Project-team EXMO

INRIA, Evaluation of Theme “Représentation et traitement des données et des connaissances”

October 2015

Project-team acronym :	**EXMO**
Project-team title :	**Computer-mediated exchange of structured knowledge**
	Échanges de connaissance structurée médiatisés par ordinateur
Scientific leader :	**Jérôme Euzenat**
Research center :	**INRIA Grenoble Rhône-Alpes (Montbonnot Saint-Martin)**
Common project-team with :	Laboratoire d’Informatique de Grenoble (**LIG**)
URL:	http://exmo.inria.fr/

1 Personnel

Personnel (October 2011)

	Misc.	INRIA	CNRS	University	**Total**
DR(1) / Professors		1			**1**
CR(2) / Assistant Professor				1	**1**
Permanent Engineer(3)					
Temporary Engineer(4)					
PhD Students		2			**2**
Post-Doc.		1			**1**
Total		**4**		**1**	**5**
External Collaborators				1	**1**
Visitors ($>$ 1 month)					

Personnel (October 2015)

	Misc.	INRIA	CNRS	University	**Total**
DR / Professors		1			**1**
CR / Assistant Professor				2	**2**
Inria Starting Position					
Permanent Engineer					
Temporary Engineer		1			**1**
PhD Students		2		1	**3**
Post-Doc.					
Total		**4**		**3**	**7**
External Collaborators					
Visitors ($>$ 1 month)					

(1) “Senior Research Scientist (Directeur de Recherche)” (2) “Junior Research Scientist (Chargé de Recherche)”
(3) “Civil servant (CNRS, INRIA, ...)” (4) “Associated with a contract (Ingénieur Expert or Ingénieur Associé)”

Changes in staff

	Misc.	INRIA	CNRS	University	total
DR / Professors					
CR / Assistant Professors				1	1
Inria Starting Positions					
Arrival				2013	

Current composition of the project-team (October 2015):

Jérôme Euzenat, DR1, INRIA
Jérôme David, Associate professor[1], U. Pierre Mendès-France, Grenoble
Manuel Atencia Arcas[1], Associate professor, U. Pierre Mendès-France, Grenoble
Nicolas Guillouet, Engineer Technology transfer, INRIA
Armen Inants, PhD student, INRIA
Tatiana Lesnikova, PhD student, U. Joseph Fourier, Grenoble
Adam Sanchez, PhD student, INRIA (ANR Lindicle grant)

Current position of former project-team members (including PhD students; 2011-2015):

Jose-Luis Aguirre Cervantes (Post-doc), Engineer, GenoStar, Grenoble
Luz-Maria Priego Roche (Post-doc), Unemployed
Cássia Trojahn dos Santos (Post-doc), Associate professor[1], U. Toulouse
Melisachew Wudage Chekol (PhD), Post doctoral researcher, U. Mannheim
Zhengjie Fan (PhD), Research engineer, China mobile, Beijing

Last INRIA enlistments

None during the period.

Team History

EXMO has been created as an INRIA project in 2003. It was a founding team of the LIG (Laboratoire d'informatique de Grenoble) in 2006.
The initial EXMO topic concerned interoperability on the web in its full generality. During our 2005 INRIA evaluation, this topic has been refocussed on "ontology matching and alignment". At the 2011 INRIA evaluation, we have been encouraged to reinforce our work on "data interlinking" (see §5).

2 Work progress

Our work is at the crossroads of:

- The challenge of "transforming the tidal wave of data into trustworthy knowledge libraries" of the "Mastering complexity: data, networks and flows" strategic objective, and
- The "human and knowledge: emergence, mediation and education" topic of the "Sciences that serve individuals, society and knowledge" strategic objective

of the INRIA 2013-2017 strategic plan: "towards INRIA 2020".

[1] Maître de conférences = Lecturer with tenure

2.1 Keywords

knowledge representation, semantics of knowledge representation, ontologies, semantic web, knowledge transformation, ontology matching, ontology alignment, data interlinking, link keys, RDF, RDF Path, OWL, SPARQL, PSPARQL, Alignment API, Alignment server, OntoSim, Aroma.

2.2 Context and overall goal of the project

The semantic web blends the communication capabilities of the web with knowledge representation. Expressing formalised knowledge on a computer is useful, not exclusively for the consumption of the computer, but for communication. The goal of EXMO is the development of theoretical, experimental and software tools for communicating formalised knowledge.
There is no reason why knowledge expressed on the web should be in a single format or by reference to a single vocabulary (or ontology). In order to interoperate, representations have to be matched and transformed. We build on our experience of alignments as representing the relationships between ontologies. Such alignments may be used for generating knowledge transformations (or any other kind of mediators) used for interoperating or interlinking data.
We focus on the design of an alignment infrastructure and on the investigation of alignment properties when they are used for reconciling ontologies. More precisely, our work may be grouped in three areas:

- Developing different aspects of ontology matching, including evaluation, alignment semantics and new types of matchers;
- Data interlinking, with or without relations to ontology matching;
- Dynamic evolution of knowledge networks.

Our work is naturally applied in all contexts in which ontologies are used for expressing knowledge that has to be communicated. It is more directly focussed on the infrastructure of the semantic web and the web of data.

2.3 Objectives for the evaluation period

We summarise here the objectives given at the last evaluation (*in italics*) and indicate to what extent they have been achieved (and where to find them in this report). The full text of these objectives is reported in Appendix A. Sections 2.4 to 2.7 adopt a simpler organisation which corresponds better to the work achieved.

Alignment foundations

Three specific topics were considered:

Semantics of network of ontologies *Our goal is to develop further the semantics of networks of ontologies. We aim at considering two research directions: (i) using the semantics for defining what a peer in a peer-to-peer system can know; (ii) developing the principles of revision in networks of ontologies.* Although in a different way, we have achieved a weight-based semantics for alignments (§2.4.3, p. 6), and global revision definition for networks of ontologies (§2.7.3, p. 13).

Algebraic manipulation of alignments *The use of algebras of relations offers a formal intermediate position. Hence we want to fully define the possible operations on alignments, including composition, reasoning and combining alignments.* This topic started late due to the difficulty to hire someone to tackle it, but finally the theoretical foundations are laid out (§2.4.3, p. 7).

Distances between ontologies *We want to propose computable generalisations of our alignment space measures. We also plan to investigate agreement and disagreement between ontologies proposed by M. d'Aquin.* Distances between ontologies have not been investigated the way we would have liked it due to lack of people to work on it [23]. We still have good ideas to develop.

Alignment infrastructure

API improvement *This will require two major redesigns: integrating algebras of relations at the core of the API, and integrating better reasoning mechanisms within the API. Under the pressure of application using it (see Applications and Objective 3), we will certainly make EDOAL evolve.* API improvements have indeed been made. The main progress has been the evolution of EDOAL towards expressing link keys (§2.5.3, p. 9) and SPARQL query generation from link keys (which was in the next topic). The integration of algebras of relations is only partially done.

Alignment services for ambient computing *We will further improve the Alignment server and mobile library so that applications could use them easily. In particular, we want to better integrate SPARQL query generation and manipulation from EDOAL.* Most of these topics have been covered by the previous item. We also developed a remarked framework for mobile semantic web including an Android version of the Alignment API [24, 47] which is not further reported here. This work had stopped for focussing on our core competences.

Evaluation of matchers *The work on matching evaluation shall continue through SEALS and OAEI.* The evaluation of matchers has been tremendously successful thanks to the SEALS project (§4.3, p. 23). *We are also working on automatic generation of new test sets allowing to generate networks of aligned ontologies.* In particular, our work on test generation is integrated in the Alignment API and routinely used in OAEI evaluations (§2.4.3, p. 6).

Exploitation and data interlinking

Data interlinking *We are planning to apply and extend our expertise in ontology matching to linked data co-reference resolution, i.e. for determining whether different URIs refer to the same entity.* This has undoubtfully been the largest change in our activities, to the extent that this sub-objective is now one of our main activities. We have developed similarity-based, link key-based, import-by-query and cross-lingual techniques for data interlinking (§2.5.3, p. 9). In particular, the work developed around the notion of link keys is deeply original.

Query generation from EDOAL *Hence, we need to provide adapted EDOAL to SPARQL transformations that can be used for extracting data and eventually generating links.* This has been successfully implemented in the Alignment API. Thanks to the concept of link key it is possible to generate SPARQL queries which, in turn, will generate links (§2.5.3, p. 9). *Another issue is related to minimality [...] It is also related to the problem of confluence in rewriting systems.* Such aspects finally have not been investigated.

Query containment *It is necessary to deal with query or correspondence comparison. We are currently investigating the static analysis of RDF manipulation in cooperation with the WAM team.* This work has been thoroughly investigated in the framework of various combination of query and ontology languages (§2.6.3, p. 11).

Other issues

Better matchers *We would like to build better matchers [...] However, this can only be achieved if substantial resources are dedicated to these tasks.* As anticipated, this topic was subject to manpower availability that we did not have. However, we have successfully further developed our work on context-based matching which was one of these aspects (§2.4.3, p. 6).

Applications *We would like to contribute to the adoption and dissemination of semantic web technology more broadly. In fact, we would like to deliver it in everybody's pocket.* We have invested quite a lot of effort in this transversal topic: developing a mobile semantic platform (see above: "Alignment services for ambient computing"), promoting semantic technologies for smart cities through the Ready4SmartCities project (§4.3, p. 23), and above all outreaching and writing proposals. We now resolved to concentrate our efforts on topics on which we can make the difference.

This should keep us busy for the next four years. After this, we expect that the most challenging problems to be considered will revolve around the dynamics of networks of ontologies (how they preserve coherence while evolving).
Here we are...

2.4 Objective 1: Ontology matching and alignments

Ontology matching is the task of finding correspondences between the entities of two ontologies [45]. Its result is expressed as an alignment. An alignment is a set of relations, e.g., equivalence, subsumption, between entities, e.g., concepts, from two different ontologies. We have pursued our work in various directions: evaluation of ontology matchers, semantics of alignments, extension of alignment languages and experimentation with matching through intermediate ontologies.

2.4.1 Personnel

Permanent researchers: Jérôme Euzenat, Jérôme David, Manuel Atencia
Post-doctoral researchers: Cássia Trojahn dos Santos$^{\rightarrow 2012-02}$, Jose-Luis Aguirre$^{2012-01\rightarrow 2012-10}$, Luz-Maria Priego$^{2013-12\rightarrow 2014-05}$
PhD students: Armen Inants$^{2012-12\rightarrow}$

2.4.2 Project-team positioning

The ontology matching field is still under active development [11]: we have a large attendance to the ontology matching workshop and the OAEI evaluations. We managed to deliver the second edition of our Ontology matching book (6 years after the first one) [45]. EXMO had a strong experience in ontology matching. In our research, we cover the whole ontology matching spectrum: foundations, tool support, matchers and evaluation. As mentioned before, we had not sufficient resource to invest in developing a new standalone ontology matcher.
The evaluation activity has been increasingly cooperative, hence we have been at the centre of the main evaluation efforts with our colleagues from Mannheim, Oxford, Trento, and others.
In contrast, our work on extending alignments languages towards more expressivity (EDOAL, algebras of relations) is quite unique.
Several semantics of alignments have been proposed: with common domains as in database, with disjoint domains as in ϵ-connections (U. Liverpool), with relations between domains as in DDL (U. Trento) and by equalisation to a common domain like in our work. It is well acknowledged that they are alternatives rather than competitor and subsequent efforts, including ours, are developed so that they could accommodate them all. One of the major competitors on foundations and tool

support is Till Mossakovski's group (Magdeburg U.). We have competing tools and server and complementary theories supporting them. They have integrated our Alignment API in the last version of their OntoHub framework.
The NCBO Bioportal (Stanford U.), targeting more specifically bioinformatics, and Synopsis (U. Trento) are competing alignment servers.

