and domain experts; (5) enable application developers to select views of the ontology for their specific domain and use case. ACIMOV adopts the standard git-based approach for coding, leveraging agility and DevOps principles. It has been designed to be operationalized using collaborative software development platforms such as Github or Gitlab, and tooling with continuous integration and continuous deployment workflows (CI/CD workflows) that run syntactic and semantic checks on the repository, specialize modules, generate and publish the ontology documentation [64].

## 7.4 Analyzing and Reasoning on Heterogeneous Semantic Graphs

### 7.4.1 Corese Semantic Web Factory

**Participants:** Rémi Ceres, Fabien Gandon, Olivier Corby.

Corese [107], an open-source Semantic Web platform, implements W3C languages such as RDF, RDFS, OWL RL, SHACL, SPARQL, and extensions including SPARQL Function, SPARQL Transformation, and SPARQL Rule.

In the enhancement of Corese's distribution, two new interfaces, Corese-GUI and Corese-Command, were launched on Flathub. Additionally, a one-click installation script for Corese-Command is now available for Linux and MacOS.

The documentation of Corese has been fully updated and is accessible at <sup>3</sup>.

The new interface, Corese-Command, supplements existing ones such as Corese-Library, Corese-GUI, Corese-Server, and Corese-Python. Corese-Command, evolving from the previous Corese-CLI, enables command-line usage of Corese. It encompasses subcommands for converting RDF file formats, running SPARQL queries, performing SHACL validation on RDF datasets, and executing SPARQL queries on remote endpoints. Improvements in file loading now allow handling of local files, URLs, or directories.

All interfaces have been unified to support Corese configuration files in properties format.

Enhancements include bug fixes in Corese-Python, addition of Markdown result format for SPARQL, and N-Quads RDF serialization.

Relevant websites include the Corese project site at [Corese Web site](#) and the GitHub repository at [Corese github URL](#).

### 7.4.2 SHACL Extension

**Participants:** Olivier Corby, Iliana Petrova, Fabien Gandon, Catherine Faron.

In the context of a collaboration with Stanford University, we worked on extensions of W3C SHACL Shape Constraint Language <sup>4</sup>.

We conducted a study on large, active, and recognized ontology projects (ex. Gene Ontology, Human Phenotype Ontology, Mondo Disease Ontology, Ontology for Biomedical Investigations, OBO Foundry, etc.) as well as an analysis of several existing tools, methodologies and guidelines for ontological engineering.

<sup>3</sup>[Documentation of Corese](#)

<sup>4</sup>see [SHACL Recommendation](#)



As a result we identified several sets of ontology validation constraints that fall into six big clusters: **i)** formalization/modeling checks; **ii)** terminological/writing checks; **iii)** documentation/ editorial practices, terminology-level checks; **iv)** coherence between terminology and formalization; **v)** metamodel-based checks; **vi)** integration/interoperability/data checking. These can be further refined depending on whether they are specific to RDFS/OWL meta-model, domain/ontology specific, or Linked Data specific. This precise categorization of the ontology validation constraints allowed us to analyze the needs and impact of the extension we are targeting in terms of semantic expressiveness, computational complexity of the validation, and current syntax of the SHACL language.

We then concentrated on the formalization of the semantic extensions and their validation methods and came up with a proposal of a corresponding syntactic extensions of SHACL.

The formal specification of the identified extensions enabled us to proceed with the implementation of a prototype plugin for Protégé (Stanford's widely used ontology editor) based on the Corese engine and which extends the SHACL standard with these newly proposed capabilities.

### 7.4.3 Automatic Assessment of Text Difficulty

**Participants:** Molka Dhouib, Catherine Faron.

In the continuation of our work in the field of educational resource recommendation, we explored the possibility of automatically assess text difficulty. For many educational applications, the difficulty of a text is a key information. Today, the best results for this task are obtained by using NLP and deep learning techniques. However, the use of these methods can result in the loss of statistical linguistic information that is important for determining text readability more accurately. In this context, we propose an approach for assessing text readability by combining neural network models with linguistic features extracted from the text and integrated into the model to improve the quality of the neural network models. Experimental results show that this combination outperforms state of the art approaches [87].

### 7.4.4 On the Automatic Assessment of Natural Language Expert Explanations in Medicine

**Participants:** Santiago Marro, Theo Alkibiades Collias, Elena Cabrio, Serena Villata.

The importance of explanations in decision-making, particularly in the medical domain, has been widely recognized. However, the evaluation of the quality of these explanations remains a challenging task. In this work, we propose a novel approach for assessing and evaluating the reasons provided in explanations about clinical cases. Our approach leverages an external knowledge base and a defined prevalence function to score each reason based on its pertinence in the domain. By applying a deterministic prevalence function, we ensure total transparency of the reasons assessment, facilitating a precise explanation of the rationale behind the scoring hierarchy of each reason. We demonstrate the effectiveness of our approach in clinical cases, where medical experts explain the rationale behind a specific diagnosis and why other potential diagnoses are dismissed. Our methodology provides a nuanced and detailed evaluation of the explanation, contributing to a more comprehensive understanding of the decision-making process [68]

### 7.4.5 Explanatory Argumentation in Natural Language for Correct and Incorrect Medical Diagnoses

**Participants:** Benjamin Molinet, Santiago Marro, Elena Cabrio, Serena Villata.

In the context of the ANTIDOTE project, we investigated the generation of natural language argument-based explanations in medicine. A huge amount of research is carried out nowadays in Artificial Intelligence to propose automated ways to analyze medical data with the aim to support doctors in delivering

medical diagnoses. However, a main issue of these approaches is the lack of transparency and interpretability of the achieved results, making it hard to employ such methods for educational purposes. It is therefore necessary to develop new frameworks to enhance explainability in these solutions. We present a novel full pipeline to generate automatically natural language explanations for medical diagnoses. The proposed solution starts from a clinical case description associated with a list of correct and incorrect diagnoses and, through the extraction of the relevant symptoms and findings, enriches the information contained in the description with verified medical knowledge from an ontology. Finally, the system returns a pattern-based explanation in natural language which elucidates why the correct (incorrect) diagnosis is the correct (incorrect) one.

#### 7.4.6 ACTA Module Upgrade

**Participants:** Cristian Cardellino, Theo Alkibiades Collias.

For the ANTIDOTE project we were in charge of a major refactoring of the original ACTA module <sup>5</sup>, enabling the latest Large Language Models (LLM) models from Hugging Face. The new module is a standalone Python Library<sup>6</sup> available under Apache License. We are also in the process of refactoring the code for the web application of ACTA to better integrate it with the new modules and writing a code easier to maintain.

#### 7.4.7 Qualitative evaluation of arguments in persuasive essays

**Participants:** Elena Cabrio, Serena Villata, Santiago Marro.

Argumentation is used by people both internally, by evaluating arguments and counterarguments to make sense of a situation and take a decision, and externally, e.g., in a debate, by exchanging arguments to reach an agreement or to promote an individual position. In this context, the assessment of the quality of the arguments is of extreme importance, as it strongly influences the evaluation of the overall argumentation, impacting on the decision making process. The automatic assessment of the quality of natural language arguments is recently attracting interest in the Argument Mining field. However, the issue of automatically assessing the quality of an argumentation largely remains a challenging unsolved task.

Our contribution is twofold: first, we present a novel resource of 402 student persuasive essays, where three main quality dimensions (i.e., *cogency*, *rhetoric*, and *reasonableness*) have been annotated, leading to 1908 arguments tagged with quality facets; second, we address this novel task of argumentation quality assessment proposing a novel neural architecture based on graph embeddings, that combines both the textual features of the natural language arguments and the overall argument graph, i.e., considering also the support and attack relations holding among the arguments. Results on the persuasive essays dataset outperform state-of-the-art and standard baselines' performance. See [68, 69, 100].

#### 7.4.8 RDF Mining

**Participants:** Ali Ballout, Catherine Faron, Rémi Felin, Andrea Tettamanzi.

The Shapes Constraint Language (SHACL) is a W3C recommendation which allows to represent constraints in RDF and validate RDF data graphs against these constraints. A SHACL validator produces a validation report whose result is false for a shape graph as soon as there is at least one node in the

<sup>5</sup>Acta module web site

<sup>6</sup>Acta library on Github

RDF data graph that does not conform to the shape. This Boolean result of the validation of an RDF data graph against an RDF shape graph is not suitable for discovering new high-potential shapes from the RDF data. Therefore, we have proposed a probabilistic framework to accept shapes with a realistic proportion of nodes in an RDF data graph that does not conform to it. Based on this framework, we have also proposed an extension of the SHACL validation report to express a set of metrics including the generality and likelihood of shapes and we define a method to test a shape as a hypothesis test [61, 60].

The task of evaluating the fitness of a candidate axiom against known facts or data is known as *candidate axiom scoring*. Being able to accurately score candidate axioms is a prerequisite for automatic schema or ontology induction, but can also be useful for ontology and/or knowledge graph validation. Accurate axiom scoring heuristics are often heavy to compute, which is a big problem if one wants to exploit them in iterative search methods like level-wise generate-and-test or evolutionary algorithms, where large numbers of candidate axioms need to be scored. We have tackled the challenge of learning a predictive model as a surrogate to reasoning, that predicts the acceptability of candidate class axioms, that is fast to execute yet accurate enough to be used in such settings. For this purpose, we have leveraged a semantic similarity measure extracted from the subsumption hierarchy of an ontology and we have proven that our proposed method is able to learn the acceptability labels of candidate OWL class axioms with high accuracy and that it can do so for multiple types of OWL class axioms [53].

#### 7.4.9 Capturing Geospatial Knowledge from Real-Estate Advertisements

**Participants:** Lucie Cadorel, Andrea Tettamanzi.

In the framework of a CIFRE thesis with Septeo Proptech, we have proposed a workflow to extract geographic and spatial entities based on a BiLSTM-CRF architecture (Bidirectional Long Short-Term Memory Conditional Random Field) with a concatenation of several text representations and to extract spatial relations, to build a structured Geospatial knowledge base. This pipeline has been applied to the case of French housing advertisements, which, generally provide information about a property's location and neighbourhood. Our results show that the workflow tackles French language and the variability and irregularity of housing advertisements, generalizes Geoparsing to all geographic and spatial terms, and successfully retrieves most of the relationships between entities from the text [86].

Text representations are widely used in NLP tasks such as text classification. Very powerful models have emerged and been trained on huge corpora for different languages. However, most of the pre-trained models are domain-agnostic and fail on domain-specific data. We performed a comparison of different text representations applied to French Real Estate classified advertisements through several text classification tasks to retrieve some key attributes of a property. Our results demonstrate the limitations of pre-trained models on domain-specific data and small corpora, but also the strength of text representation, in general, to capture underlying knowledge about language and stylistic specificities [58].

#### 7.4.10 ISSA: semantic indexing of scientific articles and advanced services

**Participants:** Anna Bobasheva, Catherine Faron, Aline Menin, Franck Michel, Marco Winckler.

As a continuation of the ISSA project that ended in 2022, in 2023 we started the **ISSA 2** project. In line with open science and the FAIR principles, ISSA 2 proposes a generic method to explore the data available in an open archive in order to 1) extract new knowledge, 2) exploit this knowledge with a bibliometric objective and 3) propose services to researchers and documentalists, in particular in terms of bibliometrics and information retrieval. It relies on and extends the outcomes of the ISSA project, in particular the semantic index of the publications of an open archive (metadata, descriptors and named entities mentioned in the text, linked to standard knowledge bases). The proposed methods exploit data mining techniques as well as techniques for the construction, publication and exploitation of knowledge

graphs. In addition to Agritrop, CIRAD's open archive that served as a use case in ISSA, ISSA 2 considers the HAL instance of the UR EuroMov Digital Health in Motion.

In 2023, we produced a new version of the indexing pipeline (6.1.14) that extends the scope of data represented in RDF, supports multi-instance pipelines, and facilitates the deployment process. It has been set in production and now performs the weekly indexing of both archives: Agritrop and HAL EuroMov Digital Health in Motion. We also developed a new search interface capable of exploiting the rich semantics of reference vocabularies to provide more intelligent search results than simple keyword-base searches. Finally, we developed a novel visualization technique based on polymorphic glyphs to support the visual exploration of hierarchical spatio-temporal data [70].

The pipeline and search interface are DOI-identified and available under an open license on public repositories. The knowledge graph produced by the pipeline for the Agritrop archive is also made public as a downloadable dump (DOI: [10.5281/zenodo.10381606](https://doi.org/10.5281/zenodo.10381606)) and through a public SPARQL endpoint. We published new results in a journal article: [44].

#### 7.4.11 IndeGx: A Model and a Framework for Indexing Linked Datasets and their Knowledge Graphs with SPARQL-based Test Suits

**Participants:** Pierre Maillot, Olivier Corby, Catherine Faron, Fabien Gandon, Franck Michel.

The joint exploitation of RDF datasets relies on the knowledge of their content, of their endpoints, and of what they have in common. Yet, not every dataset contains a self-description, and not every endpoint can handle the complex queries used to generate such a description.

As part of the ANR DeKaloG, we proposed a standard-based approach to generate the description of a dataset. The description generated as well as the process of its computation are expressed using standard vocabularies and languages. We have implemented our approach into a framework, called IndeGx, to automatically generate the description of datasets and endpoints and collect them in an index. We have experimented IndeGx on a set of 339 active knowledge bases. The method and the results of our experimentation have been published in a journal article: [38].

Several visualizations were also generated from IndeGx and are available online: [IndeGx Web Site](#).

#### 7.4.12 Metadatamatic: A tool for the description of RDF knowledge bases

**Participants:** Pierre Maillot, Olivier Corby, Catherine Faron, Fabien Gandon, Franck Michel.

During the experimentations done as part of the development of IndeGx, we observed that less than 10% of accessible RDF datasets offer a description of their provenance. We theorize that this is partly due to the difficulty for data providers to learn how to create such descriptions. Some initiatives such as the COST KG working group, or a European Commission action have proposed guides consisting of lists of mandatory and recommended classes and properties to use.

We propose Metadatamatic, an online tool based on the before-mentioned initiatives to generate a KB description. Metadatamatic aims to bypass this learning problem and help further the description of a KB. Metadatamatic generates a description using well-established vocabulary from a simple Web form. It also offers to extract some parts of the description, such as the list of vocabularies used in the data, automatically from the content of the KB. We hope that this tool will lead to an improvement of the usability of Linked Data in general.

Metadatamatic was presented as a conference poster ([67]) and is available online: [Metadatamatic](#)

#### 7.4.13 Learning Pattern-Based Extractors from Natural Language and Knowledge Graphs: Applying Large Language Models to Wikipedia and Linked Open Data

**Participants:** Célian Ringwald, Fabien Gandon, Catherine Faron, Franck Michel, Hanna Abi Akl.

Whether automatically extracted from structured elements of articles or manually populated, the open and linked data published in DBpedia, and Wikidata offer rich and structured complementary views of the textual descriptions found in Wikipedia. However, the unstructured text of Wikipedia articles contains a lot of information that is still missing in DBpedia and Wikidata. Extracting them would be interesting to improve the coverage and quality of these knowledge graphs (KG) and this would have an important impact on all downstream tasks.

