



Group leader: Georges Quénot (DR CNRS).

The MRIM group studies methods for satisfying user information needs. These needs have largely evolved over time, as well as contexts in which they are formulated and the methods by which they are satisfied.

The context of Information Retrieval (IR) has broadened not only to include all types of data (texts, images, video, etc.), but also to include all types of users, and IR has been widely used through other applications (filtering, recommending systems, social networks, etc.). More recently, concerns about the fairness and the transparency of IR systems has emerged and the recent GDPR European Union regulation has introduced a “right to explanation” on algorithmic decision-making [GF17], which all IR systems or methods are subject to. This is especially challenging, as most of them now involve “deep learning black boxes”.

As in many other domains, learning-based methods, with, as a corollary, the use and building of large annotated collections, have become dominant for almost all aspects of IR. Such point should not however be considered as exclusive, as other fundamental aspects of information retrieval (like new extensions of existing models of IR) are still under concern in this research field.

MRIM has conducted research on the following themes (details are given in section 4):

- Modeling of IR systems
- Semantic indexing
- Multimedia Indexing
- Transparency and explainability
- Social Networks and Personalization
- IR in Under-Resourced Languages
- Mobile and Cultural Heritage
- Medical IR
- IR for E-tourism
- IR and Smart Cities
- Evaluation of IR systems

Recommendations after the previous evaluation. During the previous evaluation, the experts made the following recommendations:

The Expert Panel recommends that engineering work be distinguished from research work.

This is not easy to do as we have contracts with industrial partners and as we have to participate to evaluation campaigns to validate our work, both of which requires significant engineering work. However, we did our best to separate engineering work from research work by hiring more engineers on our contracts.

The MRIM team should be reinforced by an engineer and devote more time to its work on new information retrieval models.

This is indeed a strong need for MRIM. Unfortunately, we have little control over the recruitment and the affectation of research engineers and very few positions were available so we were not able to get one, even part-time. However, we improved the balance between engineering and research tasks.

Profile of activities. The activities of MRIM are divided between:

- knowledge production (25 journal and 120 conference and workshop articles published in the 2014-2018 period, organization of challenges, conferences and summer schools);
- release of code and resources (annotated data sets);
- valorization and transfer (Milliworks, SATT Linksium);
- involvement in the university life (Ahmed Lbath, CA member);
- training through research via Master and PhD supervision (14 PhDs defended in 2014-2018).

MRIM contributed to production of scientific knowledge through 25 journal articles and two books. The most significant articles were related to a multi-branch neural net architecture (ensemble learning with joint training, in Neurocomputing) [J1*], to the use of formal logic for modeling information retrieval (ACM Computing Surveys) [J2*], to an hybrid query expansion model for text and micro-blog information retrieval (Information Retrieval Journal) [J3*], to a system architecture for heterogeneous moving-object trajectory meta-model using generic sensors (IEEE Systems Journal) [J5*], to a study of a dynamic network model for smart city data-loss resilience (IEEE Access) [J4*], and to the organization of the semantic indexing task at TRECVID during six years (invited paper in ITE Transactions on Media Technology and Applications) [J18].

MRIM contributed to the formation through research of PhD (14 defended over the period) and Master (over 30 defended during the period) students. Several PhDs were co-supervised with professors from foreign universities.

MRIM co-organized, jointly with the AMA team, the 2018 (40th) edition of the European Conference on Information Retrieval (ECIR). This conference is ranked A in the Core classification and gathered 250 participants. We also organized the 2014 edition of the Autumn School on Information Retrieval and its Applications (EARIA). MRIM members are members of the organizing committee of ACM SIGIR 2020 Conference (Paris). This conference is ranked A*, and gathered 800 to 1000 participants.

MRIM released or contributed to the release of several data collections: TRECVID Semantic Indexing (INS, 2010-2015) and Ad hoc Video Search (AVS, 2016-2021), CLICIDE, GUIMUTEIC, CLEF eHealth 2014-2018 and CLEF CMC 2016-2017.

Regarding industrial transfer, MRIM sold an Image Indexing System to the GlobeVIP company.

Ethiopia is an emerging country that have decided to start a local PhD and research program. The goal of University of Addis-Ababa is to start a local PhD graduation. For this reason, MRIM has been involved in this new scientific program. More precisely, we are in charge of the course "Information Retrieval Agent" for the Master in Computer Science of this University. We have also participated to the first local PhD Graduation in Computer Science in the track of Natural Language Processing [PhD6].