2.4.3 Scientific achievements

We have pursued our work on ontology matching in different directions.

Evaluation and test generation Evaluation of ontology matching algorithms involves to confront them with test ontologies and to compare results. For assessing the degree of achievement of current ontology matching algorithms we have co-organised yearly, since 2004, evaluation events and set up the Ontology Alignment Evaluation Initiative [27, 12, 22, 26]. Year after year, we have seen the field as a whole evolve towards better tools [52].
During the evaluation period, we contributed to automating ontology matching evaluation in the framework of the SEALS project (see §4.3, p. 23). This involved:

- designing and developing a web-based evaluation service that allows for participants to run their own evaluations and manipulate the results in a direct feedback cycle [40, 74];
- supporting participants to implement the standard interface for running their tools in the evaluation service [72, 54].

We introduced new benchmarks for testing multilingual ontology matching [10].
We have also provided a modular test generation framework enabling to generate ontology matching tests from different seed ontologies and with different levels of difficulty. We showed that we were able to reproduce our own OAEI Benchmark results both with the original seed ontology and with other ontologies. We also assessed experimentally the properties of these tests.
The work on test generation has been published in *Journal of web semantics* [6]; the work on multilingual evaluation as well [10].

Weighted alignment semantics Alignment correspondences are often assigned a weight or confidence factor by matchers. Nonetheless, few semantic accounts have been given so far for such weights. We have proposed a formal semantics for weighted correspondences between different ontologies [14]. It is based on a classificative interpretation of correspondences: if o and o' are two ontologies used to classify a common set X, then alignments between o and o' are interpreted as encoding how elements of X classified in the concepts of o are re-classified in the concepts of o', and weights are interpreted as measures of how precise and complete re-classifications are. This semantics is justifiable for extensional matchers. We have proven that it is a generalisation of the semantics of absolute correspondences, and we have provided properties that relate correspondence entailment with description logic constructors.
This work has been presented at ISWC 2012 [14].

Context-based matching Context-based matching finds correspondences between entities from two ontologies by relating them to other resources [45]. We designed a general view of context-based matching by analysing such existing matchers [9]. This view is instantiated in a path-driven approach that (*i*) anchors the ontologies to external ontologies, (*ii*) finds sequences of entities (path) that relate entities to match within and across these resources, and (*iii*) uses algebras of relations for combining the relations obtained along these paths. Parameters governing such a system were identified and made explicit.

We conducted experiments with different parameter configurations in order to assess their influence. In particular, experiments confirm that restricting the set of ontologies reduces the time taken at the expense of recall and F-measure. Increasing path length within ontologies increases recall and F-measure as well. In addition, algebras of relations (see below) allow for a finer analysis, which shows that increasing path length provides more correct or non precise correspondences, but marginally increases incorrect correspondences.
This work has been published in *Journal on data semantics* [9].

Alignment algebras We have shown previously that algebras of relations are useful in managing ontology alignments. They make it possible to aggregate alignments disjunctively or conjunctively and to propagate alignments within a network of ontologies. The previously considered algebra of relations contains taxonomical relations between classes only. We have tackled the problem of combining two or more calculi over disjoint universes into a single calculus [33]. The problem is important because ontology matching deals with various kinds of ontological entities: concepts, individuals, properties. We have defined the notion of heterogeneous qualitative calculus based on an algebraic construct called Schröder category. A Schröder category is to binary relations over heterogeneous universes what a relation algebra is to homogeneous ones. We have established the connection between homogeneous and heterogeneous qualitative calculi by defining two mutually inverse transition operators. We have designed an algorithm for combining two homogeneous calculi with different universes into a single calculus. This has been applied to alignment relations [33]. It is, first, able to deal with empty classes, and, second, incorporates all qualitative taxonomical relations that occur between individuals and concepts, including the relations "is a" and "is not". We have proved that this algebra is coherent with respect to the simple semantics of alignments.
The last part of this work will be presented at ISWC 2015 [33].

2.4.4 Collaborations

The work on context-based matching has been achieved thanks to a collaboration with Angela Locoro (U. Genova).

The semantics for weighted alignments is the result of joint work with Luciano Serafini, Chiara Ghidini (FBK Trento) and Alex Borgida (Rutgers U.).

The work on evaluation is that of a whole community which has been involved in SEALS and the OAEI effort. We can name: Christian Meilicke, Heiner Stuckenschmidt, Dominique Ritze (U. Mannheim), Ernesto Jimenez-Ruiz (U. Oxford), Patrick Lambrix (U. Linköping).

2.4.5 External support

The FP7 project SEALS (§4.3, p. 23) has been a strong support for the work on matcher evaluation;

The FP7 Ready4SmartCities project (§4.3, p. 23) supported the further development of the Alignment API and server and promoting this technology.

2.4.6 Self assessment

The main objectives that we put forth four years ago have been met. Evaluation still receives a lot of attention (and the software developed in the SEALS project is still improved), though we have progressively reduced our activity on this topic. The work on foundations (weighted semantics and algebra combination) as well as the work on context-based matching has been very strong. A couple of papers still have to be published. Although we may come back on some ontology

matching issues if necessary, we consider that the planned EXMO work program has been achieved and we will mostly dedicate our work to other topics.

2.5 Objective 2: Data interlinking

The web of data uses semantic web technologies to publish data on the web in such a way that they can be interpreted and connected together [70]. It is thus critical to establish links between these data, both for the web of data and for the semantic web that it contributes to feed. We are working on different techniques for interlinking data using keys extraction, rule processing, machine learning, and natural language processing.

2.5.1 Personnel

Permanent researchers: Manuel Atencia, Jérôme David, Jérôme Euzenat
PhD students: Zhengjie Fan$^{\rightarrow 2014-03}$, Tatiana Lesnikova$^{2012-10\rightarrow}$, Adam Sanchez$^{2013-12\rightarrow}$

2.5.2 Project-team positioning

Linked data interlinking is a relatively new field. It can benefit of our experience in ontology matching for two reasons: (*i*) both problems share features and solutions, (*ii*) data interlinking can use alignments to reduce their search space. Hence, our investment in the field is quite natural. However, there are problems specific to data interlinking such as the size of data sources.
Our outstanding main competitor in the domain is likely the team of Axel-Cyrille Ngonga Ngomo (U. Leipzig) who has developed numerous techniques for data interlinking. Other involved groups are Chris Bizers's (U. Mannheim), from which we use the Silk tool for similarity-based data interlinking. One of our specificity, which applies to most of our work, is to privilege first symbolic approaches, eventually supported by numeric techniques. In France, Nathalie Pernelle, Fatiha Saïs and colleagues (LRI, Orsay) have developed key discovery approaches.
One of the strengths of the team is the development of original approaches based on link keys, and their integration in our EDOAL alignment language. No comparable approach has been developed (keys require the assumption of their preservation across data set, this is not the case of link keys).

2.5.3 Scientific achievements

Our contributions on data interlinking proposes new interlinking techniques and take advantage of alignments in data interlinking.

Key-based data cleansing We have proposed a method for analysing web datasets based on key dependencies. Keys are sets of properties which uniquely identify individuals (instances of a class). We have refined the notion of database keys in a way which is more adapted to the context of description logics and the openness of the semantic web [15].
In order to better deal with web data of variable quality, we have introduced the definition of pseudo-keys [18]. We have also designed and implemented an algorithm for discovering pseudo-keys. Experimental results show that, even for a large dataset such as DBpedia, the runtime of the algorithm is still reasonable [25]. This work has allowed to detect automatically duplicates within wikipedia.
The work on pseudo-keys has been presented at EKAW 2012 [18]. The work on different types of keys has been presented at ICCS 2014 [15].

Link keys However, ontologies do not necessarily come with key descriptions and they may reveal useless when interlinking data. We have refined the notion of link keys introduced in [45] in weak, plain and strong link keys. Then, we have shown how such link keys may be used for deducing equality statements (links) between individuals across data sources in the web of data.
We have designed an algorithm to generate first a small set of candidate link keys [16]. Depending on whether some of the, valid or invalid, links are known, we defined supervised and non supervised measures for selecting the appropriate link keys. The supervised measures approximate precision and recall on a sample, while the non supervised measures are the ratio of pairs of entities a link key covers (coverage), and the ratio of entities from the same data set it identifies (discrimination). We have experimented with these types of measures, showing the accuracy and robustness of both.
This approach has been adapted to the simpler context of relational databases, and we have shown how candidate link keys can be encoded in the formal concept analysis framework [17]. We are pursuing this work with full link keys.
Link keys have been implemented within the Alignment API, i.e., they can be associated to class correspondences. It is then possible to automatically generate SPARQL Construct queries which generate links between entities.
The work on link keys has been presented at ECAI 2014 [16].

An import-by-query algorithm for data interlinking We have proposed a rule-based approach to infer all *certain* sameAs and differentFrom facts that are logically entailed from a given set of domain constraints and facts. Our main contribution is a novel algorithm, called Import-By-Query, that enables the scalable deployment of such an approach in the decentralised setting of linked data [13]. The main challenge is to identify the data, possibly distributed over several datasets, useful for inferring sameAs and differentFrom statements of interest. For doing so, Import-By-Query alternates steps of sub-query rewriting and of tailored querying the linked data cloud in order to import data as specific as possible for inferring or contradicting given target sameAs and differentFrom statements. It is an extension of the well-known query-subquery algorithm for answering Datalog queries over deductive databases. Experiments conducted on a real-world dataset have demonstrated the feasibility of this approach and its usefulness in practice for data linkage and disambiguation.
Additionally, and in line with the problem of dealing with uncertainty in linked data, we have proposed a probabilistic mechanism of trust that allow peers in a semantic peer-to-peer network to select the peers that are better suited to answer their queries, when query reformulation based on alignments may be unsatisfactory due to unsoundness or incompleteness of alignments [4].
The import-by-query approach is the result of the PhD thesis of Musltafa Al Bakri. It has been presented at AAAI 2015 [13]. The processing of trust in networks of ontologies has been published in *Knowledge and Information Systems* [4].

Learning similarities for data interlinking In the context of the Datalift project, we have developed a data interlinking module which generates data interlinking scripts from ontology alignments [62]. For that purpose, we have integrated existing technologies within the Datalift platform [32, 41]: the Alignment API (§3.2, p. 15), for taking advantage of the EDOAL language, and Silk, developed by Frei Universtität Berlin, for processing linking scripts. We have further developed an algorithm able to determine potential attribute correspondences of two classes depending on their features. For that purpose, we use k-means or k-medoïds clustering to identify groups of properties which can be compared. This provides property correspondences used to construct a Silk script which generates an initial link set. Some of the links are presented to users who assess their validity. We then use an improvement of the disjunctive version space supervised learning

method to learn a better script from the assessed links. Such a technique can be iterated until fully satisfactory links are found.
This work is the result of the PhD thesis of Zhengjie Fan [2] and has been presented at Web intelligence 2014 [31].

Cross-lingual data interlinking Another key challenge of linked data is to be able to discover links across datasets when entities are described in different natural languages. Since it is not possible to rely on direct name comparison, more global measures must be considered. In that context, we are developing an approach which represents RDF entities as (virtual) text documents and compare them using different strategies [36, 37].
In order to assess the quality of possible measures, we are currently following two directions: (*i*) a translation-based approach where the virtual documents are automatically translated [34]; (*ii*) a language-independent approach where important terms found in documents are mapped to a terminological resource [35], viz., BabelNet, to compute document similarity. We found that the former approach is more efficient on both generic and individual entities. However, the necessary depth to create virtual documents has an influence on the result quality. We conjecture that this depends on the generic-specific character of entities and have worked on testing this hypothesis.
This work has been partially published in ACM DocEng 2015 [35].