This work proposes to exploit the dual bases formed from Wikipedia pages and Linked Open Data (LOD) bases covering the same subjects in natural language and in RDF, to produce RDF extractors targeting specific RDF patterns and tuned for a given language. Therefore, the main research question is: Can we learn efficient customized extractors targeting specific RDF patterns from the dual base formed by Wikipedia on one hand, and DBpedia and Wikidata on the other hand?

The landscape of the research field drawn at the intersection of language models and knowledge graphs is very dynamic and quickly evolving. For this reason, as the first step of this work, we designed an extended systematic review of the latest NLP approaches to KG extraction.

In a second step we started to design a first dataset focused on datatype properties. We restricted the selection of our training to facts respecting a given SHACL shape and information that could be found in the Wikipedia abstract. Then to learn how to extract relations with datatype properties from natural language, we exploited pre-trained encoder-decoder models, and more precisely BART (denoising autoencoder sequence-to-sequence model). We explored several aspects of the task formulation that could impact the generation of triples in this context: the size of the model, the size of the learning sample needed to learn a given SHACL pattern, and the syntax of the triples. The current work will be published in two papers accepted at AAAI 2024 (Association for the Advancement of Artificial Intelligence).

Alongside the pursuit of the main objectives, the work started by a participation to the [International Semantic Web Research Summer School](#) and the [Generative Modeling Summer School](#) where it was presented [104]. ISWS was the opportunity to collaborate and work on a question related to generative IA, and more especially on how to render triples describing fictional characters as illustrative images [50]. This collaboration led us to attend at the [King's College Knowledge prompting hackathon](#) where we had the occasion to develop a first proposal related to visual relation extraction.

#### 7.4.14 AI agent to convert natural language questions into SPARQL queries

**Participants:** Emma Tysinger, Fabien Gandon.

An experimental knowledge graph (KG) driven framework (10.26434/chemrxiv-2023-sljbt) was recently introduced to facilitate the integration of heterogeneous data types, encompassing both experimental data (mass spectrometry annotation, results from biological screening and fractionation) as well as meta-data available on the Web (such as taxonomies and metabolite databases). Although this KG efficiently encapsulates the different data structures and semantic relationships, retrieving specific information through structured or visually queries or even programmatically, is not trivial. In a collaborative project KG-Bot we designed and implemented an AI agent that can convert natural language questions into SPARQL queries and programmatic data-mining tasks and generate adapted visualization. By leveraging the potential of emerging Large Language Models (LLMs) to understand semantic relationships encapsulated in KGs and mentioned in the questions, the agent autonomously iterates to construct a SPARQL query of any submitted natural language question. After retrieving the necessary information from the KG, the agent provides a preliminary interpretation of the results in natural language, along with relevant visualizations and statistics [82].

#### 7.4.15 Hybridizing machine learning and knowledge graphs: vision, model evaluation, and knowledge injection

**Participants:** Pierre Monnin.

Knowledge graphs (KGs) are nowadays largely adopted, representing a successful paradigm of how symbolic and transparent AI can scale on the World Wide Web. However, they are generally tackled by Machine Learning (ML) and mostly numeric based methods such as graph embedding models (KGEMs) and deep neural networks (DNNs). The latter methods have been proved efficient but lack some major characteristics such as interpretability and explainability. Conversely, these characteristics are intrinsically supported by symbolic AI methods and artefacts, thus motivating a research effort to hybridize machine learning and knowledge graphs.

In this direction, in a vision paper, we introduced some of the main existing methods for combining KGs and ML and highlighted research gaps and perspectives that we deem promising and currently under-explored for the involved research communities [36]. These span from KG support for prompting Large Language Models, integration of KG semantics in ML models to symbol-based methods, interpretability of ML models, and the need for improved benchmark datasets.

We investigated the evaluation of KGEMs for the task of link prediction which aims at predicting the missing tail of a triple  $(h, r, ?)$  or the missing head of a triple  $(?, r, t)$ . Usually, their performance is assessed using rank-based metrics such as Hits@K, MR, or MRR, which evaluate their ability to give high scores to ground-truth entities. We extended this evaluation framework with a semantic perspective by proposing the metric Sem@K that measures the capability of models to predict valid entities w.r.t. domain and range constraints [37]. We designed different versions of Sem@K depending on the presence or not of types, potentially hierarchically organized. We also evaluated a wide array of KGEMs which showed that Sem@K offers an additional perspective beside rank-based metrics, and that some families of models perform better than others.

We proposed one possible way to inject knowledge represented as RDFS constructs into KGEMs to learn embeddings that better capture semantics [65]. More specifically, we introduced *protographs* that are built from relation signatures (*i.e.*, domain and range) and the class hierarchy. Embeddings are learned on such protographs and are supposed to encapsulate the semantics of the KGs. Protograph-based embeddings are then used to bootstrap embeddings learned on the actual KG, yielding better performance in downstream tasks (link prediction, entity clustering, and node classification). These results highlight the interest of injecting knowledge into KGEMs, a current and vibrant direction of the research community.

#### 7.4.16 Semantic Overlay Network for Linked Data Access

**Participants:** Fabien Gandon, Mahamadou Toure.

We proposed and evaluated MoRAI (Mobile Read Access in Intermittent internet connectivity), a distributed peer-to-peer architecture organized in three levels dedicated to RDF data exchanges by mobile contributors. We presented the conceptual and technical aspects of this architecture as well as a theoretical analysis of the different characteristics. We then evaluated it experimentally and results show the relevance of considering geographical positions during data exchanges and of integrating RDF graph replication to ensure data availability in terms of requests completion rate and resistance to crash scenarios [101].

## 8 Bilateral contracts and grants with industry

### 8.1 Bilateral contracts with industry

### Curiosity Collaborative Project

**Participants:** Catherine Faron, Oscar Rodríguez Rocha, Molka Dhouib.

Partner: TeachOnMars.

This collaborative project with the TeachOnMars company started in October 2019. TeachOnMars is developing a platform for mobile learning. The aim of this project is to develop an approach for automatically indexing and semantically annotating heterogeneous pedagogical resources from different sources to build up a knowledge graph enabling to compute training paths, that correspond to the learner's needs and learning objectives.

### CIFRE Contract with Doriane

**Participants:** Andrea Tettamanzi, Rony-Dupuy Charles.

Partner: Doriane.

This collaborative contract for the supervision of a CIFRE doctoral scholarship, relevant to the PhD of Rony-Dupuy Charles, is part of Doriane's Fluidity Project (Generalized Experiment Management), the feasibility phase of which has been approved by the Terralia cluster and financed by the Région Sud-Provence Alpes Côte d'Azur and BPI France in March 2019. The objective of the thesis is to develop machine learning methods for the field of agro-vegetation-environment. To do so, this research work will take into account and address the specificities of the problem, i.e. data with mainly numerical characteristics, scalability of the study object, small data, availability of codified background knowledge, need to take into account the economic stakes of decisions, etc. To enable the exploitation of ontological resources, the combination of symbolic and connective approaches will be studied, among others. Such resources can be used, on the one hand, to enrich the available datasets and, on the other hand, to restrict the search space of predictive models and better target learning methods.

The PhD student will develop original methods for the integration of background knowledge in the process of building predictive models and for the explicit consideration of uncertainty in the field of agro-plant environment.

### CIFRE Contract with Kinaxia

**Participants:** Andrea Tettamanzi, Lucie Cadorel.

Partner: Kinaxia.

This thesis project is part of a collaboration with Kinaxia that began in 2017 with the Incertimmo project. The main theme of this project was the consideration of uncertainty for a spatial modeling of real estate values in the city. It involved the computer scientists of the Laboratory and the geographers of the **ESPACE Laboratory**. It allowed the development of an innovative methodological protocol to create a mapping of real estate values in the city, integrating fine-grained spatiality (the street section), a rigorous treatment of the uncertainty of knowledge, and the fusion of multi-source (with varying degrees of reliability) and multi-scale (parcel, street, neighbourhood) data.

This protocol was applied to the Nice-Côte d'Azur metropolitan area case study, serving as a test bed for application to other metropolitan areas.

The objective of this thesis, which was carried out by Lucie Cadorel under the supervision of Andrea Tettamanzi, was, on the one hand, to study and adapt the application of methods for extracting knowledge from texts (or text mining) to the specific case of real estate ads written in French, before extending them to other languages, and, on the other hand, to develop a methodological framework that makes it possible to detect, explicitly qualify, quantify and, if possible, reduce the uncertainty of the extracted information, in order to make it possible to use it in a processing chain that is finalized for recommendation or decision making, while guaranteeing the reliability of the results.

## Plan de Relance with Startin'Blox

**Participants:** Pierre-Antoine Champin, Fabien Gandon, Maxime Lecoq.

Partner: [Startin'Blox](#).

The subject of this project is to investigate possible solutions to build on top of the Solid architecture capabilities to discover services and access distributed datasets. This would rely on standardized search and filtering capabilities for the SOLID PODs, as well as on traversal or federated SPARQL query solving approaches to design a pilot architecture. We also intend to address performance issues via caching or indexing strategies in order to allow a deployment of the Solid ecosystem on a web scale.

## CP4SC Project - sub contract

**Participants:** Fabien Gandon, Rémi Ceres.

Partner: Atos.

Initiated in January 2023, the CP4SC project is a collaborative effort with Atos that focuses on developing a Cloud Platform for Smart Cities. This innovative platform is designed to integrate and analyze data from a wide range of city services and smart devices. As a single, secure access point for data management and dissemination, it promotes innovation and interdepartmental cooperation in urban areas. The project targets sustainable urban development by addressing three key areas: optimizing resource consumption in various living environments through energy management; reducing greenhouse gas emissions via mobility management and alternative transportation solutions; and enhancing environmental conservation and public health through strategic management and observation of the Earth's environment. In that context Atos has a contract with Wimmics to study the use of semantic technologies and CORESE in managing the data of this Smart City scenario.

## 9 Partnerships and cooperations

### 9.1 International initiatives

#### 9.1.1 Associate Teams in the framework of an Inria International Lab or in the framework of an Inria International Program

##### PROTEMICS

**Participants:** Olivier Corby, Catherine Faron, Fabien Gandon, Iliana Petrova.

**Title:** Protégé and SHACL extension to support ontology validation

**Duration:** 2020 -> 2023

**Coordinator:** Rafael Gonçalves (rafael.goncalves@stanford.edu)

**Partners:**

- Stanford University Stanford (USA)

**Inria contact:** Fabien Gandon



**Summary:** We propose to investigate the extension of the structure-oriented SHACL validation to include more semantics, and to support ontology validation and the modularity and reusability of the associated constraints. Where classical Logical (OWL) schema validation focuses on checking the semantic coherence of the ontology, we propose to explore a language to capture ontology design patterns as extended SHACL shapes organized in modular libraries. The overall objective of our work is to augment the Protégé editor with fundamental querying and reasoning capabilities provided by CORESE, in order to assist ontology developers in performing ontology quality assurance throughout the life-cycle of their ontologies

### 9.1.2 Participation in other International Programs

#### MERI-Wimmics collaboration

**Participants:** Antoine Vidal-Mazuy, Michel Buffa.

**Visited institution:** MERI team at CCRMA/Stanford

**Country:** USA

**Dates:** 4-22 May, 2023

**Context of the visit:** scientific collaboration on Web Audio Modules.

**Description:** Research stay and lectures.

## 9.2 International research visitors

### 9.2.1 Visits of international scientists

#### Other international visits to the team

##### Matteo Palmonari

**Status** Associate Professor

**Institution of origin:** University of Milano Bicocca

**Country:** Italy

**Dates:** from 26-06-23 to 25-07-23

**Context of the visit:** Research stay to investigate the application of LLMs to the argument mining task.

**Mobility program/type of mobility:** Research stay funded by 3IA Côte d'Azur

##### Cristhian Figueroa

**Status** Assistant Professor

**Institution of origin:** University del Cauca

**Country:** Colombia

**Dates:** from 23-06-23 to 07-07-23

**Context of the visit:** Research stay to investigate the construction of a knowledge graph from agro-meteorological data related to cafe growing in Colombia.

**Mobility program/type of mobility:** Research stay funded by University del Cauca

### 9.2.2 Visits to international teams

#### Research stays abroad

**Marco Winckler**

**Visited institution:** University of Bari

**Country:** Italy

**Dates:** from 23-10-23 to 19-11-23

**Context of the visit:** Invited Professor by the University of Bari

**Mobility program/type of mobility:** research stay/lecture

### 9.3 European initiatives

#### 9.3.1 H2020 projects

##### AI4Media

**Title:** AI4Media

**Duration:** 2020 - 2024

**Coordinator:** The Centre for Research and Technology Hellas (CERTH)

**Partners:**

*see consortium*

**Inria contact:** through 3IA

**Summary:** AI4Media is a 4-year-long project. Funded under the European Union's Horizon 2020 research and innovation programme, the project aspires to become a Centre of Excellence engaging a wide network of researchers across Europe and beyond, focusing on delivering the next generation of core AI advances and training to serve the Media sector, while ensuring that the European values of ethical and trustworthy AI are embedded in future AI deployments. AI4Media is composed of 30 leading partners in the areas of AI and media (9 Universities, 9 Research Centres, 12 industrial organizations) and a large pool of associate members, that will establish the networking infrastructure to bring together the currently fragmented European AI landscape in the field of media, and foster deeper and long-running interactions between academia and industry.

#### 9.3.2 Other european programs/initiatives

##### HyperAgents - SNSF/ANR project

**Title:** HyperAgents

**Duration:** 2020 - 2024

**Coordinator:** Olivier Boissier, MINES Saint-Étienne

**Partners:**

- MINES Saint-Étienne (FR)
- INRIA (FR)
- Univ. of St. Gallen (HSG, Switzerland)

**Inria contact:** Fabien Gandon

**Summary:** The HyperAgents project, Hypermedia Communities of People and Autonomous Agents, aims to enable the deployment of world-wide hybrid communities of people and autonomous agents on the Web. For this purpose, HyperAgents defines a new class of multi-agent systems that use hypermedia as a general mechanism for uniform interaction. To undertake this investigation, the project consortium brings together internationally recognized researchers actively contributing to research on autonomous agents and MAS, the Web architecture, Semantic Web, and to the standardization of the Web. Project Web site: [HyperAgents Project](#)

#### ANTIDOTE - CHIST-ERA project

**Title:** ANTIDOTE

**Duration:** 2020 - 2024

**Coordinator:** Elena Cabrio, Serena Villata

**Partners:**

- University of the Côte d'Azur (Wimmics Team)
- Fondazione Bruno Kessler (IT)
- University of the Basque Country (ES)
- University of Leuven (Belgium)
- University of Lisbon (PT)

**Summary:** Providing high quality explanations for AI predictions based on machine learning requires to combine several interrelated aspects, including, among others: selecting a proper level of generality/specificity of the explanation, considering assumptions about the familiarity of the explanation beneficiary with the AI task under consideration, referring to specific elements that have contributed to the decision, making use of additional knowledge (e.g. metadata) which might not be part of the prediction process, selecting appropriate examples, providing evidences supporting negative hypothesis, and the capacity to formulate the explanation in a clearly interpretable, and possibly convincing way. According to the above considerations, ANTIDOTE fosters an integrated vision of explainable AI, where low level characteristics of the deep learning process are combined with higher level schemas proper of the human argumentation capacity. ANTIDOTE will exploit cross-disciplinary competences in three areas, i.e. deep learning, argumentation and interactivity, to support a broader and innovative view of explainable AI. Although we envision a general integrated approach to explainable AI, we will focus on a number of deep learning tasks in the medical domain, where the need for high quality explanations, both to clinicians and to patients, is perhaps more critical than in other domains. Project Web site: [Antidote Project](#)

## 9.4 National initiatives

### ANR D2KAB

**Participants:** Olivier Corby, Catherine Faron, Fabien Gandon, Franck Michel, Nadia Yacoubi Ayadi.

Partners: LIRMM, INRAE, IRD, ACTA

D2KAB is an ANR project which started in June 2019 until June 2024, led by the LIRMM laboratory (UMR 5506). Its general objective is to create a framework to turn agronomy and biodiversity data into knowledge –semantically described, interoperable, actionable, open– and investigate scientific methods and tools to exploit this knowledge for applications in science and agriculture. Within this project the Wimmics team is contributing to the lifting of heterogeneous dataset related to agronomy coming from the different partners of the project and is responsible to develop a unique entry point with semantic querying and navigation services providing a unified view on the lifted data.