3.1 Team members

| <i>Permanent Members</i> | | | |
|--------------------------|--------------------|--------------|-------------|
| <i>Name</i> | <i>Institution</i> | <i>Grade</i> | <i>Date</i> |
| Georges Quénot | CNRS | DR | 1/1/1998 - |
| Catherine Berrut | UGA | Prof. | 1/1/1985 - |
| Marie-Christine Fauvet | UGA | Prof. | 1/1/2003 - |
| Ahmed Lbath | UGA | Prof. | 1/3/2012 - |
| Jean-Pierre Chevallet | UGA | MCF | 1/9/1993 - |
| Philippe Mulhem | CNRS | CR | 1/12/2003 - |
| Nathalie Denos | UGA | MCF | 1/9/1998 - |
| Lorraine Gœuriot | UGA | MCF | 1/9/2014 - |

| <i>Non Permanent Members</i> <i>(PhDs who defended were removed from this list)</i> | | |
|--|---------------|-----------------------|
| <i>Name</i> | <i>Status</i> | <i>Date</i> |
| Nadia Derbas | Engineer | 1/11/2013 - 30/4/2014 |
| Camille Persson | Engineer | 1/10/2015 - 30/9/2016 |
| Mateusz Budnik | Engineer | 1/1/2016 - 31/12/2016 |
| Bahjat Safadi | Engineer | 1/10/2014 - 30/4/2017 |
| Johann Poignant | Engineer | 1/3/2016 - 31/8/2017 |
| Nawal Ould-Amer | PhD | 1/10/2014 - |
| Adil Sheikh | PhD | 1/9/2015 - |
| Anuvabh Dutt | PhD | 1/10/2016 - |
| Seydou Doumbia | PhD | 1/10/2016 - |
| Édouard Bahri | PhD | 1/1/2017 - |
| Jibril Frej | PhD | 1/10/2017 - |
| Khalid Halba | PhD | 27/4/2018 - |

During the period, the team recruited Lorraine Gœuriot as Maître de Conference. Marie-Christine Fauvet was recognized as a disabled worker and is on “congé de longue maladie” since 2016. Nathalie Denos was “chargé de mission au Ministère de l’Éducation Nationale” until 2016 and since then full time “mise à disposition” there. The list of the 14 PhD students that defended during the period is given in section 1.3.3 of the annex.

3.2 Team organization and scientific animation

We have team meetings once a week, each Thursday afternoon. These mostly consist in talks by invited speakers or by MRIM members about their recent work followed by discussions about how this work can be improved by or joint with the work of other team members. We also have talks or discussions that report about a conference issue, a particularly important paper or some methodology aspect or some useful tools. In order to share our knowledge, PhD students are co-supervised within the group.

As many research activities in the team involve very intensive computations and also require very large storage space, we have developed an internal facility mutualized with the GETALP team. It includes computing nodes (currently approximately 200 CPU cores and 20 GPU boards) and storage servers (currently 280 Tbytes). Though we also have access for some projects to larger shared structures like Grid’5000 (national) or CIMENT (local), this facility gives us more flexibility, reactivity and storage. This is very important for the participation to evaluation campaigns that are themselves essential for our research. The facility is regularly upgraded whenever possible using contractual funding.

3.3 Financial resources

MRIM's financial resources come from a number of contracts, the list and the amounts of which are detailed in the annex of this document, in sections 1.1.7 for academic research grants and 1.2.1 for industrial collaborations. In total, the team received about 1420 k€ over the period, of which about 1200 k€ from academic research grants and about 220 k€ from industrial collaborations. This is globally less than the previous period but this one was exceptional thanks to the Quaero project (1400 k€ alone).

3.4 Ethics

Parity; scientific integrity; health and safety; sustainable development and taking into account environmental impacts; intellectual property and economic intelligence

Information Retrieval is a very sensible subject in terms of ethics and from many respects and this in an important concern for the MRIM team, in particular regarding privacy, fairness and accountability. Privacy is a major concern for people involved in data collection and annotation; this includes query logs or personal data, even for publicly released information. Regarding fairness, we investigate whether search engines actually give "equal chances" to content providers and how to design unbiased systems in this respect. Regarding accountability, we work on systems able to explain their decisions, as required by the recently released GDPR [GF17].

We have a 1:1 parity between permanent members. However, the situation is much less good in terms of PhD and Engineer recruitment. We believe that we have a fair selection process in this respect and that the imbalance comes from the applications. In the future, we will take further steps to encourage women applications.