2.5.4 Collaborations

The work on key detection was started with François Scharffe when he was in the team and pursued after he joined LIRMM. It has been further investigated with Michel Chein, Michel Leclère and Madalina Croitoru (GRAPHIK, Montpellier) and Fatiha Saïs and Nathalie Pernelle (LRI, Orsay) in the context of the ANR Qualinca project (to which EXMO does not participate).

The import-by-query algorithm for data interlinking is part of the thesis work of Mustafa Al Bakri, supervised by Marie-Christine Rousset and Manuel Atencia in a joint collaboration with the SLIDE team (LIG).

Cross-lingual data interlinking is performed in the context of the ANR Lindicle project in collaboration with Juanzi Li (Tsinghua U.) and Zhichun Wang (Beijing Normal U.).

We also collaborate with Gilles Sérasset and Andon Tchechmedjiev (LIG Getalp team) to use their DBnary resource in cross-lingual data interlinking.

We have started working with Amedeo Napoli (INRIA-Orpailleur) on the FCA characterisation of link keys.

2.5.5 External support

The work on similarity learning has been performed in the context of the ANR Datalift project (§4.3, p. 22).

The work on multilingual data interlinking is performed in that of the ANR Lindicle project (§4.3, p. 22).

2.5.6 Self assessment

Data interlinking is a new and important topic. We have two particular strengths to claim: (*i*) our extensive knowledge of ontology matching and moreover the capacity to relate matching and interlinking, and (*ii*) the development of the concept of link keys. The import-by-query approach is complementary to the link key approach and it would be interesting to put them together.
On similarity learning and cross-lingual data interlinking, we made local contributions which provide a better understanding of what works for data interlinking. However, the current lack

of a comprehensive and accepted benchmarks [28] as well as the need to develop robust software requires workforce. We may want to join forces with other teams to share the load.
The work on link keys is a very original approach for which we have many results to come. We know what remains to be achieved in order to unveil its full potential (in particular link key co-extraction) and plan to push it to the limit.

2.6 Objective 3: Querying the semantic web

We had developed work on path-based RDF querying and using it for evaluating SPARQL in presence of RDF Schema [3]. The WAM (now TYREX) team has a solid experience on μ-calculus interpretation of XPath. We worked together on static analysis of SPARQL queries over RDF (and RDFS) and, more precisely, query containment.

2.6.1 Personnel

Permanent researcher: Jérôme Euzenat
PhD student: Melisachew Wudage Chekol$^{\rightarrow 2013-12}$

2.6.2 Project-team positioning

We have been among the first to formulate the query containment problem for SPARQL and moreover to provide concrete and proved solutions for it. The competition in this matter comes from Claudio Guttierez (U. Chile), Marcelo Arenas, Jorge Perez (PUC Chile). They introduced the slightly different notion of query subsumption that we have not considered (but could) and restricted their work on different fragments of SPARQL, e.g., avoiding projection. We have independently provided complementary results, none covering the whole SPARQL spectrum. More work is now developed, specifically on worst-case complexity issues. This work is somewhat connected to ontology-based data access. However, we tend to consider the problems by extending the expressiveness of query and ontology languages, to the expenses, sometimes, of incompleteness, instead of reducing it. Reducing the expressiveness is quite important to actually perform (data-related) querying, though we concentrate on (data-independent) static analysis.

2.6.3 Scientific achievements

Query containment modulo schemas We study query containment, i.e., determining whether, for any graph, the answers to a query are contained in those of another query. This problem is very important for query optimisation purposes, and will be even more in case of distributed queries. In the SPARQL context, it can be equally useful for distributing federated queries or for implementing schema-based access control.
We have reduced SPARQL query containment to satisfiability in the μ-calculus. To that extent, we proposed an encoding of RDF graphs as labelled transition systems and SPARQL queries and ontologies as propositional μ-calculus formulas. They allow to translate query evaluation to graph traversing through the modalities of the logic. We have proved the correctness of the encoding.
It is then possible to use solvers of this logic to test query containment of SPARQL queries under RDFS [19] and OWL schema [20] constraints, with paths or under particular entailment regimes. We have also implemented the proposed techniques on top of a general μ-calculus solver.
This work is the result of Melisachew Wudage Chekol's PhD thesis [1]. It has been published in AAAI 2012 [20] and IJCAR 2012 [19].

Benchmarks for query containment In order to experimentally assess implementation strengths and limitations, we provided a first SPARQL containment test benchmark [21]. We studied the query demographics on DBPedia logs to design benchmarks for relevant query containment solvers. We tested available solvers on their domain of applicability on three different benchmark suites and found that (*i*) tested solutions are overall functionally correct, (*ii*) in spite of its complexity, SPARQL query containment is practicable for acyclic queries, (*iii*) state-of-the-art solvers are at an early stage both in terms of capabilities and implementation.
This work has been published in ISWC 2013 [21]. The benchmarks, results and software are available at http://sparql-qc-bench.inrialpes.fr.

2.6.4 Collaborations

The work on query containment has been developed in collaboration with Nabil Layaïda and Pierre Genevès (TYREX).

2.6.5 External support

None.

2.6.6 Self assessment

The work on semantic web querying has always been a side work with respect to EXMO, except when we use it for translating alignments. It has provided worthwhile results which are taken over by work from other teams. Work on query containment is now pursued in the TYREX project.

2.7 Objective 4: Dynamics of the semantic web

Huge quantities of data described by ontologies and linked together are being made available. These are generated in an independent manner by autonomous providers such as individuals or companies. They are heterogeneous and their joint exploitation requires connecting them, ending up in a mesh of reticulated knowledge.
However, data and knowledge have to evolve facing changes in what they represent, changes in the context in which they are used and connections to new data and knowledge sources. As their production and exchange are growing larger and more connected, their evolution is not anymore compatible with manual curation and maintenance. We work towards their continuous evolution as it is critical to their sustainability. To that extent, we investigate methods for sparingly propagating changes, belief revision of networks of ontologies and applying cultural evolution techniques to the evolution of knowledge.

2.7.1 Personnel

Permanent researchers: Jérôme Euzenat, Manuel Atencia, Jérôme David
PhD students: Adam Sanchez Ayte$^{2013-12\rightarrow}$

2.7.2 Project-team positioning

There has been studies of management of local evolution, e.g., Lilljana Stojanovic (U. Karlsruhe). There are post-hoc statistical study of evolution done by Bijan Parsia and colleagues (U. Manchester) or more global statistical study of the evolution of the semantic web or linked data. We are trying to make a difference by considering together changes over data, ontologies, links and alignments. The topic of a holistic evolution of knowledge structure like the semantic web has not been studied to our knowledge.

Ontology revision using description logics has been studied by Giorgos Flouris (FORTH) without considering alignments. The topic of alignment repair has been recently implemented within three different systems (LogMap, Alcomo, AML). The thesis of Christian Meilicke (U. Mannheim) laid the foundations of such systems, but without direct relation to belief revision. Our approach to revision encompasses more complex alignment semantics and whole networks of ontologies whilst being expressed as belief revision.
Finally, cultural language evolution has been successfully pioneered by Luc Steels and his colleagues (Sony CSL and VU Brussels). The closer to ontologies that have been considered are lexicons organised in taxonomy. We are taking inspiration form this work to apply it to knowledge representation, which had not being attempted before. Language have different characteristics: we do not tackle symbol grounding, knowledge is not observable. We are in contact with the ESSENCE Marie Curie network which is involved in such research.

2.7.3 Scientific achievements

Different approaches are currently explored.

Evolution of ontology networks and linked data We are considering the global evolution of knowledge made of interdependent ontologies, data, alignments and links. Our goal is to be able to maintain this structure with respect to the processes which are involved in its construction: logical inference, ontology matching, link key extraction, link generation, etc. This can be put at work, for instance, in an Alignment server publishing ontology alignments induced from links in linked data. In turn, these alignments may be used to extract link keys which would provide more links. After each change, it is necessary to recover stability.
Our initial work is focused on how data changes cause alignment evolution. In this regard, we are developing techniques for circumscribing the elements affected by the change and evaluating the need for change propagation, i.e, most of the time a simple change will not trigger link key recomputation (§2.5.3, p. 9).

Revision in networks of ontologies Networks of ontologies are made of a collection of logic theories, called ontologies, related by alignments. They arise naturally in distributed and open contexts in which theories are developed and maintained independently, such as the semantic web. This calls for reconsidering problems, such as revision, in this context: given a set of ontologies connected by alignments, how to evolve them such that they account for new information. In networks of ontologies, inconsistency may come from two different sources: local inconsistency in a particular ontology or alignment, and global inconsistency between them. Belief revision is well-defined for dealing with ontologies; we have investigated how it can apply to networks of ontologies. We formulated revision postulates for alignments and networks of ontologies based on an abstraction of existing semantics of networks of ontologies. We showed that revision operators cannot be simply based on local revision operators on both ontologies and alignments. We adapted the partial meet revision framework to networks of ontologies and show that it indeed satisfies the revision postulates. Finally, we considered strategies based on network characteristics for designing concrete revision operators.
This work has been accepted for publication in *Artificial intelligence* [7].

Cultural alignment repair Alignments between ontologies may be established through agents holding such ontologies attempting at communicating and taking appropriate actions when communication fails. This approach, that we call cultural knowledge evolution, has the advantage of not assuming that everything should be set correctly before trying to communicate and of being able to overcome failures. We have tested this approach on alignment repair, i.e., the improvement

of incorrect alignments. For that purpose, we performed a series of experiments in which agents react to mistakes in alignments. Agents only know about their ontologies and alignments with others and they act in a fully decentralised way. We showed that cultural repair is able to converge towards successful communication through improving the objective correctness of alignments. The obtained results are on par with a baseline of state-of-the-art alignment repair algorithms.
Results have been presented at the 3rd ESWC workshop on Debugging ontologies and ontology mappings (WoDOOM) [30] (and selected as the best paper [46]). The benchmarks, results and software are available at http://lazylav.gforge.inria.fr.

2.7.4 Collaborations

Part of the work on cultural alignment revision is now developed further with Kate Revoredo and Fernanda Baião (Universidade Federal do Estado do Rio de Janeiro, BR).

2.7.5 External support

The work on evolution of ontology networks and linked data is performed in the ANR Lindicle project (§4.3, p. 22).

2.7.6 Self assessment

This work is at its beginning but it is, in our opinion, very important. The cultural evolution approach is very original and it provided promising results with alignment repair. It is definitely a direction in which we have a unique position and that we want to pursue further. The work on revision in networks of alignments supersedes largely the other approaches to alignment repair. The classical approaches of change propagation and belief revision should not be neglected as they can be combined with cultural evolution techniques.
This is the direction we want to put our efforts on in the forthcoming years (see §6, p. 24).

3 Knowledge dissemination

3.1 Publications

	2012	2013	2014	2015	Total
PhD Thesis	1		1		2
H.D.R.(*)					
Journal	2	2	2	2	8
Conference proceedings(**)	11	6	8	3	28
Book chapter	2		2	1	5
Book (written)		1			1
Book (edited)	3	1	1		5
Patent					
General audience papers					
Technical report	1	1	1		3
Deliverable	4	1	5	3	13

(*) HDR Habilitation à diriger des Recherches – (**) Conference with a program committee

Main journals (of the domain):

Journal of web semantics [6, 5, 10]
Journal on data semantics [9]

Semantic web journal
Artificial intelligence [7]
Journal of artificial intelligence research

Main conferences (of the domain):

International Semantic Web Conference (ISWC, ≈20%) [14, 21, 33]
European Semantic Web Conference (ESWC, ≈20%)
International Joint Conference on Artificial Intelligence (IJCAI, ≈15%)
European Conference on Artificial Intelligence (ECAI, ≈30%) [16]

We selected a few publications spanning all our four research directions and presenting substantial results of the period:

[16] Manuel Atencia, Jérôme David, Jérôme Euzenat, Data interlinking through robust linkkey extraction, in: Proc. 21st european conference on artificial intelligence (ECAI), Praha (CZ), pp15-20, 2014

[9] Angela Locoro, Jérôme David, Jérôme Euzenat, Context-based matching: design of a flexible framework and experiment, *Journal on data semantics* 3(1):25-46, 2014

[7] Jérôme Euzenat, Revision in networks of ontologies, *Artificial intelligence* 228:195-216, 2015

[14] Manuel Atencia, Alexander Borgida, Jérôme Euzenat, Chiara Ghidini, Luciano Serafini, A formal semantics for weighted ontology mappings, in: Proc. 11th International semantic web conference (ISWC), Boston (MA US), pp17-33, 2012

[19] Melisachew Wudage Chekol, Jérôme Euzenat, Pierre Genevès, Nabil Layaïda, SPARQL query containment under RDFS entailment regime, in: Proc. 6th International joint conference on automated reasoning (IJCAR), Manchester (UK), pp134-148, 2012

Most of our papers are available from http://exmo.inria.fr/papers/ or from http://hal.inria.fr.