Web site: [D2KAB Project](#)

## ANR DeKaloG

**Participants:** Olivier Corby, Catherine Faron, Fabien Gandon, Pierre Maillot, Franck Michel.

Partners: Université Nantes, INSA Lyon, Inria Center at Université Côte d'Azur

DeKaloG (Decentralized Knowledge Graphs) is an ANR project until June 2024 that aims to: (1) propose a model to provide fair access policies to KGs without quota while ensuring complete answers to any query. Such property is crucial for enabling web automation, i.e. to allow agents or bots to interact with KGs. Preliminary results on web preemption open such perspective, but scalability issues remain; (2) propose models for capturing different levels of transparency, a method to query them efficiently, and especially, techniques to enable web automation of transparency. (3) propose a sustainable index for achieving the findability principle.

[DeKaloG Web site](#)

## ANR ATTENTION

**Participants:** Serena Villata, Elena Cabrio, Xiaou Wang, Pierpaolo Goffredo.

The ANR project ATTENTION started in January 2022 with Université Paris 1 Sorbonne, CNRS (Centre Maurice Halbwachs), EURECOM, Buster.Ai. The coordinator of the project is CNRS (Laboratoire I3S) in the person of Serena Villata.

In the ATTENTION project, we propose to address the urgent need of designing intelligent semi-automated ways to generate counter-arguments to fight the spread of disinformation online. The idea is to avoid the undesired effects that come with content moderation, such as overblocking, when dealing with disinformation online, and to directly intervene in the discussion (e.g., Twitter threads) with textual arguments that are meant to counter the fake content as soon as possible, and prevent it from further spreading. A counter-argument is a non-aggressive response that offers feedback through fact-bound arguments, and can be considered as the most effective approach to withstand disinformation. Our approach aims at obtaining high quality counter-arguments while reducing efforts and supporting human fact-checkers in their everyday activities.

## ANR CIGALA

**Participants:** Serena Villata, Elena Cabrio, Pierpaolo Goffredo.

The ANR ASTRID CIGALA project (December 2022 - 30 Months) "Controversy and influence in the Ukraine war: a study of argumentation and counter-argumentation through Artificial Intelligence" aims to bring together, in a multi-disciplinary way, the research fields of discourse analysis in the humanities and social sciences, with those of automatic extraction of natural language arguments from text (argument mining) in Artificial Intelligence.

Partners are Ecole de l'Air et de l'Espace (EAE), which is also the coordinator of the project, and CNRS (Laboratoire I3S).

## ANR CROQUIS

**Participants:** Andrea Tettamanzi.

The ANR project CROQUIS (March 2022 - 48 months) with CRIL (Lens) and HSM (Montpellier). The coordinator of the project is Salem Benferhat (CRIL). The local coordinator for Laboratoire I3S is Andrea

Tettamanzi. The local unit involves two other members of I3S which are not part of WIMMICS, namely Célia da Costa Pereira and Claude Pasquier. The contribution of Wimmics is focused on addressing the problem of incomplete and uncertain data.

Web site: [CROQUIS Project](#)

#### ANR AT2TA

**Participants:** Pierre Monnin.

Partners: Université de Lorraine (LORIA), Inria Paris (HeKA team), Université Paul Sabatier (IRIT), IHU Imagine, Université Côte d'Azur (I3S), Infologic

The ANR project AT2TA started in February 2023 until February 2026. The coordinator of the project is Miguel Couceiro (LORIA, Université de Lorraine). The local coordinator for I3S / Wimmics is Pierre Monnin. The project aims to develop an analogy-based machine learning framework and to demonstrate its usefulness in real case scenarios. Within the project, the Wimmics team is contributing by investigating the potential usages of analogy-based frameworks with and for knowledge graphs, and the associated adequate representation spaces.

Web site: [AT2TA project](#)

#### ISSA (AAP Collex-Persée)

**Participants:** Franck Michel, Anna Bobasheva, Olivier Corby, Catherine Faron, Aline Mennin, Marco Winckler.

Partners: CIRAD, Mines d'Alès

The [ISSA project](#) started in October 2020 and is led by the CIRAD. It aims to set up a framework for the semantic indexing of scientific publications with thematic and geographic keywords from terminological resources. It also intends to demonstrate the interest of this approach by developing innovative search and visualization services capable of exploiting this semantic index. Agritrop, Cirad's open publications archive, serves as a use case and proof of concept throughout the project. In this context, the primarily semantic resources are the Agrovoc thesaurus, Wikidata and GeoNames.

Wimmics team is responsible for (1) the generation and publication of the knowledge graph representing the indexed entities, and (2) the development of search/visualization tools intended for researchers and/or information

#### CROBORA: Crossing Borders Archives: understanding the circulation of images of Europe (ANR)

**Participants:** Marco Winckler, Aline Menin.

Coordinator: Matteo Treleani, SicLab, Université Côte d'Azur

The [CROBORA project](#) (ANR-20-CE38-0002) project led by the Sic.Lab laboratory at EUR CREATES, University of Côte d'Azur, funded by ANR from 2021 to 2024. CROBORA studies the circulation of archive images in the media space. The main hypothesis of the project is that what determines the circulation of archives thus constituting the visual memory of the European construction is not a decision that lies solely in the hands of the authors of the reuses (journalists, audiovisual professionals) but rather the consequence of a series of mediations that can be technical (the availability of archives for example), interprofessional (the relationship between archival institutions and the media), cultural (the use of a document for a purpose in one country or another), historical (an archive sequence can change its meaning over time) etc. The general objective of the project is therefore to understand the logics governing the circulation of audiovisual archives. The project aims to respond to the following sub-objectives: 1. to understand which audiovisual fragments are reused in the media to talk about Europe; 2. to build a

cartography of frequently used symbolic images; 3. to analyze the representations carried by these images; 4. to understand their trajectory in order to see how they are reshaped diachronically and according to different media, countries and institutions; and finally 5. to identify which mediations determine their readjustments.

Wimmics team is responsible for (1) the development of visualization tools for supporting the exploration of the CROBORA dataset, and (2) investigating algorithms for optimizing the visual search in the dataset.

## 9.5 Regional initiatives

### 3IA Côte d'Azur

**Participants:** Elena Cabrio, Catherine Faron, Fabien Gandon, Freddy Limpens, Andrea Tettamanzi, Serena Villata.

3IA Côte d'Azur is one of the four “Interdisciplinary Institutes of Artificial Intelligence”<sup>7</sup> that were created in France in 2019. Its ambition is to create an innovative ecosystem that is influential at the local, national, and international level. The 3IA Côte d'Azur institute is led by Université Côte d'Azur in partnership with major higher education and research partners in the region of Nice and Sophia Antipolis: CNRS, Inria, INSERM, EURECOM, ParisTech MINES, and SKEMA Business School. The 3IA Côte d'Azur institute is also supported by ECA, Nice University Hospital Center (CHU Nice), CSTB, CNES, Data Science Tech Institute and INRAE. The project has also secured the support of more than 62 companies and start-ups.

We have four 3IA chairs for tenured researchers of Wimmics and several grants for PhD and postdocs.

We also have an industrial 3IA Affiliate Chair with the company Mnemotix focused on the industrialisation and scalability of the CORESE software.

### Automazoo

**Participants:** Arnaud Barbe, Molka Dhouib, Catherine Faron.

Coordinators: Catherine Faron and Marco Corneli (CEPAM, UCA)

**Automazoo** was a 1-year project funded by the Académie d'Excellence “Homme, Idées et Milieux” of Université Côte d'Aur. Its overall aim was to adopt, adapt and combine methods from automatic language processing, knowledge representation and reasoning, and machine learning to analyze, classify and automate the semantic annotation of ancient texts.

### HISINUM

**Participants:** Arnaud Barbe, Molka Dhouib, Catherine Faron, Andrea Tettamanzi.

Coordinators: Muriel Dal Pont Legrand (GREDEG), Mélanie Plouviez (CRHI), Catherine Faron (I3S) and Arnaud Zucker (CEPAM)

**HISINUM** is a 3-year project funded by the Académie d'Excellence “Homme, Idées et Milieux” of Université Côte d'Aur. The aim of this project is to reflect on how digital humanities are renewing research practices and the issue of data in the humanities and social sciences, and on the epistemological impact of the new tools and their ability to change disciplinary boundaries.

<sup>7</sup>3IA Côte d'Azur Web site

## 10 Dissemination

### 10.1 Promoting scientific activities

#### 10.1.1 Scientific events: organisation

##### General chair, scientific chair

- Catherine Faron: Co-chair of the [1st Workshop on Controlled Vocabularies and Data Platforms for Smart Food Systems \(SmartFood 2023\)](#), colocated with [ER 2023](#), November 6, 2023, Lisbon, Portugal.
- Pierre Monnin: co-general chair of the [Neuro-symbolic AI workshop](#) (supported by GT MHyIA of GDR RADIA and AfiA, and co-located with [ECSQARU 2023](#)), September 19, 2023, Arras, France.
- Andrea Tettamanzi: co-chair of the [1st Workshop on AI-driven heterogeneous data management: Completing, merging, handling inconsistencies and query-answering \(ENIGMA 2023\)](#), Co-located with [KR 2023](#), September 3-4, 2023, Rhodes, Greece

##### Member of the organizing committees

- Elena Cabrio and Serena Villata: Local organizers of the “[Natural Language Argumentation](#)” event (GDR TAL and GDR RADIA), November 27th, 2023 in Sophia Antipolis.
- Aline Menin: PhD Day at the [Extended Reality Research and Creative Center \(XR2C2\)](#), July 7, 2023, Cannes, France.

#### 10.1.2 Scientific events: selection

##### Chair of conference program committees

- Elena Cabrio: Co-chair of Diversity and Inclusion, [The 17th Conference of the European Association of Computational Linguistics \(EACL 2023\)](#).
- Catherine Faron: Program chair of the [French Conference on Knowledge Extraction and Management \(EGC 2023\)](#), January 16-20, 2023, Lyon, France.
- Serena Villata: Co-chair of Tutorial and Workshop, [20th International Conference on Principles of Knowledge Representation and Reasoning, KR 2023](#) September 2-8, 2023, Rhodes, Greece
- Marco Winckler:
  - Technical Program Chair of [the 15th The ACM SIGCHI Symposium on Engineering Interactive Computing Systems \(EICS'2023\)](#), June 27-30, 2023, Swansea, UK.
  - Technical Program Chair of the [INTERACT 2023 - IFIP TC 13 International Conference on Human Computer Interaction](#), August 28th - September 1st 2023, University of York, UK.

##### Member of the conference program committees

- Elena Cabrio: Area Chair of: the 2023 Conference on Empirical Methods in Natural Language Processing (EMNLP), the 61st Annual Meeting of the Association for Computational Linguistics (ACL), ACL ARR Rolling Reviews; Member of the Senior Program Committee of: International Joint Conference on Artificial Intelligence (IJCAI), European Conference on Artificial Intelligence (ECAI), Association for the Advancement of Artificial Intelligence (AAAI) .
- Molka Dhouib: [Ingénierie de connaissances, European Semantic Web Conference](#)
- Catherine Faron: Senior member of Program Committees of International Semantic Web Conference (ISWC) 2023 and European Semantic Web Conference (ISWC) 2023 (resource track), member of Program Committees of European Semantic Web Conference (ISWC) 2023 (research track), Semantics 2023, FOIS 2023, ER 2023, SWODCH 2023

- Fabien Gandon: member of Program Committees of European Conference on Artificial Intelligence (ECAI) 2023, European Semantic Web Conference (ISWC) 2023, International Joint Conference on Artificial Intelligence (IJCAI) 2023, International Semantic Web Conference (ISWC) 2023 and TheWebConf 2023
- Alain Giboin: Member of the scientific committee of the MODACT 2023 conference (11-13 May 2023, Paris) on the modeling of human activity
- Pierre Monnin: member of Program Committees of International Semantic Web Conference (ISWC) 2023, ECML-PKDD 2023, TheWebConf 2024
- Aline Menin: [IEEE International Conference on Virtual Reality and 3D User Interfaces](#), [Eurographics Conference on Visualization](#), [ACM SIGCHI Symposium on Engineering Interactive Computing Systems](#)
- Franck Michel: member of Program Committees of Semantics 2023, TheWebConf 2024
- Andrea Tettamanzi: GECCO, ICAAWI-IAT'23.
- Serena Villata: Area Chair of: the 2023 Conference on Empirical Methods in Natural Language Processing (EMNLP), the 61st Annual Meeting of the Association for Computational Linguistics (ACL), ACL ARR Rolling Reviews; Area chair of: International Joint Conference on Artificial Intelligence (IJCAI), European Conference on Artificial Intelligence (ECAI), Association for the Advancement of Artificial Intelligence (AAAI) .
- Marco Winckler: member of the Program Committees of INTERACT 2023, ACM EICS 2023, EuroVis 2023, ICWE 2023, IS-EUD 2023, ACM IUI 2023, WISE 2023, SVR 2023.

### 10.1.3 Journal

#### Member of the editorial boards

- Elena Cabrio: Italian Journal of Computational Linguistics (IJCoL ISSN 2499-4553).
- Catherine Faron: [Transactions On Graph Data and Knowledge \(TGDK\)](#), [Data and Knowledge Engineering \(DKE\)](#), [Revue Ouverte d'Intelligence Artificielle \(ROIA\)](#)
- Pierre Monnin: [Transactions On Graph Data and Knowledge \(TGDK\)](#)
- Célian Ringwald: Editor at [Programming Historian](#).
- Serena Villata: [Artificial Intelligence and Law](#), [Argument and Computation](#) and [Journal of Web Semantics](#)
- Marco Winckler: [Journal of Web Engineering \(River Publishers\)](#), [Interacting with Computers \(Oxford Press\)](#), [Behaviour and Information Technology \(Taylor and Francis\)](#), [PACM Proceedings on Human-Computer Interaction \(ACM Sheridam\)](#), [IFIP Advances in Information and Communication Technology \(Springer\)](#).