We indeed do our best to ensure a high integrity standard in our research practices, for instance by enforcing rigorous and transparent experimental protocols in the evaluation of our methods, and whenever possible through participation to open international challenges.

We also have a commitment to the open source approach and we release our significant software implementations and annotated data collections unless when working with industrial partners that require us not to.

Modeling of IR Systems. Modelling an Information Retrieval Systems (IRS) consists in building a formal description of an IRS module (Analysis, Indexing, Matching, or Ranking). We worked on modelling matching and ranking IRS activities by using logical models. The use of logic for this modelling, rises from the following hypothesis: a document is an answer to a query, if there exists a logical deduction chain that starts from the document and ends to the query. This deduction chain can be a fuzzy one, i.e. a probability to deduce the query from the documents and used for ranking. We have proposed a new IR logic matching model using logical Boolean lattice mixed with a probabilistic function over this lattice. This modelling enables matching functions to be decomposed into a direct matching function (deduction from the document to the query), and a reverse matching function, that evaluate the strength of the deduction from the query to the document. This was the PhD work of Karam Abdulahhad [PhD12] and also part of it is included in a survey published at C-CAM surveys [J2*].

The term mismatch problem, which happens when query terms fail to appear in relevant documents to the query, is a long standing problem in information retrieval. However, it is not clear how often term mismatch happens in retrieval, how important it is for retrieval, or how it affects retrieval performance. An essential component for achieving term mismatch probability reduction is the knowledge resource that defines terms and their relationships. A variety of knowledge resources have been exploited, in our proposals, in order to produce effective modifications on documents or queries. More particularly, we proposed a query expansion approach based on neural language models. Neural language models are proposed to learn term vector representations, called distributed neural embeddings. And we obtained impressive results comparing with state of the art approaches in term similarity tasks. This was the PhD work of Mohannad Almasri [PhD5] and was also published in ECIR [C13*].

Semantic indexing. Semantic Information Retrieval [JF2] deals with the usage of knowledge sets to annotate documents and solve queries. The idea behind this research track lies in the following hypothesis: using only computed word distribution statistics, may not be sufficient to find the semantic links that join query to relevant documents in some IR situations. These situations are, for example, when high search precision is required, like in specialized domains (ex: medical), or when document textual descriptors are too short for efficient term statistic evaluations. In these specific situations, the usage of such explicit knowledge resources can make the difference. Semantic indexing is then the usage of a specific knowledge resources (Ex: UMLS¹, Wordnet², etc.) to automatically annotate textual document with entries to theses resources. The matching can then exploit the links in theses semantic resources to enhance the matching quality. For example, a painting only described by the text "the last supper", can match a query about "painting and Da Vinci" only if the system has this link encoded in some resources.

Term based, or semantic based, resources must be large enough and consistent to be correctly exploited by IR systems. Unfortunately, these resource are rare, because very costly to build manually. Also, because most of existing resources are oriented toward human usage, they are neither complete nor consistent. For example, this is the case for the UMLS Meta-thesaurus: despite its very large term and multi-term coverage of the medical domain, in more that 18 languages, and the strong human effort to maintain this resource of about 1 million of "concepts", and the twice a year constant updates, this medical domain specific resource is not consistent. For example, the simple 'is-a' hierarchical relationship have some looping path. The main problem in such a resource is a lack of explicit mathematical logical semantics. Hence, the important first step to exploit these very large and valuable resources, like UMLS, is first to 'clean' them and force their structure to respect some logical language. They can then be partially transformed into ontology expressed on explicit logical language. This work has been done in the Demeke Ayele PhD thesis [PhD6], in which he has proposed a Knowledge Acquisition Framework from Unstructured Bio-medical Knowledge Sources. The work is continuing with the PhD of Jibril

¹<https://www.nlm.nih.gov/research/umls/>

²<https://wordnet.princeton.edu>

Frej, that started in September 2017. This research track is oriented toward the usage of Machine Learning for transforming discrete knowledge set into a continuous multidimensional space (embedding). A proposal to extend the IR Language Model using learned term embedding [C33] has been proposed in CORIA 2018.

Multimedia Indexing. Several works were conducted in the context of multimedia indexing.

In the context of the QCompere project, we developed methods for the multimodal recognition of persons in video documents [JF1*, J25, J16, C82]. We worked on the design of multimodal descriptors for violent scenes detection in movies [C116, C23*, PhD11]. We worked on the optimization of visual descriptors by simultaneously significantly reducing their size and increasing their performance in classification tasks [J24, PhD11], on methods for improving the efficiency of multi-label classification for large scale multimedia indexing [C87], and on methods for the fusion of a large number of multimodal descriptors [Ch1*].