3.2 Software

	A	SO	SM	EM	SDL	OC DA-CD-MS-TPM	loc
Alignment API	4	4	4	4	4	4	35k
OntoSim	3	4	3↘2	2	4	4	8k
Aroma	3	4	3↘2	2	4	4	4k
Melinda	2	4	1↗3	1↗3	2↗4	4	7k
Lazy lavender	2↗4	4	2	n/a	2↗4	4	2k

See https://semeval.inria.fr/37/previewInfo4Experts.

Alignment API The Alignment API and server[2] is composed of alignment formats, including the expressive alignment format EDOAL, an API for manipulating (generating, parsing, rendering, trimming, evaluating) these alignments, a library implementing this API and a server for sharing and storing alignments and networks of ontologies on the web. This API provides a high level of interoperability between systems providing and requiring alignments. It has been adopted by more than 50 other teams worldwide[3], both by team implementing matchers and teams manipulating them, and is used in the Ontology Alignment Evaluation Initiative (§2.4.3, p. 6). It is distributed since 2003 under the LGPL license and registered after the Software Protection Agency (IDDN.FR.001.050019.000.S.P.2011.000.20900). Current version is 4.7.

[2] http://alignapi.gforge.inria.fr
[3] http://alignapi.gforge.inria.fr/impl.html

***OntoSim* library** OntoSim[4] is an API dedicated to the computation of distances between ontologies and ontology entities. It contains all ontology-space and alignment-space measures that we introduced. It also provides measures used for matching ontologies and supports the development of new measures. In particular, it provides methods for aggregating similarity matrices. It also comes with a set of distances (string, objects, collections). It is written in Java, distributed since 2009 under the LGPL license and current version is 2.4.

Aroma Aroma[5] is an ontology matcher made of an association rule discovery algorithm and several string-based similarities. This algorithm has the originality to produce both equivalence and subsumption correspondences. It has the advantages to run fast and to scale to large ontologies (several thousands of concepts). It has been integrated in the ITM commercial software developed by Mondeca. Aroma is not further developed but maintained to work with newer libraries. Aroma is written in Java, distributed since 2009 under the LGPL license and current version is 1.2.

Melinda key and link keys discovery Melinda[6] is a set of tools allowing to extract pseudo-keys and candidate link keys from RDF data. Pseudo-Keys extractor allows to find all the keys and pseudo-keys having discriminability and support values greater than given thresholds. Link key discovery algorithm extracts all link key candidates between two given RDF datasets. It also provides for each candidate its discriminability and coverage, or estimations of precision and recall if a sample of sameAs links is provided. Both tools are written in Java and distributed under the LGPL License. Key extraction is also available through a web interface and service at http://rdfpkeys.inrialpes.fr

Lazy lavender Lazy lavender[7] is a Java framework for running randomised cultural knowledge evolution experiments. It allows one to design, run, generate output and reproduce experiments. The goal of the software is to test variations of experiments and to warrant reproducibility. It is used with a online laboratory logbook to record experiment results and refer to correct software versions. It is distributed through git cloning (no release) since 2014 under the LGPL and CeCILL-C licences.

Note that in 2011, we listed two further software (both in open source under similar conditions as above), that are not further developed by us:

PSPARQL Query evaluator, has been used in the best WWW 2012 paper[8] to show that it was the fastest implementation of path evaluation (this was neither the focus of the software, nor that of the paper).

IDDL Reasoner, is still developed, under the DRAOn name[9], by its main developer Chan Le Duc and colleagues (U. Paris 8).

3.3 Valorisation and technology transfer (Socio-economic impact and transfer)

The software that we develop is distributed under open source licenses, this facilitated adoption of this software. Besides cooperation with various companies in projects (ATOS, Mondeca, etc.), we mention two specific types of actions.

[4] http://ontosim.gforge.inria.fr
[5] http://aroma.gforge.inria.fr
[6] https://gforge.inria.fr/projects/melinda
[7] http://lazylav.gforge.inria.fr
[8] Marcelo Arenas, Sebastián Conca, Jorge Pérez: Counting beyond a Yottabyte, or how SPARQL 1.1 property paths will prevent adoption of the standard, Proc. WWW conference, Lyon (FR), pp629–638, 2012
[9] http://iddl.gforge.inria.fr/

Meaning engines

Meaning Engines[10] is a start-up company, founded in particular by former EXMO post-doctoral researcher François Scharffe, whose goal is to help improve the knowledge of corporate knowledge, e.g., catalogs, costumer data, through linked data principles (the application of semantic web technology for publishing data). Among their prospective costumers are music aggregators as well as banks. We have benefited from the support of Nicolas Guillouet for developing generic connectors based on our Alignment API. They introduce two novel features: using the notion of link keys to identify identical items in a data flow and performing hybrid integration which either identifies or creates objects from the incoming flows. In fact, hybrid integration is a type of knowledge evolution that provides new interesting research problems.

Ready4SmartCities

We also had a long-standing awareness raising activity in the domains of ambient intelligence, mobile computing and now smart cities. We are convinced that most of the semantic technologies developed for the web have a key role to play there. In particular, for dealing with heterogeneity. This activity went through various collaborations and realisations in these domains. However, it took a more concrete turn with the Ready4SmartCities support action (§4.3, p. 23). In that context, we have been advocating these technologies to urban planners, municipalities, building managers, energy providers, etc. We have demonstrated concrete matching technologies in the http://al4sc.inrialpes.fr alignment server. We have taught these technologies at the Ready4SmartCities "linked data for smart cities" summer school.

3.4 Teaching

Supervision of educational programs

Jérôme David has been coordinator, with Benoît Lemaire, of option "Web, Informatique et Connaissance" of Master "Ingénierie de la Cognition, de la Création, et des Apprentissages" (UPMF, UJF & INPG, 2011-2013 and 2014-2015);

Manuel Atencia and Jérôme David are coordinators of option "Web, Informatique et Connaissance" of Master "Ingénierie de la Cognition, de la Création, et des Apprentissages" (UPMF, UJF & INPG, since academic year 2015);

Jérôme Euzenat is, with Éric Gaussier, coordinator of the "AI and the web" option of the M2R in computer science and applied mathematics (UJF & INPG, during the whole period).

Tutorials

Post-graduate level: Jérôme Euzenat, "Ontology matching", 3h, European Summer School on Ontology Engineering and the Semantic Web, Cercedilla (ES), 2012, 2013

Post-graduate level: Jérôme Euzenat and Pavel Shvaiko, "Ontology matching", 3h, Tutorial ISWC, Riva del Garda (IT), 2014

Post-graduate level: Jérôme Euzenat, "Data interlinking", 3h+3h, Tutorial+Practical, "linked data for smart cities" summer school, Cercedilla (ES), 2015

Post-graduate level: Jérôme Euzenat, "Data interlinking", 1h, Tutorial, Web intelligence summer school, Saint-Étienne (FR), 2015

Post-graduate level: Jérôme Euzenat, "Dynamic interoperability: from ontology matching to cultural knowledge evolution", 3h, Tutorial, ESSENCE summer school, Edinburgh (UK), 2015

[10] http://meaningengines.com

Post-graduate level: Manuel Atencia, Jérôme David, Philippe Genoud, “What is this thing called linked data?”, 6h, Tutorial, ACM Document engineering symposium, Lausanne (CH), 2015

University courses

Name	Course title (short)	Level	Institution	Hours (eqTD)	Academic Years
Jérôme David	Introduction à Java	Licence ESSIG	UPMF	24h	2012
Jérôme David	Programmation Avancée	Licence 3 MASS	UPMF	20h	2012
Jérôme David	Introduction aux bases de données relationnelles	Licence 2 MASS	UPMF	8h	2012
Manuel Atencia	Introduction à la programmation fonctionnelle	Licence	UJF	36h	2012
Jérôme David	Outils Informatique (c2i)	Licence 1 Socio	UPMF	24h	2012, 2013
Jérôme David	Développement Mobile	Licence Pro. SIL	UPMF	30h	2012, 2013
Jérôme David	Algorithmique et programmation par objets	Licence 2 MASS	UPMF	70h	2014
Manuel Atencia	Introduction aux technologies du Web	Licence L3 MASS	UPMF	60h	2013, 2014
Jérôme David	Programmation Java 2	Master 1 IC2A	UPMF	30h	2012, 2013, 2015
Jérôme David	JavaEE	Master 2 IC2A	UPMF	30h	2012, 2013
Jérôme David	Interface Homme Machine 2	Master 2 IC2A	UPMF	30h	2012, 2013
Manuel Atencia, Jérôme David	Web Sémantique	Master 2 IC2A	UPMF	30h	2012
Jérôme David	Documents XML	Master 1 IC2A	UPMF	30h	2015
Manuel Atencia	Langages et technologies du Web 1	Master 1 IC2A	UPMF	30h	2013, 2014
Manuel Atencia	Langages et technologies du Web 2	Master 1 IC2A	UPMF	30h	2013, 2014
Jérôme David	Développement Web Mobile	Master 2 IC2A	UPMF	30h	2012, 2013, 2015
Manuel Atencia	Web Sémantique	Master 2 IC2A	UPMF	30h	2013, 2014
Jérôme David	JavaEE	Licence Pro. SIL	UPMF	25h	2015
Jérôme David	Introduction à Python	Licence ESSIG	UPMF	24h	2015
Jérôme Euzenat	Semantic web: from XML to OWL	Master 2 Info.	UJF & INPG	22h	2012, 2013, 2014, 2015

PhD theses

Our students are registered to the doctoral school on mathematics, information science and technology, informatics of the University of Grenoble.

PhD (defended: 2012-12-19): Melisachew Wudage Chekol, Graph queries for the semantic web, supervisors: Jérôme Euzenat and Nabil Layaïda

PhD (defended: 2014-04-04): Zhengjie Fan, Ontology-based data interlinking, supervisors: Jérôme Euzenat and François Scharffe

PhD (defended: 2014-12-15): Mustafa Al-Bakri, Uncertainty-sensitive reasoning over the web of data, supervisors: Marie-Christine Rousset and Manuel Atencia

PhD (in progress): Tatiana Lesnikova, Cross-lingual data interlinking, supervisors: Jérôme Euzenat and Jérôme David

PhD (in progress): Armen Inants, Ontology alignment algebra, supervisor: Jérôme Euzenat

PhD (in progress): Adam Sanchez Ayte, Ontology alignment and data interlinking evolution on the web of data, supervisors: Jérôme Euzenat and Jérôme David

Our previous doctoral students are associate professors[1] (Sébastien Laborie in Pau, Antoine Zimmermann at École des mines de Saint-Étienne and Faysal Alkhateeb at Yarmouk University) or engineer (Jérôme Pierson at Hurence). Some post-doctoral students in the team also found associate professor[1] positions (Cássia Trojahn in Toulouse, Chan Le Duc in Paris).