#### Reviewer - reviewing activities

- Catherine Faron: [Journal of Web Semantics \(JOWS\)](#), [ACM Journal on Computing and Cultural Heritage \(JOCCH\)](#)
- Fabien Gandon: [ACM Transactions on the Web](#), [Springer Heritage Science](#), [Springer Nature Journal](#)
- Pierre Monnin: [Semantic Web Journal](#), [Journal of Web Semantics](#), [International Journal of Approximate Reasoning](#), [Annals of Mathematics and Artificial Intelligence](#), [AI Communications](#), [Communications Medicine](#), [SoftwareX](#)
- Aline Menin: [Journal on Virtual Reality](#)



- Franck Michel: [Biodiversity Data Journal](#), [Data and Knowledge Engineering](#)
- Andrea Tettamanzi: Artificial Intelligence.
- Marco Winckler: [SoftwareX](#), [International Journal of Human-Computer Interaction](#), [Information and Software Technology](#), [Personal and Ubiquitous Computing](#)
- Molka Dhouib: [Data and Knowledge Engineering](#)

#### 10.1.4 Invited talks

- Elena Cabrio: “Processing Natural Language to Extract, Analyze and Generate Arguments from Texts”. [SEPLN 2023. 39th International Conference of the Spanish Society for Natural Language Processing](#). Jaen (Spain), September 2023.
- Catherine Faron: “Description et interrogation de l'écosystème de graphes de connaissances D2KAB ” at [Séminaire résidentiel INRAE Semantic Linked Data](#), October 10, 2023, Agde, France.
- Fabien Gandon
  - “[Walking our Way to the Web](#)” at the Conference on the Web: Scientific Creativity, Technological Innovation and Society, XXVIII Conference on Contemporary Philosophy and Methodology of Science, March 9-10, 2023, University of A Coruña (Spain)
  - “Weaving a Web of Augmented Intelligence.”, Keynote Computer Science Insight (CSI) Talk University St. Gallen, June 29, 2023
  - Panel “[Conférence IA et Règles européennes](#)”, November 14, 2023, Palais des Rois Sardes, Nice
- Pierre Monnin
  - Neuro-Symbolic Approaches for the Knowledge Graph Lifecycle. [MALOTEC seminar](#), LORIA, December 18, 2023, Nancy, France.
  - Knowledge Graphs and Analogical Reasoning: from the Example of Zero-Shot KG Bootstrapping to Perspectives. [Workshop on analogies: from learning to explainability](#) (supported by GT MHyIA & EXPLICON – GdR RADIA), November 27, 2023, Arras, France.
  - “Using Wikidata to bootstrap a new knowledge graph: selecting relevant entities using analogical pruning” at the [Wikidata Modelling Days 2023](#), Virtual Event, November 30, December 1-2, 2023. [Recording](#).
- Aline Menin: “Visualization of spatio-temporal data: challenges for user perception and representation.” at the [IVU Laboratory seminar](#). November 20, 2023, Bari, Italy.
- Marco Winckler: “Information Visualisation and Analytical Provenance: the missing gaps.” at the [IVU Laboratory seminar](#). October 6, 2023, Bari, Italy.
- Molka Dhouib: “Construction d'un graphe de connaissance à partir des annotations manuelles de textes de zoologie antique” at [Symposium MaDICS, action RoCED](#). May 24, 2023, Troyes, France.
- Serena Villata
  - “Formal and Natural Arguments for Effective Explanations” at the 33rd Int. Conference on Automated Planning and Scheduling (ICAPS-2023), July 2023, Prague.
  - “Artificial Argumentation for Humans” at the franco-japanese conference "Les nouveaux chemins de l'art et de la culture", October 3, 2023, Tokyo, Japan.
  - “Argumentation and Dialogue” at the Ecole d'Automne IA2 in AI and Management of Heterogeneous Information and Data, Octobre 30 - Novembre 3, 2023 Le Lazaret, Sète

### 10.1.5 Leadership within the scientific community

- Catherine Faron: participant [Dagstuhl Seminar, 23081 Agents on the Web 2023-02-19 to 2023-02-24](#)
- Fabien Gandon: participant [Dagstuhl Seminar, 23081 Agents on the Web 2023-02-19 to 2023-02-24](#)
- Marco Winckler: Steering Committee Chair for the [IFIP TC13 INTERACT Conference series](#).

### 10.1.6 Scientific expertise

- Elena Cabrio: Member of the Comité du suivi doctoral, INRIA Sophia Antipolis; Member of the Bureau of the Académie 1 of IDEX UCA JEDI; Member of the Conseil d'Administration (CA) of the French Association of Computational Linguistics (ATALA).
- Catherine Faron: Member of the HCERES committee in charge of evaluating Paris-Cité training offer; expert for the European commission (MSCA 2023 and CL4 calls); expert for Institut Français de Bioinformatique (ELIXIR node)
- Fabien Gandon: European Science Foundation (ESF): Evaluation of a FWO research project, Evaluation of a FWO post-doctoral project; Reviewer PEPR Santé Numérique.
- Serena Villata: Member of the ANR committee to evaluate the research projects on AI (CE23), Member of the Bureau of CEP of Inria SAM.
- Marco Winckler: Member of the Comité du suivi doctoral (Paris Saclay, UFSCar).

### 10.1.7 Research administration

- Fabien Gandon : Leader of the Wimmics team ; co-president of scientific and pedagogical council of the Data Science Technical Institute ([DSTI](#)) W3C Advisory Committee Representative (AC Rep) for Inria.
- Serena Villata: Deputy scientific director of 3IA Côte d'Azur.
- Marco Winckler : Leader of the SPARKS team of the CNRS laboratory I3S (UMR 7271).

### 10.1.8 Promoting Open Science practices

Too often, the methods described in research papers are not reproducible because the code and/or data are simply not provided. As a result, it is hardly possible to verify the results and build upon these works. The Open Science movement is meant to fix this by fostering the unhindered spreading of the results, methods and products of scientific research. It is based on the open access to publications, data and source codes.

To make the team members aware of these issues, in 2023 we gave a 2-hour presentation on the principles of Open Science, the goals of experiment reproducibility, with a focus on practical approaches meant to make codes and data findable (using metadata), accessible (public repositories and long-time preservation), referenceable (point to a specific version) and citable (give credit, attribution), as well as good practices to cite others' codes and data.

Franck Michel. [article+code+data]: A virtuous tryptic towards reproducible research. 2023. [Slides](#).

## 10.2 Teaching - Supervision - Juries

### 10.2.1 Teaching

**Participants:** Michel Buffa, Elena Cabrio, Olivier Corby, Catherine Faron, Fabien Gandon, Aline Menin, Amaya Nogales Gómez, Andrea Tettamanzi, Serena Villata, Marco Winckler, Molka Dhoub, Benjamin Molinet, Célian Ringwald, Pierre Monnin, Anaïs Ollagnier.

- Michel Buffa:
  - Licence 3, Master 1, Master 2 Méthodes Informatiques Appliquées à la Gestion des Entreprises (MIAGE) : Web Technologies, Web Components, etc. 192h.
  - DS4H Masters 3D games programming on Web, JavaScript Introduction: 40h.
- Olivier Corby:
  - Licence 3 IOTA UniCA 25 hours Semantic Web
  - Licence 3 IA DS4H UniCA 25 hours Semantic Web
- Catherine Faron :
  - Master 2/5A SI PNS: Web of Data, 32 h
  - Master 2/5A SI PNS: Semantic Web 32h
  - Master 2/5A SI PNS: Ingénierie des connaissances 15h
  - Master DSAI UniCA: Web of Data, 30h
  - Master 1/4A SI PNS and Master2 IMAFA/5A MAM PNS: Web languages, 28h
  - Licence 3/3A SI PNS and Master 1/4A MAM PNS: Relational Databases, 60h
  - Master [Data ScienceTech Institute \(DSTI\)](#): Data pipeline, 50h.
- Fabien Gandon :
  - Master: Integrating Semantic Web technologies in Data Science developments, 72 h, M2, [Data ScienceTech Institute \(DSTI\)](#), France.
  - Tutorial Inria Academy - Inria Chile, “CORESE and Semantic Web Standards”, 1h, 10/10/2023
  - Tutorial Inria Academy, “CORESE and Semantic Web Standards”, 6h, 28/11/2023
  - Inria Academy, “Introduction to CORESE”, Tutorial at [World AI Cannes Festival \(WAICF\)](#), 10/02/2023
  - “Introduction à l’Intelligence Artificielle - Représentation et traitement de connaissances”, Catherine Faron, Fabien Gandon, 3.5 hours, EFELIA Côte d’Azur, 3IA Côte d’Azur, 2/11/2023
- Aline Menin :
  - Master 2, Data Visualization, 7h éq. TD (CM/TD), UniCA, MBDS DS4H, France.
  - Polytech 5ème année, UniCA, 13.5h (CM/TP), Data visualization.
  - BUT 2, IUT Nice Côte d’Azur, 160h éq. TD (CM/TD), “Développement efficace”, “Qualité de développement”, and “Développement des Applications avec IHM”.
- Molinet Benjamin:
  - Fabron BUT-1 - Intro NLP - 42h TD.
  - EMSI Casablanca, Master IA2 - Natural Language Processing, 20h CM/TD
  - Master I Computer Science, Text Processing in AI, 2 hours.
- Amaya Nogales Gómez:
  - Master 1, Data Sciences & Artificial Intelligence, UniCA, 20h (CM/TD), Security and Ethical Aspects of Data.
  - Licence 2, Licence Informatique, UniCA, 36h (TP), Structures de données et programmation C.
- Serena Villata:

- Master II Droit de la Création et du Numérique - Sorbonne University: Approche de l'Elaboration et du Fonctionnement des Logiciels, 15 hours (CM), 20 students.
- Master 2 MIAGE IA - University Côte d'Azur: I.A. et Langage : Traitement automatique du langage naturel, 28 hours (CM+TP), 30 students.
- DUT IA et santé. Natural Language Processing, 4 hours.
- Elena Cabrio:
  - Master I Computer Science, Text Processing in AI, 28 hours.
  - Master 2 MIAGE IA: I.A. et Langage : Traitement automatique du langage naturel, 15 hours.
  - Master 1 EUR CREATES, Parcours Linguistique, traitements informatiques du texte et processus cognitifs. Introduction to Computational Linguistics, 30 hours.
  - DUT IA et santé. Natural Language Processing, 4 hours.
  - Licence 3 Sciences et Technologies: parcours Intelligence Artificielle, Natural Language Processing, 30 hours.
  - Licence 2 IUT, Introduction to AI, 30 hours.
  - Licence 1 IUT, Introduction to Database and SQL, 133 hours.
- Andrea Tettamanzi
  - Licence: Introduction à l'Intelligence Artificielle, 45 h ETD, L2, UniCA, France.
  - Master: Logic for AI, 30 h ETD, M1, UniCA, France.
  - Master: Web, 30 h ETD, M1, UniCA, France.
  - Master: Algorithmes Évolutionnaires, 24.5 h ETD, M2, UniCA, France.
  - Master: Modélisation del l'Incertitude, 24.5 h ETD, M2, UniCA, France.
- Marco Winckler
  - Licence 3: Event-driven programming, 45 h ETD, UniCA, Polytech Nice, France.
  - Master 1: Methods and tools for technical and scientific writing, Master DSAI, 15 h ETD, UniCA, DS4H, France.
  - Master 2: Introduction to Scientific Research, 15 h ETD, UniCA, Polytech Nice, France.
  - Master 2: Information Visualisation, 34 h ETD, UniCA, Polytech Nice, France.
  - Master 2: Data Visualization, 15 h ETD, UniCA, MBDS DS4H, France.
  - Master 2: Design of Interactive Systems, 34 ETD, UniCA, Polytech Nice, France.
  - Master 2: Evaluation of Interactive Systems, 34 ETD, UniCA, Polytech Nice, France.
  - Master 2: Multimodal Interaction Techniques, 15 ETD, UniCA, Polytech Nice, France.
  - Master 2: coordination of the TER (Travaux de Fin d'Etude), UniCA, Polytech Nice, France.
  - Master 2: coordination of the track on Human-Computer Interaction at the Informatics Department, UniCA, Polytech Nice, France.
- Molka Dhoub
  - Licence 3/3A SI PNS: Relational Databases, 32,5h (TD).
  - Master 1/4A SI PNS: Web languages, 10h (TD)
  - Licence 3/LPI: Introduction to Web of Data and Semantic Web, 16h (TD)
- Célian Ringwald:
  - Licence 3/3A SI PNS: XML technologies, 10h (TD).
  - Licence 3/3A SI PNS: Bases de données relationnelles, 30h (TD).

- 1ere Année BUT : Introduction bases de données relationnelles, 40h (TD).
- Pierre Monnin:
  - Master 1/Applied Foreign Languages: Introduction to Artificial Intelligence. 12h CM, 12h TD
  - Master 1/Adult Education: Introduction to Artificial Intelligence. 12h CM
  - Polytech Nice 5th year / EUR DS4H Master 2: Machine Learning & Semantic Web. 1h CM, 1.5h TD
  - Polytech Nice 5th year / EUR DS4H Master 2: Machine Learning tutoring. 10h TD
  - TELECOM Nancy 3rd year apprenticeship students: Artificial Intelligence. 10h CM, 8h TD
- Anaïs Ollagnier:
  - Licence 2 IUT, Introduction to AI, 48h TD.
  - Master 1/2/Phd EUR LEXSOCIETE. Introduction to AI Applied to Law and Implications for Administration and Public Service. 10h CM
  - Master 1/2/Phd EUR LEXSOCIETE. Introduction to AI Applied to Law. 12h CM, 8h TD
  - Master 1/2/Phd EUR ELMI. Artificial Intelligence and societal tranformation. 10h CM
- Arnaud Barbe:
  - 1ere Année BUT : Introduction aux bases de données relationnelles et SQL, 40h TD.

### E-learning

- Mooc: Fabien Gandon, Olivier Corby & Catherine Faron, Web of Data and Semantic Web (FR), 7 weeks, [FUN](#), Inria, France Université Numérique, self-paced course 41002, Education for Adults, 17496 learners registered at the time of this report and 855 certificates/badges, [MOOC page](#).
- Mooc: Fabien Gandon, Olivier Corby & Catherine Faron, Introduction to a Web of Linked Data (EN), 4 weeks, [FUN](#), Inria, France Université Numérique, self-paced course 41013, Education for Adults, 5952 learners registered at the time of this report, [MOOC page](#).
- Mooc: Fabien Gandon, Olivier Corby & Catherine Faron, Web of Data (EN), 4 weeks, [Coursera](#), self-paced course Education for Adults, 5134 learners registered at the time of this report, [MOOC page](#).
- Mooc: Michel Buffa, HTML5 coding essentials and best practices, 6 weeks, [edX MIT/Harvard](#), self-paced course Education for Adults, more than 500k learners at the time of this report (2015-2022), [MOOC page](#).
- Mooc: Michel Buffa, HTML5 Apps and Games, 5 weeks, [edX MIT/Harvard](#), self-paced course Education for Adults, more than 150k learners at the time of this report (2015-2022), [MOOC page](#).
- Mooc: Michel Buffa, JavaScript Introduction, 5 weeks, [edX MIT/Harvard](#), self-paced course Education for Adults, more than 250k learners at the time of this report (2015-2022), [MOOC page](#).