We worked on the detection of several concepts simultaneously in images or in video shots. We found that just combining detection scores using an appropriate function performs at least as well as directly training detectors for frequent combinations of two or three target concepts [C118, C24*, J15, PhD10]. We investigated the use of conceptual and temporal relations for concept detection in video shots. We exploited both the implicit (co-occurrences) and explicit (generic-specific, exclusion) relations between concepts. For the temporal aspect, we exploit the fact that, in videos when a concept appears (or not) in one shot, it is more likely (or less likely) to appear in the few previous or next shots. We obtained significant performance improvement for both types of relations [C86, J23, PhD10]. We also worked on active learning for minimizing the annotation effort per person recognition in video documents [C107, C111, PhD3*].

In the previously described work, we used deep learning as this became the most efficient approach for most tasks but we also did specific studies on its use in multimedia indexing. We compared engineered (classical) features versus learned ones. Using many classical ones can be on par with learned ones and that fusing both can perform even better [C85, J20, PhD3*]. However, this approach is costly and unlikely to remain advantageous in the future, given the rapid progresses of deep learning methods. We also investigated the use of concept hierarchies for improving their recognition [C4*, C41]. In order to investigate whether the obtained gain was really due to the use of relations between concepts or to an ensemble learning, effect, we indirectly discovered that doing ensemble learning with a joint training could significantly improve the detection accuracy. [C35, C30, J1*].

Three PhDs (Nadia Derbas [PhD11], Abdelkader Hamadi [PhD10], and Mateusz Budnik [PhD3*]) related to multimedia indexing were defended over the period and one (Anuvabh Dutt) is ongoing.

Transparency and explainability. Transparency and explainability are two very important elements related to information access and artificial learning. When we focus on the transparency of information retrieval systems during their design, axiomatic constraints may be used. Such aspects were studied for personalization purposes in [C43] and for image search in [C88]. Others directions studied in the team are related to the transparency of classical web search engines “by testing”, i.e., without any knowledge of their internal features. We propose a global framework (the first one to our knowledge) to define and run experiments on “non-cooperative” search engines. It defines appropriate metrics for evaluating dissimilarities between search results, and protocols for evaluating the dependence of returned results on various parameters through constraints on test queries. A first preliminary experiment [C26] was awarded the best paper in the french conference CORIA 2019, and larger experiments are currently run.

A second research direction is dedicated to the explainability of by convolutional neural networks classification results. Our current focus are explanations through the use of data from the training set, knowing that this set is not always available. This work is achieved through several M2R work in the context of LIG emergence projects.

Social Networks and Personalization. The research work dedicated to social networks and personalization aims at defining methods and models integrating information coming from the user

(or his interaction), in a way to enhance the quality of the retrieval. Our work is based on parsimonious information retrieval models: such probabilistic language models (PLM) seek to build compact and precise term distributions by eliminating stop words and non-essential terms. PLM were successfully applied for relevance feedback to capture relevant terms from feedback document to expand a query. Our approach consists in expanding such models in a way to support one important kind of social networks, namely the tagging systems that allow users to assign tags to documents (like web pages), in a way to generate a tag-based personalized parsimonious information retrieval model, PTPLM. Because personalization is used for expanding queries using terms of documents, one part of our extension integrates in the optimization process of PLM the links between user's tags and documents terms through the use of word embeddings. The results presented in [C5*, C11*] show that our proposal outperforms state of the art. This work has been mainly supported by a regional PhD grant (Nawal Ould-Amer, on-going, RESPIR project).

Information retrieval models adapted to social media, and in particular microblogs, has been explored in several directions. The first one studies the applicability of integration of classical IR techniques with user-user relationships during the CLEF Social Book Search evaluation campaigns [C83]. A second direction is explored with the team AMA (Massih-Reza Amini) within an LIG Emergence project. Work carried out within this project has been published in [C32, C54, C3*].

Mining micro-blogs requires to handle properly noisy and duplicated data. Handling noise and detecting online e-activism has been explored as part of a collaboration with researchers from PACTE laboratory (laboratoire de sciences sociales). Preliminary results have been published in [C29].