3.5 General audience actions

Jérôme Euzenat gave a presentation on *Library resources and the web of data: the missing link* at the Millenium club (group of librarians) days, Montbonnot (FR), 2012-03-28.

Jérôme Euzenat coordinated the INTech technology watch seminar on *L'ouverture des données: technologies et usages (Open data: technology and use)*, 7 presentations, 12 demonstrations, 80 participants, Montbonnot, 2012-06-05.

Jérôme Euzenat has been interviewed by Françoise Breton, *L'émergence d'un web des connaissances*, published on the INRIA Grenoble Rhône-Alpes web site, 2012-09-25.

Jérôme Euzenat gave a presentation on *Les données liées ouvertes (linked open data): publier des données réutilisables* at the open data days of the FREMIT Federation, IMT/IRIT, Toulouse (FR), 2013-11-05.

Jérôme Euzenat coordinated the CNRS development days (JDEV) sessions on *web programming*, 4 presentations, 6 tutorials, 5 workshops, 80 participants, Palaiseau, 2013-9-4–6.

Jérôme Euzenat gave a presentation on *Communication et adaptation: la fabrique de la communication flexible* au Cycle de conférences Canopé (ex-CRDP) ≪La culture partagée≫, Grenoble (FR), 2014-04-02

Jérôme Euzenat gave a talk to the 7es journées de l'interopérabilité des applications d'entreprise (JIAE) on "Aligner les ontologies pour communiquer (matching ontologies to communicate)", Saint-Étienne (FR), 2014-05-16.

Jérôme David participates to the working group ≪Créer et maintenir une interconnexion sémantique des grands référentiels culturels: le "Graphe Culture"≫ of the ministry of Culture[11].

Jérôme Euzenat gave a talk to the Grilog networking business meeting "Big data" on "Publication et exploitation des données avec les technologies sémantiques (Publishing and exploiting data with semantic technologies)", Meylan (FR), 2015-04-28.

[11] http://cblog.culture.fr/projet/2013/11/07/groupe-de-travail-metadonnees-culturelles/

3.6 Visibility

Prizes and distinctions

Jérôme Euzenat has been elected ECCAI fellow (2014)

Expertise and recruitment committees

Jérôme Euzenat has been reviewing grant demands for WWTF (AT), 2014.

Jérôme David has reviewed ANR Blanc bilatteraux, 2013.

Jérôme Euzenat has been evaluator for European projects (FP7, H5: Smart cities and sustainability communication networks, content and technology), 2012-2015

Jérôme Euzenat has been evaluator of "research products" for the Agenzia Nazionale di Valutazione del Sistema Universitario e della Ricerca (IT), 2012

Jérôme Euzenat has been evaluator for an Emergence research grant of the DigiCosme Labex, Saclay (FR), 2013

Jérôme Euzenat has been external expert for the CEDAR ANR Chair of Excellence project, Lyon (FR), 2014

Jérôme Euzenat has been scientific committee member for the TNO "Smart Appliances" study for the European commission, 2014

Jérôme David has been member of the recruitment committee of Université Pierre Mendès France for the associate professor position 27MCF291, 2012.

Jérôme Euzenat has been member of the recruitment committee of the University of Caen for the professor position 27PR2055, 2012.

Editorial boards

Jérôme Euzenat is member of the editorial board of *Journal of web semantics* (since 2015 area editor), *Journal on data semantics* and the *Semantic web journal* (for the whole period).

Jérôme Euzenat is founding member of the "Semantic Web Science Association" (steering committee for the ISWC conference series). He has been part of the committee until 2014.

Chairs

Jérôme David has been tutorial chair (with Irwin King) of the 21st International World Wide Web Conference (WWW 2012).

Jérôme Euzenat has been co-chairman (with Manfred Hauswirth and Josie Xavier Perreira) of the "Experiment and evaluation" track of the 11th "International Semantic Web Conference" (ISWC, Boston, MA US), 2012

Jérôme Euzenat was co-chair with Maria Keet of the "ontologies" area of the programme committee of the "European Semantic Web Conference (ESWC)", 2014

Programme committees

Jérôme Euzenat has been programme committee member for ISWC (2012, 2013, 2014, 2015), WWW (2012, 2013), IJCAI (2013, 2015), KR (2014), K-Cap (2011, 2013), Context (2013, 2015), ECAI (2014), AAAI (2012), ESWC (2013, 2014, 2015), EKAW (2012), FOIS (2012, 2014), ICCS (2014), ER (2012), OntoBras (2013, 2015), RFIA (2014), the Ingénierie des Connaissances conference (IC 2014)

Cássia Trojahn dos Santos has been programme committee member for the International Workshop on Evaluation of Semantic Technologies (iWest 2012), the Ontology Matching workshop (OM 2012).

Jérôme David has been programme committee member for ESWC (2013, 2014, 2015), MobiWIS (2012), IC (2013, 2014, 2015), the Ontology matching workshop (OM 2012, 2013, 2014, 2015), Artificial Intelligence meets the Web of Data workshop (2012) and Downscaling the Semantic Web workshop (2012), the Graph Based Structures for Knowledge Representation and Reasoning Workshop (GKR@IJCAI 2013).

Manuel Atencia has been programme committee member of ISWC (2013, 2014), IJCAI (2015), ECAI (2014), ESWC (2015), the Ontology matching workshop (OM 2013), the Graph Based Structures for Knowledge Representation and Reasoning Workshop (GKR@IJCAI 2013), IC (2014)

Tatiana Lesnikova has been programme committee member for the Ontology matching workshop (OM 2014).

Organisation

Jérôme Euzenat organises yearly (with Pavel Shvaiko) the "Ontology matching" workshop at the ISWC conference (2012-2015).

EXMO (and other colleagues from Oxford, Trento, Mannheim, Linköping, Milano, Amsterdam, Galway and the Open university) organises yearly the Ontology alignment evaluation initiative (OAEI 2012-2014). OAEI 2015 is now organised by Ernesto Jimenez Ruiz (Oxford U.).

Jérôme David is member of the organisation committee of the "Web of data summer school", Saint-Étienne (FR), 2014-2015.

4 Funding

4.1 Funding external to Inria

(k€)	2012	2013	2014	2015
National initiatives				
ANR Datalift	36	36	18	
ANR Lindicle		65	65	65
European projects				
FP7 INFRASTRUCTURE SEALS	82			
FP7 SA Ready4SmartCities		20	42	23
Industrial contracts: none				
Scholarships				
PhD (UJF)	15	30	40	15
Post Doc				
Other external funding: none				
Total	133	151	155	103

4.2 Inria competitive funding

(k€)	2012	2013	2014	2015
INRIA Research Initiatives: none				
Associated teams: none				
Scholarships				
Internship				
PhD[3]		40	40	40
Post Doc[4]				
Technological development: none				
Other Inria competitive funding				
ITI		8	34	34
Total		48	74	74

[1] INRIA Cooperative Research Initiative (ARC = Action de Recherche Cooperative).
[2] Large-scale Initiative Action (AE = Action d'Envergure nationale).
[3] INRIA doctoral research contract (CORDI-S, doctorat Inria sur subvention).
[4] INRIA postdoctoral research contract ('post-doc Inria sur subvention").
[5] Junior engineer supported by INRIA.
[6] Technological Development Action (ADT = Action de Développement Technologique).
[7] Software Development Operation (ODL = Opération Développement Logiciel). Last ODLs ended in 2010.

4.3 Main contracts and grants

National initiatives

ANR Datalift EXMO has coordinated with LIRMM the DATALIFT project which has produced a platform for publishing governmental data as linked data. EXMO was particularly involved in the generation of links between datasets (see §2.5, p. 8).

- Project name: DATALIFT
- Program: ANR-ContInt
- Instrument: platform
- Duration: September 2010 – March 2014
- Coordinator: INRIA EXMO/François Scharffe
- Participants: Jérôme Euzenat, Zhengjie Fan, Jérôme David
- Other partners: LIRMM, INRIA-Wimmics, Eurecom, ATOS, Mondeca, INSEE, IGN, FING
- Web site: http://www.datalift.org

ANR Lindicle EXMO carries out with Tsinghua university (Knowledge engineering group), Beijing (CN), the Lindicle project whose goal is to investigate crosslingual data interlinking between French, English and Chinese data sources (see §2.5, p. 8).

- Project name: Linking data in cross-lingual environment
- Program: ANR-Blanc international 2
- Duration: January 2013 – December 2016
- Coordinator: INRIA EXMO/Jérôme David
- Participants: Jérôme Euzenat, Manuel Atencia Arcas, Jérôme David, Tatiana Lesnikova, Adam Sanchez Ayte
- Other partners: Tsinghua university (CN)
- Web site: http://lindicle.inrialpes.fr

European projects

FP7 SEALS (2009-2012) EXMO was a partner of the EU infrastructure project SEALS whose goal is to provide the infrastructure for evaluating semantic technologies. Jérôme Euzenat was vice-coordinator in charge of the research area and Cássia Trojahn led the work package on ontology matching evaluation. More particularly, EXMO has provided an infrastructure for evaluating ontology matching systems and algorithms, integrated in the SEALS platform.

- Project name: Semantic Evaluation At Large Scale
- Type: CAPACITIES (Infrastructures)
- Call: Scientific Data Infrastructure
- Instrument: Combination of COLLABORATIVE PROJECTS and COORDINATION and SUPPORT ACTIONS (CPCSA)
- Duration: June 2009 – June 2012
- Coordinator: Universidad Politecnica de Madrid (ES)
- Others partners: Innsbruck (AT), STI (AT), U. Sheffield (UK), U. Zürich (CH), U. Mannheim (DE), U. Oxford (UK), TZI (DE)
- Web site: http://seals-project.eu

FP7 Ready4SmartCities: EXMO is a partner of the FP7 coordination and support action Ready4SmartCities which aims at increasing awareness and interoperability for the adoption of ICT and semantic technologies in energy system to obtain a reduction of energy consumption and CO_2 emission at smart cities community level through innovative relying on RDF and innovation outcomes and ICT-based solutions. We are interested in promoting the use of semantic technologies in this context. We are more particularly in charge of displaying the benefits of maintaining alignments between available ontologies.

- Project acronym: Ready4SmartCities
- Project name: ICT Roadmap and Data Interoperability for Energy Systems in Smart Cities
- Type: CAPACITIES
- Call: ICT-2013.6.4 - Optimising Energy Systems in Smart Cities
- Instrument: Coordination and Support Action
- Objective: Optimising Energy Systems in Smart Cities
- Duration: October 2013 – September 2015
- Coordinator: D'appolonia Spa (IT)
- Others partners: Universidad Politecnica de Madrid (ES), CSTB (FR), CERTH (GR), VTT (FI), AIT (AT), AEC3 (UK), Politecnico di Torino (IT), Empirica (DE)
- Inria contact: Jérôme Euzenat
- Participants: Jérôme Euzenat, Luz Maria Priego-Roche, Jérôme David, Adam Sanchez Ayte
- Web site: http://www.ready4smartcities.eu

Industrial contracts: none

ARCs: none

AEs: none

Associated teams and other international projects: none

Other funding

Jérôme Euzenat is benefiting from a Special visiting researcher grant from the Brazilian Ciência sem Fronteiras program on "Methodology and algorithms for ontology refinement and matching" (2015-2017). He will be working with the team of Fernanda Baião and Kate Revoredo at the Universidade Federal do Estado do Rio de Janeiro (UNIRIO). Together they will investigate methods for evolving ontologies and alignments which involve users and agents. The goal of the project is to design methods and algorithms for both revising ontologies to represent the evolution of knowledge in a reliable manner and obtaining better quality alignments.