### 10.2.2 Supervision

#### PhDs

- PhD in progress: Ali Ballout, *Active Learning for Axiom Discovery*, Supervised by Andrea Tettamanzi, UniCA.
- PhD in progress: Lucie Cadorel, *Localisation sur le territoire et prise en compte de l'incertitude lors de l'extraction des caractéristiques de biens immobiliers à partir d'annonces*, Supervised by Andrea Tettamanzi, UniCA.

- PhD in progress: Rony Dupuy Charles, *Combinaison d'approches symboliques et connexionnistes d'apprentissage automatique pour les nouvelles méthodes de recherche et développement en agro-végétale-environnement*, Supervised by Andrea Tettamanzi, UniCA.
- PhD in progress: Rémi Felin, *Découverte évolutive d'axiomes à partir de graphes de connaissances*, UniCA, Co-supervised by Andrea Tettamanzi, Catherine Faron.
- PhD in progress: Pierpaolo Goffredo, *Fallacious Argumentation in Political Debates*, UniCA 3IA, Co-supervised by Elena Cabrio and Serena Villata.
- PhD defended: Santiago Marro, *Argument-based Explanatory Dialogues for Medicine*, UniCA 3IA, Co-supervised by Elena Cabrio and Serena Villata.
- PhD in progress: Benjamin Molinet, *Explanatory argument generation for healthcare applications*, UniCA 3IA, Co-supervised by Elena Cabrio and Serena Villata.
- PhD in progress: Benjamin Ocampo, *Subtle and Implicit Hate Speech Detection*, UniCA 3IA, Co-supervised by Elena Cabrio and Serena Villata.
- PhD in progress: Clément Quere. *Immersive Visualization Techniques for spatial-temporal data*. Co-supervised by Aline Menin, Hui-Yin Wu, and Marco Winckler.
- PhD in progress: Célian Ringwald, *Learning RDF pattern extractors for a language from dual bases Wikipedia/LOD*, Université Côte d'Azur, Co-supervised by Fabien Gandon, Catherine Faron, Franck Michel, Hanna Abi-Akl.
- PhD in progress: Florent Robert, *Analyzing and Understanding Embodied Interactions in Extended Reality Systems*. Co-supervised by Hui-Yin Wu, Lucile Sassatelli, and Marco Winckler.
- PhD in progress: Maroua Tikat, *Visualisation multimédia interactive pour l'exploration d'une base de métadonnées multidimensionnelle de musiques populaires*. Co-supervised by Michel Buffa, Marco Winckler.
- PhD in progress: Xiaoou Wang, *Counter-argumentation generation to fight online disinformation*, UniCA, Co-supervised by Elena Cabrio and Serena Villata.

### Internships and Apprenticeships

- Apprentice (IUT3) Hugo Carton, "Data visualization". Co-supervised by Arnaud Barbe, Catherine Faron, Aline Menin, and Fabien Gandon.
- Apprentice (Polytech, M2) Antoine Vidal-Mazuy, "Conception et implémentation d'un hôte de plugins audio dans le web". Co-supervised by Michel Buffa and Marco Winckler.
- Internship (Bachelor MIT) Emma Tysinger "An AI agent to query chemistry knowledge graph". Co-supervised by Louis-Félix Nothias and Fabien Gandon.
- Internship (Aix Ynov Campus, M2) Thomas Mac Vicar, "Etude et réalisation d'un outil pour la conservation, archivage et exposition des projets en réalités étendues". Co-supervised by Aline Menin and Marco Winckler.
- Internship (Polytech, M1) Aimane Elfiguigui, "L'usage des glyphes polymorphes pour assister l'exploration des données spatio-temporelles". Co-supervised by Aline Menin and Marco Winckler.
- Internship (Polytech, M1) Nicolas Audoux, "Exploration visuelle de données météorologiques". Co-supervised by Catherine Faron and Nadia Yacoubi.
- Internship (Polytech, M2) Raphaël Julien. "Creating semantic representations of the real world in augmented reality: HandyNotes". Co-supervised by Marco Winckler and Hui-Yin Wu.

- Internship (CREATES, M1) Irina Gokhaeva, “Cyberharcèlement et discours de haine implicite : construction, annotation et analyse d’un corpus extrait des réseaux sociaux”. Co-supervised by Anaïs Ollagnier and Elena Cabrio.

#### Master Projects (TER/PER Projet d’Etude et de Recherches Polytech)

- Master 2 PER, Polytech: Lynda Attouch. Co-supervised by Catherine Faron, Molka Dhouib and Pierre Monnin
- Master 2 PER, Polytech: Guillaume Méroué. Supervised by Pierre Monnin
- Master 2 PER, Polytech: Theo Jeannes, Quentin Bourdeau, Ambre Correia. Co-supervised by Aline Menin and Marco Winckler
- Master 2 PER, Polytech: Christophe Ruiz and Enzo Daval. Co-supervised by Aline Menin and Marco Winckler
- Master 2 PER, Polytech: Pauline Devictor, Kilian Bonnet, Joel Dibasso, Habib Belmokhtar. Co-supervised by Marco Winckler, Clément Quéré, and Hui-Yin Wu.
- Master 2 PER, Polytech: Hadyl Ayari. Supervised by Marco Winckler, Maroua Tikat, Catherine Faron.
- Master 1 Research Project, DS4H: Ekaterina Kostyrykina. Co-supervised by Molka Dhouib, Catherine Faron
- Master 1 Research Project, DS4H: Oumayma Missaoui. Co-supervised by Molka Dhouib, Catherine Faron
- Master 1 Research Project, DS4H: Karl Alwyn. Co-supervised by Catherine Faron and Franck Michel
- Master 1 Research Project, DS4H: Bachir Benna. Co-supervised by Catherine Faron and Franck Michel
- Master 1 Research Project, DS4H: Steven Essam. Co-supervised by Molka Dhouib, Catherine Faron
- Master 1 Research Project, DS4H: Hamza Hamdaoui. Co-supervised by Aline Menin and Marco Winckler
- Master 1 Research Project, DS4H: Hadyl Ayari. Co-supervised by Aline Menin and Marco Winckler.
- Master 1 Research Project, DS4H: Imane El Mountasser. Supervised by Anaïs Ollagnier
- Master 1 Research Project, DS4H: Ahmed Muhammad. Supervised by Anaïs Ollagnier

#### 10.2.3 Juries

- Catherine Faron
  - PhD jury member Slimane Makhlof, “Apprendre à enchérir : prédiction d’évènements rares et choix de stratégie”, Université Paris Cité, April 19, 2023.
  - PhD Comité de Suivi Individuel (CSI) member for Guilherme Santos (U. Toulouse), Nour Matta (UTT), and Maroua Tikat (UniCA)
- Fabien Gandon
  - Reviewer HDR Aurélien Béné, “Technologies logicielles pour l’instrumentation du travail intellectuel - Une informatique au service du sens”, Université de Technologie de Troyes, 21/06/2023
  - Reviewer HDR Sabrina Kirrane, “Web Knowledge Governance: Legal Knowledge Representation and Automated Compliance Checking”, WU Vienna University, 20/11/2023

- PhD Comité de Suivi Individuel (CSI) member for Pierpaolo Goffredo and Kevin Mottin
- Marco Winckler
  - PhD Jury member, “An Architecture for Music Recommendation Systems Based on Context of Interaction and User Experience”. December 20, 2023. Universidade Federal de São Carlos, Brazil.
  - PhD Comité de Suivi Individuel (CSI) member for Maylon PIRES MACEDO, Eliezer Emanuel BERNART, and Théo BOUGANIM.
- Anaïs Ollagnier, jury de l’édition 2023 du prix de la thèse ATALA décerné conjointement à Tobias Mayer (I3S, Université Côte d’Azur, Sophia Antipolis) et Thimothée Mickus (ATILE, Université de Lorraine/CNRS, Nancy).

## 10.3 Popularization

### 10.3.1 Articles and contents

- Fabien Gandon: Interview pour “Comment l’IA bouleverse l’art”, Beaux Arts Magazine, May 2023 P.1, 52-56,
- Article in the ERCIM News 133: “[IndeGx: An Index of Linked Open Datasets on the Web](#)” by Pierre Maillot, Catherine Faron, Fabien Gandon and Franck Michel (Inria)

### 10.3.2 Education

- Anaïs Ollagnier
  - “Désinformation et propagation des rumeurs sur les réseaux sociaux”. Lycée Dumont d’Urville, December 6, 2023, Toulon, France.
  - “Validation et diffusion des données scientifiques”. Collège l’Eganaude, April 14, 2023, Biot, France.

### 10.3.3 Interventions

- Fabien Gandon: performed 3 sessions “Chiche!” the 24th of January 2023 and 4 other sessions the 8th of December 2023
- Anaïs Ollagnier: “IA et évaluation, outil de triche ou/et de correction?”. Dans le cadre de la Journée Pédagogique de l’EUR Arts et Humanités 2023

## 11 Scientific production

### 11.1 Major publications

- [1] D. Allemang, J. Hendler and F. Gandon. *Semantic Web for the Working Ontologist*. 3. ACM, June 2020. DOI: [10.1145/3382097](https://doi.org/10.1145/3382097). URL: <https://hal.inria.fr/hal-02939606>.
- [2] A. Ben othmane. ‘CARS - A multi-agent framework to support the decision making in uncertain spatio-temporal real-world applications’. Université Côte d’Azur, 12th Oct. 2017. URL: <https://hal.archives-ouvertes.fr/tel-01680984>.
- [3] F. Berthelon. ‘Emotion modelization and detection from expressive and contextual data’. Université Nice Sophia Antipolis, 16th Dec. 2013. URL: <https://tel.archives-ouvertes.fr/tel-00917416>.
- [4] K. R. Bouzidi. ‘Semantic web models to support the creation of technical regulatory documents in building industry’. Université Nice Sophia Antipolis, 11th Sept. 2013. URL: <https://tel.archives-ouvertes.fr/tel-00876366>.



- [5] E. Cabrio. 'Artificial Intelligence to Extract, Analyze and Generate Knowledge and Arguments from Texts to Support Informed Interaction and Decision Making'. Habilitation à diriger des recherches. Université Côte d'Azur, Oct. 2020. URL: <https://hal.inria.fr/tel-03084380>.
- [6] L. Costabello. 'Context-aware access control and presentation of linked data'. Université Nice Sophia Antipolis, 29th Nov. 2013. URL: <https://tel.archives-ouvertes.fr/tel-00934617>.
- [7] P. F. Diallo. 'Sociocultural and temporal aspects in ontologies dedicated to virtual communities'. COMUE Université Côte d'Azur (2015 - 2019); Université de Saint-Louis (Sénégal), 16th Sept. 2016. URL: <https://tel.archives-ouvertes.fr/tel-01402394>.
- [8] A. E. A. Djebri. 'Uncertainty Management for Linked Data Reliability on the Semantic Web'. Université Côte D'Azur, 24th Feb. 2022. URL: <https://hal.archives-ouvertes.fr/tel-03679118>.
- [9] A. Ettorre. 'Towards an interpretable model of learners in a learning environment based on knowledge graphs'. Université Côte d'Azur, 28th Nov. 2022. URL: <https://theses.hal.science/tel-04060324>.
- [10] M. Fell. 'Natural language processing for music information retrieval : deep analysis of lyrics structure and content'. Theses. Université Côte d'Azur, May 2020. URL: <https://tel.archives-ouvertes.fr/tel-02587910>.
- [11] F. Gandon. 'Distributed Artificial Intelligence And Knowledge Management: Ontologies And Multi-Agent Systems For A Corporate Semantic Web'. Université Nice Sophia Antipolis, 7th Nov. 2002. URL: <https://tel.archives-ouvertes.fr/tel-00378201>.
- [12] R. Gazzotti. 'Knowledge graphs based extension of patients' files to predict hospitalization'. Theses. Université Côte d'Azur, Apr. 2020. URL: <https://hal.archives-ouvertes.fr/tel-02959420>.
- [13] P. Goffredo, E. Cabrio, S. Villata, S. Haddadan and J. Torres Sanchez. 'DISPUTool 2.0: A Modular Architecture for Multi-Layer Argumentative Analysis of Political Debates'. In: *Proceedings of the AAAI Conference on Artificial Intelligence*. AAAI Conference on Artificial Intelligence. Vol. 37. 13. Washington, DC, United States, 26th June 2023, pp. 16431–16433. DOI: [10.1609/aaai.v37i13.27069](https://doi.org/10.1609/aaai.v37i13.27069). URL: <https://hal.science/hal-04349796>.
- [14] P. Goffredo, M. Espinoza, S. Villata and E. Cabrio. 'Argument-based Detection and Classification of Fallacies in Political Debates'. In: *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing*. SINGAPORE, Singapore: Association for Computational Linguistics, Dec. 2023, pp. 11101–11112. DOI: [10.18653/v1/2023.emnlp-main.684](https://doi.org/10.18653/v1/2023.emnlp-main.684). URL: <https://hal.science/hal-04260221>.
- [15] P. Goffredo, S. Haddadan, V. Vorakitphan, E. Cabrio and S. Villata. 'Fallacious Argument Classification in Political Debates'. In: *Thirty-First International Joint Conference on Artificial Intelligence IJCAI-22*. Vienna, Austria: International Joint Conferences on Artificial Intelligence Organization, 23rd July 2022, pp. 4143–4149. DOI: [10.24963/ijcai.2022/575](https://doi.org/10.24963/ijcai.2022/575). URL: <https://hal.science/hal-03873412>.
- [16] N. Halliwell. 'Evaluating and improving explanation quality of graph neural network link prediction on knowledge graphs'. Université Côte d'Azur, 23rd Nov. 2022. URL: <https://theses.hal.science/tel-03907696>.
- [17] R. Hasan. 'Predicting query performance and explaining results to assist Linked Data consumption'. Université Nice Sophia Antipolis, 4th Nov. 2014. URL: <https://tel.archives-ouvertes.fr/tel-01127124>.
- [18] M. Lefrançois. 'Meaning-Text Theory lexical semantic knowledge representation : conceptualization, representation, and operationalization of lexicographic definitions'. Université Nice Sophia Antipolis, 24th June 2014. URL: <https://tel.archives-ouvertes.fr/tel-01071945>.
- [19] A. Macina. 'SPARQL distributed query processing over linked data'. COMUE Université Côte d'Azur (2015 - 2019), 17th Dec. 2018. URL: <https://tel.archives-ouvertes.fr/tel-02340700>.
- [20] N. Marie. 'Linked data based exploratory search'. Université Nice Sophia Antipolis, 12th Dec. 2014. URL: <https://tel.archives-ouvertes.fr/tel-01130622>.