5 Follow-up of the previous evaluation

Here is the recommendation from the previous evaluation:

> **Recommended actions:** We strongly recommend to extend the research team with additional personnel, both from INRIA and by involving university researchers. While this should be done in a way that supports the aims of the project, we recommend to use recruiting also to enrich the general background of the project e.g. by employing a researcher from LOD or with a more formal background in knowledge representation (such as description logic). As noted above, we also recommend to carefully broaden the scope of the project at this point, into the direction of LOD. EXMO should try to build a joint project with one of the key open data actors in France such as Data Publica and ETALAB.

These three recommendations have been implemented as far as possible:

- The hiring of Manuel Atencia on a lecturer position can be considered as a fulfilment of the first recommendation. However, we deplore that we have not been able to recruit at INRIA in spite of highly qualified candidates.
- The project scope has definitely been broadened in the indicated direction as a large part of our activity has been indeed devoted to data interlinking (see §2.5, p. 8), a subtopic of Linked Open Data (LOD).
- We have worked within the Datalift project with companies like ATOS and Mondeca. We are transferring part of our results, and in particular the generation of links from link keys, to the Meaning engine start-up company (see §3.3, p. 17).

6 Objectives for the next four years

Self-assessment

The semantic web is nowadays a gigantic network of ontologies related together by alignments expressing data which are linked together across datasets. It is successful in many respects: huge amounts of data are now published openly, linked and exploited at web scale; key web players (e.g., Google, Facebook) are now exploiting the same technologies to create very large knowledge

graphs for answering search queries or recommending products to consumers and consumers to companies.
EXMO has taken its share in the development of semantic web technologies. The strength of the EXMO team is its focus at the edge of a handful of topics. EXMO has been successful and impactful on ontology matching these past years. The new data interlinking topic is attracting interest and generating new ideas. We have, in the last period, maintained our global volume of publication but largely reinforced journal publications at a very good level (AIJ, JWS, JoDS, TKDE, KAIS, MTAP). Globally, the team, of now three permanent people, is well regarded and the PhD students that we have trained have reached a good visibility.
The structure of the semantic web, inherited from its logical foundations, is relatively rigid. However, it has to be able to evolve, otherwise it will become an obstacle to communication rather than a facilitator. One important question of the coming years is thus to provide this flexibility and fluidity to the semantic web by serendipitously allowing stakeholders to communicate efficiently through it.

Perspectives for the research team

The EXMO project is about to end and we plan to start a new project. The semantic web adventure was in part transferring knowledge that had been acquired in the previous decades. Now that this technology has been embraced, it is time for us to address a new ambitious cutting-edge program. We will propose a new project which considers how agents (people or programs) may effectively communicate though having heterogeneous views of the world, e.g., because they have different experiences. For that purpose, we will adopt the cultural evolution strategy in which agents evolve their knowledge through continuously interacting in a specific environment. This provides a situated basis, and in this case an interaction basis, to knowledge evolution. This does not necessarily require to drop the logical heritage, which is very useful for drawing safe conclusions and highlighting conflicts. This certainly should make it more alive: able to adapt when it becomes an obstacle, but able to do it rationally, i.e., by taking into account the reasons for being so and the reasons for conflicting. Instead of creating the ultimate ontology, our objective is the observation of an *unstable* and dynamic knowledge structure.
This long-term research program raises many questions such as:

- By which mechanisms, is it possible to articulate symbolic and interaction-based techniques so that they show convergence to a representation enabling successful communication?
- Under which conditions and with which mechanisms can multiple knowledge models coexist? I.e., what are the models and in which contexts, is it possible that they do not merge?
- Which mechanisms provide which level of interoperability among agents?
- How such models can gradually evolve or co-evolve when facing environment change?

We plan to tackle these problems from three different investigation angles that complement each others:

- A theoretical approach which studies the mechanisms by which it is possible to evolve formal knowledge.
- An experimental approach based on the theories of cultural evolution showing how, by continuously trying and failing to communicate, agents converge towards mutually agreeable states, i.e., efficient for communicating;
- Eventually, a social science approach in which observed social communication models could be used as inspiration and as reference for human and social plausibility.

This research program is fundamental; however, in a world where most artefacts, including scientific theories, are designed by teams from different horizons, it could be applied to supporting this design process.

7 Bibliography of the project-team

This list is disjoint from that of the previous evaluation. Most of our papers are available from http://exmo.inria.fr/papers/.

Doctoral Dissertations and Habilitation Theses

[1] M. CHEKOL, Static analysis of semantic web queries, Thèse d'informatique, Université de Grenoble, Grenoble (FR), 2012, [hal:tel-00834448].

[2] Z. FAN, Concise pattern learning for RDF data sets interlinking, Thèse d'informatique, Université de Grenoble, Grenoble (FR), 2014, [hal:tel-00986104].

Articles in International Journal

[3] F. ALKHATEEB, J. EUZENAT, Constrained regular expressions for answering RDF-path queries modulo RDFS, *International Journal of Web Information Systems 10*, 1, 2014, p. 24–50, [doi:10.1108/IJWIS-05-2013-0013], [hal:hal-00975283].

[4] M. ATENCIA, M. AL-BAKRI, M.-C. ROUSSET, Trust in networks of ontologies and alignments, *Knowledge and Information Systems 42*, 2, 2015, p. 353–379, [doi:10.1007/s10115-013-0708-9], [hal:hal-01171782].

[5] M. ATENCIA, M. SCHORLEMMER, An interaction-based approach to semantic alignment, *Journal of web semantics 13*, 2012, p. 131–147, [doi:10.1016/j.websem.2011.12.001], [hal:hal-00768411].

[6] J. EUZENAT, M. ROSOIU, C. TROJAHN DOS SANTOS, Ontology matching benchmarks: generation, stability, and discriminability, *Journal of web semantics 21*, 2013, p. 30–48, [doi:10.1016/j.websem.2013.05.002], [hal:hal-00917906].

[7] J. EUZENAT, Revision in networks of ontologies, *Artificial intelligence 228*, 2015, p. 195–216, [doi:10.1016/j.artint.2015.07.007], [hal:hal-01188763].

[8] S. LABORIE, J. EUZENAT, N. LAYAÏDA, Semantic adaptation of multimedia documents, *Multimedia tools and applications 55*, 3, 2011, p. 379–398, [doi:10.1007/s11042-010-0552-9], [hal:hal-00781024].

[9] A. LOCORO, J. DAVID, J. EUZENAT, Context-based matching: design of a flexible framework and experiment, *Journal on data semantics 3*, 1, 2014, p. 25–46, [doi:10.1007/s13740-013-0019-z], [hal:hal-00975284].

[10] C. MEILICKE, R. GARCÍA CASTRO, F. FREITAS, W. VAN HAGE, E. MONTIEL-PONSODA, R. RIBEIRO DE AZEVEDO, H. STUCKENSCHMIDT, O. SVÁB-ZAMAZAL, V. SVÁTEK, A. TAMILIN, C. TROJAHN DOS SANTOS, S. WANG, MultiFarm: A benchmark for multilingual ontology matching, *Journal of web semantics 15*, 3, 2012, p. 62–68, [doi:10.1016/j.websem.2012.04.001], [hal:hal-00768423].

[11] P. SHVAIKO, J. EUZENAT, Ontology matching: state of the art and future challenges, *IEEE Transactions on knowledge and data engineering 25*, 1, 2013, p. 158–176, [doi:10.1109/TKDE.2011.253], [hal:hal-00917910].

International Conference/Proceedings

[12] J. L. AGUIRRE, B. CUENCA GRAU, K. ECKERT, J. EUZENAT, A. FERRARA, W. VAN HAGE, L. HOLLINK, E. JIMÉNEZ-RUIZ, C. MEILICKE, A. NIKOLOV, D. RITZE, F. SCHARFFE, P. SHVAIKO, O. SVÁB-ZAMAZAL, C. TROJAHN DOS SANTOS, B. ZAPILKO, Results of the Ontology Alignment Evaluation Initiative 2012, *in : Proc. 7^{th} ISWC workshop on ontology matching (OM), Boston (MA US)*, P. Shvaiko, J. Euzenat, A. Kementsietsidis, M. Mao, N. Noy, H. Stuckenschmidt (editors), p. 73–115, 2012, [hal:hal-00768409].

[13] M. AL-BAKRI, M. ATENCIA, S. LALANDE, M.-C. ROUSSET, Inferring same-as facts from linked data: an iterative import-by-query approach, *in : Proc. 29th Conference on Artificial Intelligence (AAAI), Austin (TX US)*, B. Bonet, S. Koenig (editors), p. 9–15, 2015, [hal:hal-01113463].

[14] M. ATENCIA, A. BORGIDA, J. EUZENAT, C. GHIDINI, L. SERAFINI, A formal semantics for weighted ontology mappings, *in : Proc. 11th International semantic web conference (ISWC), Boston (MA US)*, p. 17–33, 2012, [doi:10.1007/978-3-642-35176-1_2], [hal:hal-00768413].

[15] M. ATENCIA, M. CHEIN, M. CROITORU, J. DAVID, M. LECLÈRE, N. PERNELLE, F. SAÏS, F. SCHARFFE, D. SYMEONIDOU, Defining key semantics for the RDF datasets: experiments and evaluations, *in : Proc. 21st International Conference on Conceptual Structures (ICCS), Iasi (RO)*, p. 65–78, 2014, [doi:10.1007/978-3-319-08389-6_7], [hal:lirmm-01090357].

[16] M. ATENCIA, J. DAVID, J. EUZENAT, Data interlinking through robust linkkey extraction, *in : Proc. 21st european conference on artificial intelligence (ECAI), Praha (CZ)*, T. Schaub, G. Friedrich, B. O'Sullivan (editors), IOS press, p. 15–20, Amsterdam (NL), 2014, [doi:10.3233/978-1-61499-419-0-15], [hal:hal-01179166].

[17] M. ATENCIA, J. DAVID, J. EUZENAT, What can FCA do for database linkkey extraction?, *in : Proc. 3rd ECAI workshop on What can FCA do for Artificial Intelligence? (FCA4AI), Praha (CZ)*, p. 85–92, 2014, [hal:hal-01180937].

[18] M. ATENCIA, J. DAVID, F. SCHARFFE, Keys and pseudo-keys detection for web datasets cleansing and interlinking, *in : Proc. 18th international conference on knowledge engineering and knowledge management (EKAW), Galway (IE)*, p. 144–153, 2012, [doi:10.1007/978-3-642-33876-2_14], [hal:hal-00768412].

[19] M. CHEKOL, J. EUZENAT, P. GENEVÈS, N. LAYAÏDA, SPARQL query containment under RDFS entailment regime, *in : Proc. 6th International joint conference on automated reasoning (IJCAR), Manchester (UK)*, p. 134–148, 2012, [doi:10.1007/978-3-642-31365-3_13], [hal:hal-00749087].

[20] M. CHEKOL, J. EUZENAT, P. GENEVÈS, N. LAYAÏDA, SPARQL query containment under SHI axioms, *in : Proc. 26th American national conference on artificial intelligence (AAAI), Toronto (ONT CA)*, AAAI Press, p. 10–16, Menlo Park (CA US), 2012, [hal:hal-00749080].

[21] M. CHEKOL, J. EUZENAT, P. GENEVÈS, N. LAYAÏDA, Evaluating and benchmarking SPARQL query containment solvers, *in : Proc. 12th International semantic web conference (ISWC), Sydney (NSW AU)*, p. 408–423, 2013, [doi:10.1007/978-3-642-41338-4_26], [hal:hal-00917911].