- [21] T. Mayer. ‘Argument Mining on Clinical Trials’. Université Côte d’Azur, 16th Dec. 2020. URL: <https://tel.archives-ouvertes.fr/tel-03209489>.
- [22] Z. Meng. ‘Temporal and semantic analysis of richly typed social networks from user-generated content sites on the web’. Université Côte d’Azur, 7th Nov. 2016. URL: <https://hal.inria.fr/tel-01402612>.
- [23] F. Michel. ‘Integrating heterogeneous data sources in the Web of data’. Université Côte d’Azur, 3rd Mar. 2017. URL: <https://tel.archives-ouvertes.fr/tel-01508602>.
- [24] F. Michel, F. Gandon, V. Ah-Kane, A. Bobasheva, E. Cabrio, O. Corby, R. Gazzotti, A. Giboin, S. Marro, T. Mayer, M. Simon, S. Villata and M. Winckler. ‘Covid-on-the-Web: Knowledge Graph and Services to Advance COVID-19 Research’. In: ISWC 2020 - 19th International Semantic Web Conference. Athens / Virtual, Greece, 1st Nov. 2020. DOI: [10.1007/978-3-030-62466-8\\_19](https://doi.org/10.1007/978-3-030-62466-8_19). URL: <https://hal.archives-ouvertes.fr/hal-02939363>.
- [25] T. H. Nguyen. ‘Mining the semantic Web for OWL axioms’. Université Côte d’Azur, 2nd July 2021. URL: <https://tel.archives-ouvertes.fr/tel-03406784>.
- [26] C. Pasquier, C. Da Costa Pereira and A. G. B. Tettamanzi. ‘Extending a Fuzzy Polarity Propagation Method for Multi-Domain Sentiment Analysis with Word Embedding and POS Tagging’. In: *Frontiers in Artificial Intelligence and Applications*. ECAI 2020 - 24th European Conference on Artificial Intelligence. Vol. 325. Santiago de Compostela, Spain: IOS Press, 29th Aug. 2020, pp. 2140–2147. DOI: [10.3233/FAIA200338](https://doi.org/10.3233/FAIA200338). URL: <https://hal.inria.fr/hal-02936130>.
- [27] T. A. Pham. ‘OntoApp : a declarative approach for software reuse and simulation in early stage of software development life cycle’. Université Côte d’Azur, 22nd Sept. 2017. URL: <https://tel.archives-ouvertes.fr/tel-01680766>.
- [28] O. Seye. ‘Sharing and reusing rules for the Web of data’. Université Nice Sophia Antipolis; Université Gaston Berger de Saint Louis, 15th Dec. 2014. URL: <https://hal.inria.fr/tel-01096306>.
- [29] M. Tounsi Dhouib. ‘Knowledge engineering in the sourcing domain for the recommendation of providers’. Université Côte d’Azur, 26th Mar. 2021. URL: <https://tel.archives-ouvertes.fr/tel-03336353>.
- [30] M. Toure. ‘Local peer-to-peer mobile access to linked data in resource-constrained networks’. Université Côte d’Azur; Université de Saint-Louis (Sénégal), 4th Oct. 2021. URL: <https://tel.archives-ouvertes.fr/tel-03514271>.
- [31] D. M. Tran. ‘Discovering multi-relational association rules from ontological knowledge bases to enrich ontologies’. Université Côte d’Azur; Université de Danang (Vietnam), 23rd July 2018. URL: <https://tel.archives-ouvertes.fr/tel-01926812>.

## 11.2 Publications of the year

### International journals

- [32] F. Amardeilh, S. Aubin, S. Bernard, C. Faron, S. Bravo, R. Bossy, F. Michel, J. Raphel and C. Roussey. ‘Combining different points of view on plant descriptions: mapping agricultural plant roles and biological taxa’. In: *Frontiers in Artificial Intelligence*. Frontiers in Artificial Intelligence 6 (2023), p. 118803. DOI: [10.3389/frai.2023.1188036](https://doi.org/10.3389/frai.2023.1188036). URL: <https://hal.science/hal-04289970>.
- [33] M. Buffa and J. Lebrun. ‘Rocking the Web with Browser-based Simulations of Tube Guitar Amplifiers’. In: *Journal of the Audio Engineering Society* 71.11 (16th Nov. 2023), pp. 753–768. URL: <https://inria.hal.science/hal-04335875>.
- [34] J.-P. Calbimonte, A. Ciortea, T. Kampik, S. Mayer, T. R. Payne, V. Tamma and A. Zimmermann. ‘Autonomy in the Age of Knowledge Graphs: Vision and Challenges’. In: *Transactions on Graph Data and Knowledge* 1.1 (2023), 22p. DOI: [10.4230/TGDK.1.1.13](https://doi.org/10.4230/TGDK.1.1.13). URL: <https://hal.science/hal-04358408>.

- [35] J. Chen, H. Dong, J. Hastings, E. Jiménez-Ruiz, V. López, P. Monnin, C. Pesquita, P. Škoda and V. Tamma. 'Knowledge Graphs for the Life Sciences: Recent Developments, Challenges and Opportunities'. In: *Transactions on Graph Data and Knowledge* 1.1 (19th Dec. 2023), pp. 1–33. DOI: [10.4230/TGDK.1.1.5](https://doi.org/10.4230/TGDK.1.1.5). URL: <https://inria.hal.science/hal-04353523>.
- [36] C. d'Amato, L. Mahon, P. Monnin and G. Stamou. 'Machine Learning and Knowledge Graphs: Existing Gaps and Future Research Challenges'. In: *Transactions on Graph Data and Knowledge* 1.1 (19th Dec. 2023), pp. 1–35. DOI: [10.4230/TGDK.1.1.8](https://doi.org/10.4230/TGDK.1.1.8). URL: <https://inria.hal.science/hal-04353543>.
- [37] N. Hubert, P. Monnin, A. Brun and D. Monticolo. 'Sem@K: Is my knowledge graph embedding model semantic-aware?' In: *Semantic Web – Interoperability, Usability, Applicability* 14.6 (11th Dec. 2023), pp. 1273–1309. DOI: [10.3233/SW-233508](https://doi.org/10.3233/SW-233508). URL: <https://inria.hal.science/hal-04344975>.
- [38] P. Maillot, O. Corby, C. Faron, F. Gandon and F. Michel. 'IndeGx: A Model and a Framework for Indexing RDF Knowledge Graphs with SPARQL-based Test Suites'. In: *Journal of Web Semantics* (19th Jan. 2023). DOI: [10.1016/j.websem.2023.100775](https://doi.org/10.1016/j.websem.2023.100775). URL: <https://hal.science/hal-03946680>.
- [39] A. Ollagnier, E. Cabrio and S. Villata. 'Unsupervised Fine-grained Hate Speech Target Community Detection and Characterisation on Social Media'. In: *Social Network Analysis and Mining* (3rd Mar. 2023). DOI: [10.1007/s13278-023-01061-4](https://doi.org/10.1007/s13278-023-01061-4). URL: <https://hal.science/hal-04014977>.
- [40] S. Pamerlon, A.-S. Archambeau, F. Pando de la Hoz, G. Olivier, F. Michel, S. Terceirie, E. Rodinson, N. Maurel, G. Martínez-Sagarra, A. Kerner, R. Vignes Lebbe, B. Schatz and P. Dupont. 'Knowledge Base on Species Life Traits : A Spanish/French Plinian Core implementation use case'. In: *Biodiversity Information Science and Standards* 7 (29th Aug. 2023), e111784. DOI: [10.3897/biss.7.111784](https://doi.org/10.3897/biss.7.111784). URL: <https://hal.science/hal-04198414>.
- [41] K. K. Patel, G. Desaulniers, A. Lodi and F. Lecue. 'Explainable prediction of Qcodes for NOTAMs using column generation'. In: *Journal of the Operational Research Society* (2023), pp. 1–11. DOI: [10.1080/01605682.2023.2181715](https://doi.org/10.1080/01605682.2023.2181715). URL: <https://hal.science/hal-04255231>.
- [42] S. Shen, M. Treleani, D. Compagno and M. Winckler. 'From Stock Shots to Ghost Data: Tracking Audiovisual Archives about the European Union'. In: *VIEW. Journal of European Television History and Culture* 12.23 (5th Sept. 2023), pp. 4–23. DOI: [10.18146/view.292](https://doi.org/10.18146/view.292). URL: <https://inria.hal.science/hal-04233316>.
- [43] M. Treleani, S. Shen, M. Winckler and A. Menin. 'Circulation et répétition des images de stock dans les médias audiovisuels La représentation visuelle des migrants dans les JT européens'. In: *Les Cahiers du numérique* 1.19 (2023). DOI: [10.3166/LCN.2023.002](https://doi.org/10.3166/LCN.2023.002). URL: <https://hal.science/hal-04293239>.

### National journals

- [44] A. Toulet, F. Michel, A. Bobasheva, A. Menin, S. Dupré, M.-C. Deboin, M. Winckler and A. Tchechmedjiev. 'ISSA : un graphe de connaissances au service de la recherche bibliographique'. In: *Revue des Nouvelles Technologies de l'Information* 39 (2023), pp. 515–522. URL: <https://imt-mines-ales.hal.science/hal-03969275>.

### Invited conferences

- [45] O. Boissier, P.-A. Champin, A. K. Chopra, N. Fornara, F. Gandon, R. Taelman and J. van den Bussche. 'Decentralized Hypermedia Ecosystems'. In: *Dagstuhl Seminar. Agents on the Web*. Vol. 23081. Wadern, Germany: Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2023, pp. 123–126. URL: <https://hal-emse.ccsd.cnrs.fr/emse-04240553>.
- [46] O. Boissier, A. Ciorcea, F. Gandon, S. Mayer and A. Ricci. 'Hypermedia Multi-Agent Systems'. In: *Dagstuhl Seminar. Agents on the Web*. Vol. 23081. Wadern, Germany: Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2023, pp. 117–118. URL: <https://hal-emse.ccsd.cnrs.fr/emse-04240524>.

- [47] M. Buffa and A. Vidal-Mazuy. ‘High-performance Web-Audio applications, the challenges of writing a DAW’. In: Réunion de lancement du projet ANR DOTS. Paris, France, 13th Oct. 2023. URL: <https://inria.hal.science/hal-04337523>.
- [48] A. Ciorcea, C. Faron, J. Fred Hübner, E. O’Neill, M. Poveda-Villalón, A. Ricci and A. Zimmermann. ‘Abstractions for Agents on the Web’. In: *Dagstuhl Seminar*. Agents on the Web. Vol. 23081. Wadern, Germany: Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2023, pp. 126–130. URL: <https://hal-emse.ccsd.cnrs.fr/emse-04240584>.
- [49] F. Gandon. ‘Finding the Critical MAS(S): Resources and Representations Needed for Weaving a Web that Hosts Linked Multi-Agent Systems’. In: *Dagstuhl Seminar 23081*. Dagstuhl Seminar 23081 : Agents on the Web. Dagstuhl, Germany, 20th Feb. 2023. URL: <https://inria.hal.science/hal-03989703>.

### International peer-reviewed conferences

- [50] R. Abu Ahmad, M. Critelli, Ş. Efeoğlu, E. Mancini, C. Ringwald, X. Zhang and A. Meroño Peñuela. ‘Draw Me Like My Triples: Leveraging Generative AI for Wikidata Image Completion’. In: *CEUR Workshop Proceedings*. Wikidata 2023 - The 4th Wikidata Workshop (at ISWC 2023). Athenes, Greece, 7th Nov. 2023. URL: <https://hal.science/hal-04262826>.
- [51] H. A. Akl. ‘A ML-LLM pairing for better code comment classification’. In: *FIRE 2023 Working Notes*. FIRE (Forum for Information Retrieval Evaluation) 2023. Goa, India, 31st Dec. 2023. URL: <https://inria.hal.science/hal-04311401>.
- [52] H. A. Akl. ‘PSYCHIC: A Neuro-Symbolic Framework for Knowledge Graph Question-Answering Grounding’. In: *The First Scholarly QALD Challenge @ ISWC 2023*. ISWC 2023 - International semantic web conference. Athens (Greece), Greece, 1st Dec. 2023. URL: <https://inria.hal.science/hal-04311370>.
- [53] A. Ballout, C. da Costa Pereira and A. G. B. Tettamanzi. ‘Predicting the Acceptability of Atomic Candidate OWL Class Axioms’. In: *IEEE Xplore*. The 22nd IEEE/WIC International Conference on Web Intelligence and Intelligent Agent Technology. 2023 IEEE International Conference on Web Intelligence and Intelligent Agent Technology (WI-IAT). Venice, Italy, 2023. DOI: [10.1109/WI-IAT59888.2023.00055](https://doi.org/10.1109/WI-IAT59888.2023.00055). URL: <https://inria.hal.science/hal-04346839>.
- [54] A. Barisic, P. Sigrist, S. Oliver, A. Sciarra and M. Winckler. ‘Driver Model for Take-Over-Request in Autonomous Vehicles’. In: UMAP 2023 - 31st ACM Conference on User Modeling, Adaptation and Personalization. UMAP ’23 Adjunct: Adjunct Proceedings of the 31st ACM Conference on User Modeling, Adaptation and Personalization. Limassol, Cyprus: ACM, 2023, pp. 317–324. DOI: [10.1145/3563359.3596994](https://doi.org/10.1145/3563359.3596994). URL: <https://hal.science/hal-04205128>.
- [55] M. Buffa and A. Vidal-Mazuy. ‘WAM-studio, a Digital Audio Workstation (DAW) for the Web’. In: *ACM digital library*. WWW 2023 - The ACM Web Conference 2023. WWW ’23 Companion: Companion Proceedings of the ACM Web Conference 2023. Austin TX USA, United States, 30th Apr. 2023. DOI: [10.1145/3543873.3587987](https://doi.org/10.1145/3543873.3587987). URL: <https://inria.hal.science/hal-04335612>.
- [56] M. Buffa, A. Vidal-Mazuy, L. May and M. Winckler. ‘WAM-Studio: A Web-Based Digital Audio Workstation to Empower Cochlear Implant Users’. In: *Springer Lecture Notes in Computer Sciences*. INTERACT 2023 - 19th IFIP TC13 International Conference. Vol. LNCS. 14142. Human-Computer Interaction – INTERACT 2023: 19th IFIP TC13 International Conference, Part I. York, United Kingdom: Springer Nature Switzerland, 25th Aug. 2023, pp. 101–110. DOI: [10.1007/978-3-031-42280-5\\_6](https://doi.org/10.1007/978-3-031-42280-5_6). URL: <https://inria.hal.science/hal-04233343>.
- [57] M. Buffa, A. Vidal-Mazuy and Q. Plet. ‘Web Audio Modules: Swiss knife for audio plugin developments on the Web Platform’. In: <https://ieeexplore.ieee.org/xpl/conhome/10335168/proceeding>. Internet of Sounds 2023. Pise, Italy, 26th Oct. 2023. DOI: [10.5281/zenodo.6769408](https://doi.org/10.5281/zenodo.6769408). URL: <https://inria.hal.science/hal-04335580>.