[22] B. CUENCA GRAU, Z. DRAGISIC, K. ECKERT, J. EUZENAT, A. FERRARA, R. GRANADA, V. IVANOVA, E. JIMÉNEZ-RUIZ, A. KEMPF, P. LAMBRIX, A. NIKOLOV, H. PAULHEIM, D. RITZE, F. SCHARFFE, P. SHVAIKO, C. TROJAHN DOS SANTOS, O. ZAMAZAL, Results of the Ontology Alignment Evaluation Initiative 2013, *in : Proc. 8th ISWC workshop on ontology matching (OM), Sydney (NSW AU)*, P. Shvaiko, J. Euzenat, K. Srinivas, M. Mao, E. Jiménez-Ruiz (editors), p. 61–100, 2013, [hal:hal-00918494].

[23] J. DAVID, J. EUZENAT, J. JUNG, Experimenting with ontology distances in semantic social networks: methodological remarks, *in : Proc. 2nd IEEE international conference on systems, man, and cybernetics (SMC), Seoul (KR)*, IEEE Computer society, p. 2909–2914, Los Alamitos (CA), 2012, [doi:ICSMC.2012.6378236], [hal:hal-00768419].

[24] J. DAVID, J. EUZENAT, M. ROSOIU, Linked data from your pocket, *in : Proc. 1st ESWC workshop on downscaling the semantic web, Hersounissos (GR)*, C. Guéret, S. Schlobach, F. Pigout (editors), p. 6–13, 2012, [hal:hal-00768418].

[25] J. DAVID, F. SCHARFFE, Détection de clefs pour l'interconnexion et le nettoyage de jeux de données, *in : Actes 23^{e} journées francophones sur Ingénierie des connaissances (IC), Paris (FR)*, p. 401, 2012, [hal:hal-00741734].

[26] Z. DRAGISIC, K. ECKERT, J. EUZENAT, D. FARIA, A. FERRARA, R. GRANADA, V. IVANOVA, E. JIMÉNEZ-RUIZ, A. KEMPF, P. LAMBRIX, S. MONTANELLI, H. PAULHEIM, D. RITZE, P. SHVAIKO, A. SOLIMANDO, C. TROJAHN DOS SANTOS, O. ZAMAZAL, B. CUENCA GRAU, Results of the Ontology Alignment Evaluation Initiative 2014, *in : Proc. 9th ISWC workshop on ontology matching (OM), Riva del Garda (IT)*, P. Shvaiko, J. Euzenat, M. Mao, E. Jiménez-Ruiz, J. Li, A. Ngonga Ngomo (editors), p. 61–104, 2014, [hal:hal-01180915].

[27] J. EUZENAT, A. FERRARA, W. VAN HAGUE, L. HOLLINK, C. MEILICKE, A. NIKOLOV, F. SCHARFFE, P. SHVAIKO, H. STUCKENSCHMIDT, O. SVÁB-ZAMAZAL, C. TROJAHN DOS SANTOS, Results of the Ontology Alignment Evaluation Initiative 2011, *in : Proc. 6th ISWC workshop on ontology matching (OM), Bonn (DE)*, P. Shvaiko, I. Cruz, J. Euzenat, T. Heath, M. Mao, C. Quix (editors), p. 85–110, 2011, [hal:hal-00781022].

[28] J. EUZENAT, A modest proposal for data interlinking evaluation, *in : Proc. 7th ISWC workshop on ontology matching (OM), Boston (MA US)*, P. Shvaiko, J. Euzenat, A. Kementsietsidis, M. Mao, N. Noy, H. Stuckenschmidt (editors), p. 234–235, 2012, [hal:hal-00768421].

[29] J. EUZENAT, Uncertainty in crowdsourcing ontology matching, *in : Proc. 8th ISWC workshop on ontology matching (OM), Sydney (NSW AU)*, P. Shvaiko, J. Euzenat, K. Srinivas, M. Mao, E. Jiménez-Ruiz (editors), p. 221–222, 2013, [hal:hal-00918495].

[30] J. EUZENAT, First experiments in cultural alignment repair, *in : Proc. 3rd ESWC workshop on Debugging ontologies and ontology mappings (WoDOOM), Hersounisos (GR)*, p. 3–14, 2014, [hal:hal-01006344].

[31] Z. FAN, J. EUZENAT, F. SCHARFFE, Learning concise pattern for interlinking with extended version space, *in : Proc. 13th IEEE/WIC/ACM international conference on web intelligence (WI), Warsaw (PL)*, D. l zak, H. Nguyen, M. Reformat, E. Santos (editors), IEEE Computer society, p. 70–77, Los Alamitos (CA), 2014, [doi:10.1109/WI-IAT.2014.18], [hal:hal-01180918].

[32] Z. FAN, Data linking with ontology alignment, *in : Proc. 9th European semantic web conference (ESWC), Heraklion (GR)*, p. 854–858, 2012, [doi:10.1007/978-3-642-30284-8_70], [hal:hal-00768422].

[33] A. INANTS, J. EUZENAT, An algebra of qualitative taxonomical relations for ontology alignments, *in : Proc. 14th International semantic web conference (ISWC), Bethleem (PA US)*, 2015, [hal:hal-01188792].

[34] T. LESNIKOVA, J. DAVID, J. EUZENAT, Interlinking English and Chinese RDF data sets using machine translation, *in : Proc. 3rd ESWC workshop on Knowledge discovery and data mining meets linked open data (Know@LOD), Hersounisos (GR)*, J. Völker, H. Paulheim, J. Lehmann, H. Sack, V. Svátek (editors), 2014, [hal:hal-01180919].

[35] T. LESNIKOVA, J. DAVID, J. EUZENAT, Interlinking English and Chinese RDF data using BabelNet, *in : Proc. 15th ACM international symposium on Document engineering (DocEng), Lausanne (CH)*, P. Genevès, C. Vanoirbeek (editors), p. 39–42, 2015, [doi:10.1145/2682571.2797089], [hal:hal-01188837].

[36] T. LESNIKOVA, Interlinking cross-lingual RDF data sets, *in : Proc. ESWC PhD symposium, Montpellier (FR)*, p. 671–675, 2013, [doi:10.1007/978-3-642-38288-8_49], [hal:hal-00917908].

[37] T. LESNIKOVA, NLP for interlinking multilingual LOD, *in : Proc. ISWC Doctoral consortium, Sydney (NSW AU)*, p. 32–39, 2013, [hal:hal-00918496].

[38] T. LESNIKOVA, Interlinking RDF data in different languages, *in : Proc. 4th workshop on Terminology and Ontology: Theories and applications (TOTh), Bruxelles (BE)*, C. Roche, R. Costa, E. Coudyzer (editors), 2014, [hal:hal-01180921].

[39] A. MERCERON, J.-M. ADAM, D. BARDOU, J. DAVID, S. LUJÁN-MORA, M. MILOSZ, Training sessions in a Master degree 'Informatics as a Second Competence', *in : Proc. IEEE international global engineering education coference (EDUCON), Berlin*, p. 190–199, 2013, [doi:10.1109/EduCon.2013.6530105], [hal:hal-00871562].

[40] L. NIXON, R. GARCÍA CASTRO, S. WRIGLEY, M. YATSKEVICH, C. TROJAHN DOS SANTOS, L. CABRAL, The state of semantic technology today: overview of the first SEALS evaluation campaigns, *in : Proc. 7th ACM international conference on semantic systems (I-semantics), Graz (AT)*, p. 134–141, 2011, [doi:10.1145/2063518.2063536], [hal:hal-00781025].

[41] F. SCHARFFE, G. ATEMEZING, R. TRONCY, F. GANDON, S. VILLATA, B. BUCHER, F. HAMDI, L. BIHANIC, G. KÉPÉKLIAN, F. COTTON, J. EUZENAT, Z. FAN, P. VANDENBUSSCHE, B. VATANT, Enabling linked data publication with the Datalift platform, *in : Proc. AAAI workshop on semantic cities, Toronto (ONT CA)*, 2012, [hal:hal-00768424].

Scientific Books (or Scientific Book chapters)

[42] F. ALKHATEEB, J. EUZENAT, Querying RDF data, *in : Graph data management: techniques and applications*, S. Sakr and E. Pardede (editors), IGI Global, Hershey (PA US), 2012, ch. 15, p. 337–356, [doi:10.4018/978-1-61350-053-8.ch015], [hal:hal-00768410].

[43] J. EUZENAT, C. LE DUC, Methodological guidelines for matching ontologies, *in : Ontology engineering in a networked world*, M. Suárez Figueroa, A. Gómez Pérez, E. Motta, and A. Gangemi (editors), Springer, Heidelberg (DE), 2012, ch. 12, p. 257–278, [doi:10.1007/978-3-642-24794-1_12], [hal:hal-00768420].

[44] J. EUZENAT, M.-C. ROUSSET, Web sémantique, *in : L'IA: frontières et applications*, P. Marquis, O. Papini, and H. Prade (editors), *3*, Cepadues, Toulouse (FR), 2014, ch. 4.

[45] J. EUZENAT, P. SHVAIKO, Ontology matching, 2nd edition, Springer-Verlag, Heidelberg (DE), 2013, [doi:10.1007/978-3-642-38721-0], [hal:hal-00918122].

[46] J. EUZENAT, First experiments in cultural alignment repair (extended version), *in : ESWC 2014 satellite events revised selected papers*, V. Presutti, E. Blomqvist, R. Troncy, H. Sack, I. Papadakis, and A. Tordai (editors), 2014, p. 115–130, [doi:10.1007/978-3-319-11955-7_10], [hal:hal-01180916].

[47] M. ROSOIU, J. DAVID, J. EUZENAT, A linked data framework for Android, *in : The Semantic Web: ESWC 2012 Satellite Events*, E. Simperl, B. Norton, D. Mladenic, E. Della Valle, I. Fundulaki, A. Passant, and R. Troncy (editors), Springer Verlag, Heidelberg (DE), 2015, ch. 15, p. 204–218, [doi:10.1007/978-3-662-46641-4_15], [hal:hal-01179146].

Books or Proceedings Editing

[48] P. CUDRÉ-MAUROUX, J. HEFLIN, E. SIRIN, T. TUDORACHE, J. EUZENAT, M. HAUSWIRTH, J. XAVIER PARREIRA, J. HENDLER, G. SCHREIBER, A. BERNSTEIN, E. BLOMQVIST (editors), Proc. 11th International semantic web conference (ISWC), Boston (MA US), *7649*, Heidelberg (DE), Springer Verlag, 2012, [doi:10.1007/978-3-642-35176-1], [hal:hal-00768414].

[49] P. CUDRÉ-MAUROUX, J. HEFLIN, E. SIRIN, T. TUDORACHE, J. EUZENAT, M. HAUSWIRTH, J. XAVIER PARREIRA, J. HENDLER, G. SCHREIBER, A. BERNSTEIN, E. BLOMQVIST (editors), Proc. 11th International semantic web conference (ISWC), Boston (MA US), *7650*, Heidelberg (DE), Springer Verlag, 2012, [doi:10.1007/978-3-642-35173-0], [hal:hal-00768415].

[50] P. SHVAIKO, I. CRUZ, J. EUZENAT, T. HEATH, M. MAO, C. QUIX (editors), Proc. 6th ISWC workshop on ontology matching (OM), Bonn (DE), 2011, [hal:hal-00781043].

[51] P. SHVAIKO, J. EUZENAT, A. KEMENTSIETSIDIS, M. MAO, N. NOY, H. STUCKENSCHMIDT (editors), Proc. 7th ISWC workshop on ontology matching (OM), Boston (MA US), 2012, [hal:hal-00768518].

[52] P. SHVAIKO, J. EUZENAT, M. MAO, E. JIMÉNEZ-RUIZ, J. LI, A. NGONGA NGOMO (editors), Proc. 9th ISWC workshop on ontology matching (OM), Riva del Garda (IT), 2014, [hal:hal-01180933].

[53] P. SHVAIKO, J. EUZENAT, K. SRINIVAS, M. MAO, E. JIMÉNEZ-RUIZ (editors), Proc. 8th ISWC workshop on ontology matching (OM), Sydney (NSW AU), 2013, [hal:hal-00918492].