- [58] L. Cadorel and A. G. B. Tettamanzi. 'A Comparative Study of Text Representations for French Real-Estate Classified Advertisements Information Extraction'. In: *CEUR Workshop Proceedings. ENIGMA 2023 : 1st Workshop on AI-driven heterogeneous data management : Completing, merging, handling inconsistencies and query-answering*. Vol. 3495. ENIGMA 2023 - AI-driven heterogeneous data management: Completing, merging, handling inconsistencies and query-answering 2023. Rhodes, Greece, 3rd Sept. 2023, pp. 55–63. URL: <https://inria.hal.science/hal-04346759>.
- [59] D. El Zein, A. Câmara, C. da Costa Pereira and A. G. B. Tettamanzi. 'RULKNE: Representing User Knowledge State in Search-as-Learning with Named Entities'. In: *CHIIR 2023 - ACM SIGIR Conference on Human Information Interaction and Retrieval. CHIIR '23: Proceedings of the 2023 Conference on Human Information Interaction and Retrieval*. Austin, TX, United States: ACM, 19th Mar. 2023, pp. 388–393. DOI: [10.1145/3576840.3578330](https://doi.org/10.1145/3576840.3578330). URL: <https://inria.hal.science/hal-04152998>.
- [60] R. Felin, C. Faron and A. G. B. Tettamanzi. 'A Framework to Include and Exploit Probabilistic Information in SHACL Validation Reports'. In: *IC2023 : 34es Journées francophones d'Ingénierie des Connaissances. IC 2023 - 34es Journées francophones d'Ingénierie des Connaissances @ Plate-Forme Intelligence Artificielle (PFIA 2023). IC2023 : 34es Journées francophones d'Ingénierie des Connaissances*. Strasbourg, France, 3rd July 2023. URL: <https://inria.hal.science/hal-04152604>.
- [61] R. Felin, C. Faron and A. G. B. Tettamanzi. 'A Framework to Include and Exploit Probabilistic Information in SHACL Validation Reports'. In: *Lecture Notes in Computer Science. ESWC 2023 - 20th International European Semantic Web Conference*. Vol. LNCS-13870. The Semantic Web 20th International Conference, ESWC 2023, Hersonissos, Crete, Greece, May 28–June 1, 2023, Proceedings. Hersonissos, Greece: Springer Nature Switzerland, 2023, pp. 91–104. DOI: [10.1007/978-3-031-33455-9\\_6](https://doi.org/10.1007/978-3-031-33455-9_6). URL: <https://inria.hal.science/hal-04031744>.
- [62] P. Goffredo, E. Cabrio, S. Villata, S. Haddadan and J. Torres Sanchez. 'DISPUTool 2.0: A Modular Architecture for Multi-Layer Argumentative Analysis of Political Debates'. In: *Proceedings of the AAAI Conference on Artificial Intelligence. AAAI-23 - 37th AAAI Conference on Artificial Intelligence*. Vol. 37. AAAI-23 Special Programs, IAAI-23, EAAI-23, Student Papers and Demonstrations 13. Washington, DC, United States, 26th June 2023, pp. 16431–16433. DOI: [10.1609/aaai.v37i13.27069](https://doi.org/10.1609/aaai.v37i13.27069). URL: <https://hal.science/hal-04349796>.
- [63] P. Goffredo, M. Chaves, S. Villata and E. Cabrio. 'Argument-based Detection and Classification of Fallacies in Political Debates'. In: *ACL Anthology. EMNLP 2023 - Conference on Empirical Methods in Natural Language Processing*. Vol. 2023.findings-emnlp.684. Findings of the Association for Computational Linguistics: EMNLP 2023. Singapore (SG), Singapore: Association for Computational Linguistics, Dec. 2023, pp. 11101–11112. DOI: [10.18653/v1/2023.emnlp-main.684](https://doi.org/10.18653/v1/2023.emnlp-main.684). URL: <https://hal.science/hal-04260221>.
- [64] F.-Z. Hannou, V. Charpenay, M. Lefrançois, C. Roussey, A. Zimmermann and F. Gandon. 'The ACIMOV Methodology: Agile and Continuous Integration for Modular Ontologies and Vocabularies'. In: *MK 2023 - 2nd Workshop on Modular Knowledge associated with FOIS 2023 - the 13th International Conference on Formal Ontology in Information Systems*. Sherbrooke, Canada, 19th July 2023. URL: <https://laas.hal.science/hal-04187236>.
- [65] N. Hubert, H. Paulheim, P. Monnin, A. Brun and D. Monticolo. 'Schema First! Learn Versatile Knowledge Graph Embeddings by Capturing Semantics with MASCHInE'. In: *ACM digital library. K-CAP 2023 - Knowledge Capture Conference 2023*. Vol. 2023. K-CAP '23: Proceedings of the 12th Knowledge Capture Conference 2023. Pensacola, United States: ACM, 5th Dec. 2023, pp. 188–196. DOI: [10.1145/3587259.3627550](https://doi.org/10.1145/3587259.3627550). URL: <https://inria.hal.science/hal-04344873>.
- [66] A. Iglesias-Molina, D. van Assche, J. Arenas-Guerrero, B. de Meester, C. Debruyne, S. Jozashoori, P. Maria, F. Michel, D. Chaves-Fraga and A. Dimou. 'The RML Ontology: A Community-Driven Modular Redesign After a Decade of Experience in Mapping Heterogeneous Data to RDF'. In: *Proceedings of ISWC 2023. ISWC 2023 - 22th International Semantic Web Conference*. Athens, Greece, 2023. URL: <https://hal.science/hal-04201661>.

- [67] P. Maillot, O. Corby, C. Faron, F. Gandon and F. Michel. ‘Metadatamatic: A Web application to Create a Dataset Description’. In: *WWW '23 Companion: Companion Proceedings of the ACM Web Conference 2023*. WWW '23 - The ACM Web Conference 2023. WWW '23 Companion: Companion Proceedings of the ACM Web Conference 2023. Austin TX, United States: ACM, 2023, pp. 123–126. DOI: [10.1145/3543873.3587328](https://doi.org/10.1145/3543873.3587328). URL: <https://hal.science/hal-04171196>.
- [68] S. Marro, T. A. Collias, E. Cabrio and S. Villata. ‘On the Automatic Assessment of Natural Language Expert Explanations in Medicine’. In: *CEUR workshop proceedings*. AIxIA 2023 - 22nd International Conference of the Italian Association for Artificial Intelligence. Vol. 3578. 2nd AIxIA Workshop on Artificial Intelligence For Healthcare (HC@AIxIA 2023). Rome, Italy, 6th Nov. 2023, pp. 83–98. URL: <https://inria.hal.science/hal-04371254>.
- [69] S. Marro, B. Molinet, E. Cabrio and S. Villata. ‘Natural Language Explanatory Arguments for Correct and Incorrect Diagnoses of Clinical Cases’. In: *SciTePress digital library*. ICAART 2023 - 15th International Conference on Agents and Artificial Intelligence. Vol. 1. Proceedings of the 15th International Conference on Agents and Artificial Intelligence - (Volume 1). Lisbon (Portugal), Portugal, 22nd Feb. 2023, pp. 438–449. DOI: [10.5220/0011927000003393](https://doi.org/10.5220/0011927000003393). URL: <https://hal.science/hal-04002207>.
- [70] A. Menin, H. Ayari, F. Michel and M. Winckler. ‘Using polymorphic glyphs to support the visual exploration of hierarchical spatio-temporal data’. In: *Lecture notes in computer science*. INTERACT 2023 - 19th IFIP International Conference of Technical Committee 13 (Human- Computer Interaction). Vol. LNCS-14145. Human-Computer Interaction – INTERACT 2023 19th IFIP TC13 International Conference, York, UK, August 28 – September 1, 2023, Proceedings, Part IV. York, United Kingdom, 28th Aug. 2023, pp. 325–329. DOI: [10.1007/978-3-031-42293-5\\_31](https://doi.org/10.1007/978-3-031-42293-5_31). URL: <https://hal.science/hal-04188896>.
- [71] F. Michel and F. Gandon. ‘Turn your web APIs into just another piece of your Knowledge Graph’. In: *KGC 2023 - The Knowledge Graph Conference*. New York (NY), United States, 12th May 2023. URL: <https://hal.science/hal-04351695>.
- [72] N. B. Ocampo, E. Cabrio and S. Villata. ‘Playing the Part of the Sharp Bully : Generating Adversarial Examples for Implicit Hate Speech Detection’. In: *Findings of the Association for Computational Linguistics: ACL 2023*. ACL 2023 - 61st Annual Meeting of the Association for Computational Linguistics. Toronto, Canada: Association for Computational Linguistics, 9th July 2023, pp. 2758–2772. DOI: [10.18653/v1/2023.findings-acl.173](https://doi.org/10.18653/v1/2023.findings-acl.173). URL: <https://hal.science/hal-04214125>.
- [73] N. B. Ocampo, E. Cabrio and S. Villata. ‘Unmasking the Hidden Meaning: Bridging Implicit and Explicit Hate Speech Embedding Representations’. In: *ACL Anthology*. EMNLP 2023 - Conference on Empirical Methods in Natural Language Processing. Vol. 2023.findings-emnlp.441. Findings of the Association for Computational Linguistics: EMNLP 2023. Singapore, France: Association for Computational Linguistics, Dec. 2023, pp. 6626–6637. DOI: [10.18653/v1/2023.findings-emnlp.441](https://doi.org/10.18653/v1/2023.findings-emnlp.441). URL: <https://hal.science/hal-04351644>.
- [74] N. B. Ocampo, E. Sviridova, E. Cabrio and S. Villata. ‘An In-depth Analysis of Implicit and Subtle Hate Speech Messages’. In: *ACL Anthology*. EACL 2023 - 17th Conference of the European Chapter of the Association for Computational Linguistics. Vol. 2023. Proceedings of the 17th Conference of the European Chapter of the Association for Computational Linguistics. Dubrovnik, France: Association for Computational Linguistics, May 2023, pp. 1997–2013. DOI: [10.18653/v1/2023.eacl-main.147](https://doi.org/10.18653/v1/2023.eacl-main.147). URL: <https://hal.science/hal-04214094>.
- [75] A. Ollagnier, E. Cabrio and S. Villata. ‘Harnessing Bullying Traces to Enhance Bullying Participant Role Identification in Multi-Party Chats’. In: *Florida Online Journals*. FLAIRS 2023 - 36th International conference of the Florida artificial intelligence research society. Florida / USA, United States, 14th May 2023. URL: <https://hal.science/hal-04037120>.

- [76] A. Ollagnier, E. Cabrio, S. Villata and S. Tonelli. 'BiRDy: Bullying Role Detection in Multi-Party Chats'. In: *Proceedings of the AAAI Conference on Artificial Intelligence*. AAAI-23 - 37th AAAI Conference on Artificial Intelligence. Vol. 37. AAAI-23 Special Programs, IAAI-23, EAAI-23, Student Papers and Demonstrations 13. Washington DC, United States: Association for the Advancement of Artificial Intelligence, 6th Sept. 2023, pp. 16464–16466. DOI: [10.1609/aaai.v37i13.27080](https://doi.org/10.1609/aaai.v37i13.27080). URL: <https://hal.science/hal-04315527>.
- [77] M. A. N. Pour, A. Algergawy, P. Buche, L. J. Castro, J. Chen, A. Coulet, J. Cufi, H. Dong, O. Fallatah, D. Faria, I. Fundulaki, S. Hertling, Y. He, I. Horrocks, M. Huschka, L. Ibanescu, D. S. Jain, E. Jiménez-Ruiz, N. Karam, P. Lambrix, H. Li, Y. Li, P. Monnin, E. Nasr, H. Paulheim, C. Pesquita, T. Saveta, P. Shvaiko, G. Sousa, C. Trojahn, J. Vataščinová, M. Wu, B. Yaman, O. Zamazal and L. Zhou. 'Results of the Ontology Alignment Evaluation Initiative 2023'. In: *Proceedings of the 18th International Workshop on Ontology Matching co-located with the 22nd International Semantic Web Conference (ISWC 2023), Athens, Greece, November 7, 2023*. 18th International Workshop on Ontology Matching (OM 2023). Vol. 3591. CEUR workshop proceedings. Athens, Greece: CEUR-WS.org, 7th Nov. 2023, pp. 97–139. URL: <https://inria.hal.science/hal-04366893>.
- [78] S. Ren, M. Buffa, L. Pottier, Y. Yu and G. Schalk. 'Reflex-in: Generate Music on the Web with Real-time Brain Wave'. In: *ACM digital library*. WWW 2023 - The ACM Web Conference 2023 : Web Creativity track. WWW '23 Companion: Companion Proceedings of the ACM Web Conference 2023. Austin TX USA, France, 30th Apr. 2023, pp. 598–600. DOI: [10.1145/3543873.3587315](https://doi.org/10.1145/3543873.3587315). URL: <https://inria.hal.science/hal-04093045>.
- [79] F. A. S. Robert, H.-Y. Wu, L. Sassatelli, S. Ramanoel, A. Gros and M. Winckler. 'An Integrated Framework for Understanding Multimodal Embodied Experiences in Interactive Virtual Reality'. In: *2023 IMX - ACM International Conference on Interactive Media Experiences*. Nantes, France, 12th June 2023, <https://dl.acm.org/conference/imx>. DOI: [10.1145/3573381.3596150](https://doi.org/10.1145/3573381.3596150). URL: <https://hal.science/hal-04102737>.
- [80] S. Sarkadi and F. Gandon. 'Interoperable AI for Self-Organisation'. In: *SaSSO 2023 Workshop - Sustainable and Scalable Self-Organisation Workshop (ASCOS 2023)*. Toronto, Canada, 25th Sept. 2023. DOI: [10.1109/ACSOS-C58168.2023.00043](https://doi.org/10.1109/ACSOS-C58168.2023.00043). URL: <https://inria.hal.science/hal-04384816>.
- [81] M. Tounsi Dhouib, Q. Merilleau, C. Guerrero, M. Corneli, C. Faron and A. Zucker. 'Automatic Semantic Classification of Ancient Zoological Texts'. In: *Journal of cultural Heritage*. IAMAHA 2023 - 1st International Conference on Artificial Intelligence and Applied Mathematics for History and Archaeology. Nice, France, 27th Nov. 2023. URL: <https://hal.science/hal-04210669>.
- [82] E. Tysinger, M. Pagni, O. Kirchhoffer, F. Mehl, F. Gandon, J.-L. Wolfender and L.-F. Nothias. 'An Artificial Intelligence Agent for Navigating Knowledge Graph Experimental Metabolomics Data'. In: *Frontiers in Metabolomics - Book of abstract* -. 2023 Swiss Metabolomics Society Annual Meeting. Frontiers in Metabolomics. Zurich, Switzerland, 15th Sept. 2023. URL: <https://hal.science/hal-04381448>.
- [83] N. Yacoubi Ayadi, C. Faron, F. Michel, F. Gandon and O. Corby. 'Computing and Visualizing Agro-Meteorological Parameters based on an Observational Weather Knowledge Graph'. In: *Companion Proceedings of the ACM Web Conference 2023*. WWW 2023 - ACM Web Conference 2023 - International World Wide Web Conference. WWW '23 Companion: Companion Proceedings of the ACM Web Conference 2023. Austin (TX), United States: ACM, 30th Apr. 2023. DOI: [10.1145/3543873.3587357](https://doi.org/10.1145/3543873.3587357). URL: <https://hal.science/hal-04230234>.
- [84] A. Zimmermann, A. Ciortea, C. Faron, E. O'Neill and M. Poveda-Villalón. 'Pody: a Solid-based approach to embody agents in Web-based Multi-Agent-Systems'. In: *EMAS 2023 - 11th International Workshop on Engineering Multi-Agent Systems*. Londres, United Kingdom, 29th May 2023. URL: <https://hal-emse.ccsd.cnrs.fr/emse-04058783>.

### National peer-reviewed Conferences

- [85] A. Barbe, M. Tounsi Dhouib, C. Faron, M. Corneli and A. Zucker. ‘Building a knowledge graph from manual annotations of ancient zoology texts’. In: IC 2023 - 34es Journées francophones d’Ingénierie des Connaissances @ Plate-Forme Intelligence Artificielle (PFIA 2023). IC2023 : 34es Journées francophones d’Ingénierie des Connaissances. Starsbourg, France, 3rd July 2023. URL: <https://inria.hal.science/hal-04156996>.
- [86] L. Cadorel, A. G. B. Tettamanzi and F. Gandon. ‘Graphes de Connaissances et Ontologie pour la Représentation de Données Immobilières Issues d’Annonces en Texte Libre’. In: IC 2023 - 34es Journées francophones d’Ingénierie des Connaissances @ Plate-Forme Intelligence Artificielle (PFIA 2023). IC2023 : 34es Journées francophones d’Ingénierie des Connaissances. Strasbourg, France, 3rd July 2023. URL: <https://inria.hal.science/hal-04153252>.
- [87] M. Tounsi Dhouib, E. Kostrykina and C. Faron. ‘Automatic definition of the level of textual difficulty of documents’. In: *Revue des Nouvelles Technologies de l’Information*. 24ème conférence francophone sur l’Extraction et la Gestion des Connaissances. Dijon, France, 2024. URL: <https://hal.science/hal-04353060>.
- [88] A. Vidal-Mazuy and M. Buffa. ‘A Digital Audio Workstation for the Web Platform’. In: *Actes des journées d’informatique musicale 2023*. Journées d’Informatique Musicale 2023. Vol. Pages 191-199. Paris, France, 26th May 2023, p. 8. URL: <https://inria.hal.science/hal-04334794>.

### Conferences without proceedings

- [89] E. Puech, A. Barbe, M. Corneli and I. Théry-Parisot. ‘A combined database of anatomical features and micrographs for selected Southern African woods and charcoals: a perspective on an identification key for anthracological study’. In: XXIVe colloque du GMPCA (Groupe des Méthodes Pluridisciplinaires Contribuant à l’Archéologie) : Archéométrie 2023. Nice, France, 17th Apr. 2023. URL: <https://hal.science/hal-04122866>.

### Scientific book chapters

- [90] A. Barisic and M. Winckler. ‘Towards User Profile Meta-Ontology’. In: *Human-Centered Software Engineering. HCSE 2023. Lecture Notes in Computer Science. LNCS*. 31st Oct. 2023. URL: <https://hal.science/hal-04210148>.
- [91] M. Buffa and J. Lebrun. ‘A browser-based WebAudio ecosystem to play real-time with simulations of historic guitar tube amps and their typical distortions’. In: *Distortion in Music Production . The Soul of Sonics*. Taylor and Francis publishers, 2nd June 2023, pp. 28–46. DOI: [10.4324/9780429356841](https://doi.org/10.4324/9780429356841). URL: <https://inria.hal.science/hal-04337603>.
- [92] A. Menin, P. Maillot, C. Faron, O. Corby, C. M. Dal Sasso Freitas, F. Gandon and M. Winckler. ‘LDViz: a tool to assist the multidimensional exploration of SPARQL endpoints’. In: *Web Information Systems and Technologies : 16th International Conference, WEBIST 2020, November 3–5, 2020, and 17th International Conference, WEBIST 2021, October 26–28, 2021, Virtual Events, Revised Selected Papers*. Vol. LNBIP - 469. LNBIP - Lecture Notes in Business Information Processing. Springer, 2023, pp. 149–173. DOI: [10.1007/978-3-031-24197-0](https://doi.org/10.1007/978-3-031-24197-0). URL: <https://hal.science/hal-03929913>.

### Edition (books, proceedings, special issue of a journal)

- [93] J. Abdelnour Nocera, M. Kristín Lárusdóttir, H. Petrie, A. Piccinno and M. Winckler, eds. *Human-Computer Interaction – INTERACT 2023 : part I: 19th IFIP TC13 International Conference, York, UK, August 28 – September 1, 2023, Proceedings, Part I*. Vol. LNCS-14142. Lecture Notes in Computer Science. Springer Nature Switzerland, 2023. DOI: [10.1007/978-3-031-42280-5](https://doi.org/10.1007/978-3-031-42280-5). URL: <https://inria.hal.science/hal-04233377>.



- [94] J. Abdelnour Nocera, M. Kristín Lárusdóttir, H. Petrie, A. Piccinno and M. Winckler, eds. *Human-Computer Interaction – INTERACT 2023 : part II: 19th IFIP TC13 International Conference, York, UK, August 28 – September 1, 2023, Proceedings, Part II*. Vol. LNCS-14143. Lecture Notes in Computer Science. Springer Nature Switzerland, 2023. DOI: [10.1007/978-3-031-42283-6](https://doi.org/10.1007/978-3-031-42283-6). URL: <https://inria.hal.science/hal-04233382>.
- [95] J. Abdelnour Nocera, M. Kristín Lárusdóttir, H. Petrie, A. Piccinno and M. Winckler, eds. *Human-Computer Interaction – INTERACT 2023 : part III: 19th IFIP TC13 International Conference, York, UK, August 28 – September 1, 2023, Proceedings, Part III*. Vol. LNCS-14144. Lecture Notes in Computer Science. Springer Nature Switzerland, 2023. DOI: [10.1007/978-3-031-42286-7](https://doi.org/10.1007/978-3-031-42286-7). URL: <https://inria.hal.science/hal-04233383>.
- [96] J. Abdelnour Nocera, M. Kristín Lárusdóttir, H. Petrie, A. Piccinno and M. Winckler, eds. *Human-Computer Interaction – INTERACT 2023 : part IV: 19th IFIP TC13 International Conference, York, UK, August 28 – September 1, 2023, Proceedings, Part IV*. Vol. LNCS-14145. Lecture Notes in Computer Science. Springer Nature Switzerland; Springer Nature Switzerland, 2023. DOI: [10.1007/978-3-031-42293-5](https://doi.org/10.1007/978-3-031-42293-5). URL: <https://inria.hal.science/hal-04233389>.
- [97] S. Benferhat, G. Casini, T. Meyer and A. G. B. Tettamanzi, eds. *ENIGMA 2023 : AI-driven heterogeneous data management: Completing, merging, handling inconsistencies and query-answering 2023: Proceedings of 1st Workshop on AI-driven heterogeneous data management: Completing, merging, handling inconsistencies and query-answering, co-located with 20th International Conference on Principles of Knowledge Representation and Reasoning (KR 2023) Rhodes, Greece, September 3-4, 2023*. ENIGMA 2023 - 1st Workshop on AI-driven heterogeneous data management: Completing, merging, handling inconsistencies and query-answering. Vol. 3495. ENIGMA 2023 - AI-driven heterogeneous data management: Completing, merging, handling inconsistencies and query-answering 2023. CEUR-WS.org, 2023. URL: <https://inria.hal.science/hal-04346716>.
- [98] A. Dix, M. Winckler and M. Jones, eds. *EICS '23 Companion: Companion Proceedings of the 2023 ACM SIGCHI Symposium on Engineering Interactive Computing Systems*. ACM, 2023. DOI: [10.1145/3596454](https://doi.org/10.1145/3596454). URL: <https://inria.hal.science/hal-04233375>.
- [99] C. Faron and S. Loudcher, eds. *Extraction et Gestion des Connaissances : EGC 2023*. Vol. E39. Editions RNTI, 2023. URL: <https://hal.science/hal-04169619>.

#### Doctoral dissertations and habilitation theses

- [100] S. Marro. 'Argumentation quality : from general principles to healthcare applications'. Université Côte d'Azur, 8th Nov. 2023. URL: <https://theses.hal.science/tel-04402347>.

#### Reports & preprints

- [101] M. Toure, F. Gandon, K. Guidedi, M. Lo and C. Guéret. *Survey of Models and Architectures to Ensure Linked Data Access*. Inria & Université Côte d'Azur, CNRS, I3S, 23rd Jan. 2023, p. 23. URL: <https://inria.hal.science/hal-03956456>.

#### Other scientific publications

- [102] A. Barbe, M. T. Dhouib, C. Faron, M. Corneli and A. Zucker. 'Zoo-KG: A knowledge graph of zoological data'. In: AIUCD 2023 - Associazione per l'Informatica Umanistica e la Cultura Digitale. Sienna (Italy), Italy, 5th June 2023. URL: <https://hal.science/hal-04235805>.
- [103] S. Ouelhadj, P.-A. Champin, S. Jean-Daubias and J. Gaillard. 'A method to semantically enrich data using annotated UML'. In: EGC 2023 - Extraction et Gestion des Connaissances. Vol. RNTI-E-39. Lyon, France, 18th Jan. 2023, pp. 635–636. URL: <https://hal.science/hal-04021565>.
- [104] C. Ringwald, F. Gandon, C. Faron, F. Michel and H. A. Akl. 'Learning RDF pattern extractors from natural language and knowledge graphs - application to Wikipedia and the LOD'. In: ISWS 2023 - International Semantic Web Research Summer School / GEMSS 2023 - Generative Modeling Summer School. Bertinoro, Italy / Copenhagen, Denmark, June 2023. URL: <https://hal.science/hal-04175511>.

### 11.3 Other

#### Scientific popularization

- [105] M. Buffa and J.-F. Trubert. ‘Playing with Web Audio Modules: the Concept of Distributive Creation and Live Performance’. In: Digital Week Conference 2023. Nice, France, 12th June 2023. URL: <https://inria.hal.science/hal-04337545>.
- [106] C. Ringwald, F. Gandon, C. Faron, F. Michel and H. Abi Akl. ‘Well-Written Knowledge Graphs: Most Effective RDF Syntaxes for Triple Linearization in End-to-End Extraction of Relations from Texts (Student Abstract)’. In: AAAI 2024. Vancouver (BC), Canada, 27th Feb. 2024. URL: <https://hal.science/hal-04384724>.

#### Softwares

- [107] [SW] R. Cérés, O. Corby, E. Demairy and F. Gandon, *Corese* version 4.5.0, 25th July 2023. Inria; CNRS; Université Côte d’Azur. LIC: CECILL-C. DOI: [10.5281/zenodo.10376241](https://doi.org/10.5281/zenodo.10376241), HAL: [hal-04170333](https://hal-04170333), URL: <https://hal.science/hal-04170333>, VCS: <https://github.com/Wimmics/corese>, SWHID: [swh:1:dir:7ec9a0bccbba1254457985de26d524d0d5c4a9c5](https://sw.hid.io/1/dir/7ec9a0bccbba1254457985de26d524d0d5c4a9c5).

### 11.4 Cited publications

- [108] R. Borgo, J. Kehrer, D. H. S. Chung, E. Maguire, R. S. Laramee, H. Hauser, M. Ward and M. Chen. ‘Glyph-based Visualization: Foundations, Design Guidelines, Techniques and Applications’. In: *Eurographics State of the Art Reports* (2013). ISBN: 1558608192, pp. 39–63. URL: <http://www.cg.tuwien.ac.at/research/publications/2013/borgo-2013-gly/>.
- [109] A. Ciortea, S. Mayer, S. Bienz, F. Gandon and O. Corby. ‘Autonomous search in a social and ubiquitous Web’. In: *Personal and Ubiquitous Computing* (June 2020). DOI: [10.1007/s00779-020-01415-1](https://doi.org/10.1007/s00779-020-01415-1). URL: <https://hal.inria.fr/hal-03135088>.
- [110] F. Gandon. ‘The three ‘W’ of the World Wide Web call for the three ‘M’ of a Massively Multidisciplinary Methodology’. In: *10th International Conference, WEBIST 2014*. Ed. by V. Monfort and K.-H. Krempels. Vol. 226. Web Information Systems and Technologies. Barcelona, Spain: Springer International Publishing, Apr. 2014. DOI: [10.1007/978-3-319-27030-2](https://doi.org/10.1007/978-3-319-27030-2). URL: <https://hal.inria.fr/hal-01223236>.
- [111] F. Gandon, M. Buffa, E. Cabrio, O. Corby, C. Faron Zucker, A. Giboin, N. Le Thanh, I. Mirbel, P. Sander, A. G. B. Tettamanzi and S. Villata. ‘Challenges in Bridging Social Semantics and Formal Semantics on the Web’. In: *5th International Conference, ICEIS 2013*. Ed. by S. Hammoudi, J. Cordeiro, L. Maciaszek and J. Filipe. Vol. 190. Angers, France: Springer, July 2013, pp. 3–15. URL: <https://hal.inria.fr/hal-01059273>.
- [112] T. Munzner. *Visualization analysis and design*. CRC press, 2014.
- [113] A. Ollagnier, E. Cabrio, S. Villata and C. Blaya. ‘CyberAgressionAdo-v1: a Dataset of Annotated Online Aggressions in French Collected through a Role-playing Game’. In: *LREC 2022 - 13th Language Resources and Evaluation Conference*. Vol. 2022.lrec-1.91. Marseille, France, June 2022, pp. 867–875. URL: <https://hal.science/hal-03765860>.
- [114] A. Toulet, F. Michel, A. Bobasheva, A. Menin, S. Dupré, M.-C. Deboin, M. Winckler and A. Tchechmedjiev. ‘ISSA: Generic Pipeline, Knowledge Model and Visualization tools to Help Scientists Search and Make Sense of a Scientific Archive’. In: *ISWC 2022 - 21st International Semantic Web Conference*. Hangzhou, China, Oct. 2022. DOI: [10.1007/978-3-031-19433-7\\_38](https://doi.org/10.1007/978-3-031-19433-7_38). URL: <https://hal.science/hal-03807744>.
- [115] N. Yacoubi Ayadi, C. Faron, F. Michel, F. Gandon and O. Corby. ‘A Model for Meteorological Knowledge Graphs: Application to Météo-France Data’. In: *ICWE 2022- 22nd International Conference on Web Engineering*. 22nd International Conference on Web Engineering, ICWE 2022. Bari, Italy, July 2022. URL: <https://inria.hal.science/hal-03619869>.

- [116] N. Yacoubi Ayadi, C. Faron, F. Michel, F. Gandon and O. Corby. *WeKG-MF: a Knowledge Graph of Observational Weather Data*. ESWC 2022 - 19th Extended Semantic Web Conference. Poster. May 2022. DOI: [10.1007/978-3-031-11609-4\\_19](https://doi.org/10.1007/978-3-031-11609-4_19). URL: <https://inria.hal.science/hal-03657694>.