Research Reports

[54] J. L. AGUIRRE, C. MEILICKE, J. EUZENAT, Iterative implementation of services for the automatic evaluation of matching tools (v2), *Deliverable number 12.5v2*, SEALS, 2012, [hal:hal-00785742].

[55] F. ALKHATEEB, J. EUZENAT, Answering SPARQL queries modulo RDF Schema with paths, *Research report number 8394*, INRIA Rhône-Alpes, Grenoble (FR), 2013, [hal:hal-00904961].

[56] M. ATENCIA, J. EUZENAT, M.-C. ROUSSET, Exploiting ontologies and alignments for trust in semantic P2P networks, *Research report number 18*, LIG, Grenoble (FR), 2011, [hal:hal-00921676].

[57] S. BIROV, S. ROBINSON, M. POVEDA VILLALÓN, M. C. SUÁREZ-FIGUEROA, R. GARCÍA CASTRO, J. EUZENAT, L. M. PRIEGO, B. FIES, A. CAVALLARO, J. PETERS-ANDERS, T. TRYFERIDIS, K. ZOI TSAGKARI, Ontologies and datasets for energy measurement and validation interoperability, *Deliverable number 3.2*, Ready4SmartCities, 2014, [hal:hal-01180923].

[58] A. CAVALLARO, F. DI GENNARO, J. EUZENAT, J. PETERS-ANDERS, A. OSELLO, B. FIES, M. SEPPONEN, Vision of energy systems for smart cities, *Deliverable number 5.2*, Ready4SmartCities, 2014, [hal:hal-01180924].

[59] M. CHEKOL, J. EUZENAT, P. GENEVÈS, N. LAYAÏDA, A benchmark for semantic web query containment, equivalence and satisfiability, *Research report number 8128*, INRIA, Grenoble (FR), 2012, [hal:hal-00749286].

[60] J. DAVID, J. EUZENAT, M. ATENCIA, Language-independent link key-based data interlinking, *Deliverable number 4.1*, Lindicle, 2015, [hal:hal-01180925].

[61] J. DAVID, J. EUZENAT, M. ROSOIU, Mobile API for linked data, *Deliverable number 6.3*, Datalift, 2012, [hal:hal-00785743].

[62] J. EUZENAT, N. ABADIE, B. BUCHER, Z. FAN, H. KHROUF, M. LUGER, F. SCHARFFE, R. TRONCY, Dataset interlinking module, *Deliverable number 4.2*, Datalift, 2011, [hal:hal-00793433].

[63] J. EUZENAT, J. DAVID, A. LOCORO, A. INANTS, T. LESNIKOVA, Context-based ontology matching and data interlinking, *Deliverable number 3.1*, Lindicle, 2015, [hal:hal-01180927].

[64] J. EUZENAT, The category of networks of ontologies, *Research report number 8652*, INRIA, Grenoble (FR), 2014, [hal:hal-01093207].

[65] R. GARCÍA CASTRO, M. POVEDA VILLALÓN, F. RADULOVIC, A. GÓMEZ PÉREZ, J. EUZENAT, L. M. PRIEGO, G. VOGT, S. ROBINSON, S. BIROV, B. FIES, J. PETERS-ANDERS, Strategy for energy measurement and interoperability, *Deliverable number 3.1*, Ready4SmartCities, 2014, [hal:hal-01180926].

[66] T. LESNIKOVA, J. DAVID, J. EUZENAT, Algorithms for cross-lingual data interlinking, *Deliverable number 4.2*, Lindicle, 2015, [hal:hal-01180928].

[67] C. MEILICKE, J. L. AGUIRRE, J. EUZENAT, O. SVÁB-ZAMAZAL, E. JIMÉNEZ-RUIZ, I. HORROCKS, C. TROJAHN DOS SANTOS, Results of the second evaluation of matching tools, *Deliverable number 12.6*, SEALS, 2012, [hal:hal-00785744].

[68] L. M. PRIEGO, J. EUZENAT, R. GARCÍA CASTRO, M. POVEDA VILLALÓN, F. RADULOVIC, M. WEISE, Strategy for Energy Management System Interoperability, *Deliverable number 2.1*, Ready4SmartCities, 2013, [hal:hal-00925628].

[69] F. SCHARFFE, J. DAVID, M. ATENCIA, Keys and pseudo-keys detection for web datasets cleansing and interlinking, *Deliverable number 4.1.2*, Datalift, 2012, [hal:hal-00785745].

[70] F. SCHARFFE, Z. FAN, A. FERRARA, H. KHROUF, A. NIKOLOV, Methods for automated dataset interlinking, *Deliverable number 4.1*, Datalift, 2011, [hal:hal-00793435].

[71] M. SEPPONEN, M. HANNUS, K. PIIRA, A. CAVALLARO, R. GARCÍA CASTRO, B. FIES, T. TRYFERIDIS, K. ZOI TSAGKARI, J. EUZENAT, F. JUDEX, D. BASCIOTTI, C. MARGUERITE, R. SCHMIDT, S. BIROV, S. ROBINSON, G. VOGT, M. WEISE, A. OSELLO, Draft of innovation and research roadmap, *Deliverable number 5.3*, Ready4SmartCities, 2014, [hal:hal-01180929].

[72] C. TROJAHN DOS SANTOS, C. MEILICKE, J. EUZENAT, Iterative implementation of services for the automatic evaluation of matching tools, *Deliverable number 12.5*, SEALS, 2011, [hal:hal-00793436].

[73] M. WEISE, M. POVEDA VILLALÓN, M. SUÁREZ-FIGUEROA, R. GARCÍA CASTRO, J. EUZENAT, L. M. PRIEGO, B. FIES, A. CAVALLARO, J. PETERS-ANDERS, K. ZOI TSAGKARI, Ontologies and datasets for energy management system interoperability, *Deliverable number 2.2*, Ready4SmartCities, 2014, [hal:hal-01180932].

[74] S. WRIGLEY, R. GARCÍA CASTRO, L. CABRAL, C. TROJAHN DOS SANTOS, C. MEILICKE, L. NIXON, M. YATSKEVICH, Design of the second evaluation campaign, *Deliverable number 3.5*, SEALS, 2011, [hal:hal-00794905].

A Full text of the previous objectives

We reproduce below, the objectives which where available in our previous evaluation report. It has been slightly reduced of circumstantial comments.
EXMO aims at building on its strengths in order to increase its contribution and impact to the field of ontology matching and alignment and the semantic web at large. For that purpose, we propose to organise our work of the next years in three lines which extends the current objectives:

Alignment foundations

On the foundation side, our ambition is to consider three specific topics:

Semantic of network of ontologies Our goal is to develop further the semantics of networks of ontologies. In particular, we aim at considering two research directions: (*i*) using the semantics for defining what a peer in a peer-to-peer system can know depending on the initial knowledge of the peers, the query language used and the alignment language considered, and (*ii*) developing the principles of revision in network of ontologies, i.e., what happens when a new correspondence or an ontology modification makes the network inconsistent.

Algebraic manipulation of alignments We currently have a view of alignment composition based either on the full semantics of network of ontologies or a "naïve" view of relations. The use of algebras of relations offers a formal intermediate position. Hence we want to fully define the possible operations on alignments, including composition, reasoning and combining alignments.

Distances between ontologies Until now our work has been limited to alignments only composed of equivalence correspondences. By taking advantage of our works on algebraic manipulation of alignments and semantic evaluation measures, we want to propose computable generalisations of our alignment space measures able to capture more precise relations. We also plan to investigate agreement and disagreement between ontologies proposed by M. d'Aquin. We want to reconsider them in order to reintroduce alignments at the core of the measures and replace the currently syntactic compatibility comparison by more semantic ones, e.g., entailment and consistency as well as weaker notions such as entailment through composition and integrity constraints. This should also provide the opportunity to compare the measures with C. Meilicke's coherence measure (based on the semantics of ontologies). We plan to experiment with these measures in matchers.

Alignment infrastructure

As usual, we want to push the theoretical work within the semantic web infrastructure in order to support applications and to improve our support to alignment management.

API improvement We aim at integrating all results of the first objective within the Alignment API (and OntoSim for the last item). This will require two major redesigns: integrating algebras of relations at the core of the API, and integrating better reasoning mechanisms within the API. Under the pressure of application using it (see Applications and Objective 3), we will certainly make EDOAL evolve.

Alignment services for ambient computing We will further improve the Alignment server and mobile library so that applications could use them easily In particular, we want to better integrate SPARQL query generation and manipulation from EDOAL to support mobile and peer-to-peer systems (see Objective 3).

Evaluation of matchers The work on matching evaluation shall continue through SEALS and OAEI. OAEI is constantly evolving: as matcher capabilities evolve, new modalities are introduced (multilingual matching, non equivalence alignments). Two years ago, we have introduced evaluation of instance matching (aiming at linked data), but there is place to more consensual improvement in this matter. We are also working on automatic generation of new test sets allowing to generate network of aligned ontologies by applying both alterations and transformations on a source ontology. This will allow us to more precisely parameterise the hardness of tests to generate.

Exploitation and data interlinking

The third objective is consolidated into one precise direction relying on previous developments. Its goal is to help using alignment in applications and especially in application to linked data.
We are planning to apply and extend our expertise in ontology matching to linked data co-reference resolution, i.e. for determining whether different URIs refer to the same entity.
So far, the development of transformations has only been carried out on simple (URI-to-URI) alignments. There are applications in which such alignments are not sufficient. We already knew it for semantic web services, but this is now clear for data interlinking. We are already experiencing this in the context of the Datalift project and colleagues in Southampton are using EDOAL for that exact same purpose. To be used in data interlinking, alignments must be expressive and a language such as EDOAL is needed. Moreover, SPARQL is the common language for processing data. Hence, we need to provide adapted EDOAL to SPARQL transformations that can be used for extracting data and eventually generating links. These same manipulation techniques may be used for extending the type of semantic peer-to-peer systems that we are currently able to consider.
This raises various interesting issues for alignment manipulation. The first natural issue is to guarantee that the EDOAL to SPARQL transformation will indeed extract the described set or to characterise what approximations have been made. For instance, if data is extracted for applying a link generation algorithm, it may not matter if more data is extracted. Another issue is related to minimality: if several correspondences may apply to the same piece of data, it should be guaranteed that their application link the same object. This type of problem has been investigated in database schema mapping. It is also related to the problem of confluence in rewriting systems (we want that the order of the transformations does not influence the result).
In order to treat this problem in full, it is necessary to deal with query or correspondence comparison. We are currently investigating the static analysis of RDF manipulation in cooperation with the WAM team.

This work will be partly developed in the context of the ANR Datalift project and more precisely with U. Montpellier (F. Scharffe).

Other issues

Better matchers. We would like to build better matchers investigating several paths among which context-based matching in which the matcher takes advantage of resources external to the ontologies, reasoning about alignments for improving them, multilingual thesauri matching, and pattern-based matchers in which matching patterns are used for guiding matching. This could be based on our existing OLA and Aroma matchers.
However, this can only be achieved if substantial resources are dedicated to these tasks. As long as such resources are not available, we cannot plan to develop these aspects.
Applications. On the application side, the semantic web is slowly making progress. We would like to contribute to the adoption and dissemination of semantic web technology more broadly. In fact, we would like to deliver it in everybody's pocket. For that purpose, we are planning to curve EXMO's trajectory towards ambient and mobile applications that we have already considered.
This would happily be achieved through an appropriate European project. We have had marks of interest from the field of ambient intelligence for energy efficient buildings, so this may be a context in which to develop this work.
This should keep us busy for the next four years. After this, we expect that the most challenging problems to be considered will revolve around the dynamics of networks of ontologies (how they preserve coherence while evolving).