IR in Under-Resourced Languages. Cross-lingual information retrieval (CLIR) consists in querying in a given language a system in which documents are in another language. Dealing with several languages requires an adaptation of information retrieval models: either by translating queries or documents to get back to a monolingual context, or through multilingual matching models. Our objective here is to focus on low-resourced languages, *i.e.*, languages for which little or no linguistic knowledge and resources are available. As an illustration, a selection of use cases have been published in [C46]. In this context, solutions to CLIR are constrained by the limited resources available: machine translation tools, dictionaries or aligned corpora cannot be used. Recent work has shown that multilingual word embeddings could be built almost without any training data. We currently explore how these multilingual embedding spaces can be integrated in an information retrieval model. This work is part of Seydou Dombia PhD work, but also a collaboration with GETALP (Laurent Besacier, Didier Schwab) supported by LIG through 2 Emergence projects.

Mobile and Cultural Heritage Information access in mobility situation raises new challenges: Users are no more only concentrated on the search task, as the typical Web search engine usage. In mobility, the main task remains in the real world, like visiting a museum, being guide to seek for a good restaurant, etc. In this type of information access, context takes an important place: time, location, user activities, situation, etc. This need the studies of the integration of the user context and the current task into the IR Model. This research direction has been studied in the PhD of Mr. Kian Lam Tan in 2014 [PhD13]. In his thesis, we propose a time and location depend IR Model, in which user interest and time are modeled using the physical analogy of capacitor. A mobile game on Malaysian Cultural Heritage has also been developed during this PhD. One research direction studied by the MRIM team is related to usage of state of the art neural networks to support mobile interaction during museum visits. This work has been conducted in the context of the GUIMUITEIC FUI project. The goal of this project was to propose fully autonomous devices, due the museum constraints. The work of Maxime Portaz during his PhD [PhD1*] defined solutions for two separated problems : 1) the recognition of masterpieces, and 2) the recognition of the gestures of the visitors in way to interact with the system. For both problems, the solution proposed is to define fully Convolutional neural networks (FCNN), specifically adapted to be used on mobile devices like tablets. For the recognition of

masterpieces, a FCNN that is able to use regions is proposed. One innovation of this proposal is to be able to learn the network without the need of regions input [C38]. The detection of user's gestures is achieved through a global architecture composed of two main blocks (well adapted for mobile devices). This work has been published in the MTAP journal [J9]. This work was also used to participate to an international evaluation campaign in 2016 [C59].

Medical IR. Information retrieval in specialized domains requires to go beyond classical bag-of-words approaches. In the medical domain for instance, lay-users searching for health information may face difficulties expressing their information need or understanding documents retrieved. This issue, also known as the lexical gap, can be tackled by expanding the user queries with related terms, extracted from knowledge sources or embeddings built from medical data [C70, C71]. Considering among the relevance criteria the understandability of documents can also be considered, as a way to lower the gap between the reader's knowledge and the document. Learning to rank approaches allow to integrate such aspects in the IR process [C12*].

IR for E-tourism. This part is dedicated to the integration of web-based Information Systems. It aims at addressing issues that arise when designers want to assemble pieces of Information Systems accessible by the means of web services.

The studies involved have been conducted as a part of a broader project whose main goal is to design and implement a software system which provides context-aware personalized services for mobile users. This project was mainly funded by European Union's Erasmus Mundus project "Sustainable e-tourism, Erasmus Mundus Action 2 programme". More precisely, the issues addressed are related to mobile computing, and more specifically to system design, software architecture, expansion of information systems, distributed and heterogeneous resource access and integration (indexing and querying in highly distributed and heterogeneous environments), quality of service requirements (in presence of limited bandwidth and service discontinuity), and information synchronization when switching between connected and disconnected mode. Challenges we attempt to respond are those related to: i) new models for user's privacy preservation in clustering and context aware recommendation (Mou Lei's PhD Thesis [], funded by the European Commission), ii) semantic composition of recommended services into a composite service, and execution of the resulting composite service (Pathathai Na Lumpoon's PhD Thesis [PhD9], funded by the European Commission), iii) discovery of services and recommendation (Isaac Caicedo's PhD Thesis [PhD8], supported by COLCIENCIAS-COLFUTURO Colombia scholarship, and also founded by the University of Córdoba in Colombia) and trip planning recommendation under constraints (Uyanga SUKHBAATAR Shukhbaatar, PhD Thesis [PhD7], funded by the European Commission, joint degree with the national university of science and technology of Mongolia).

IR and Smart Cities. The research conducted in this part is mainly funded by international grants and international collaboration. It is related to the domain of cyber-physical social system. We address problems that raised scientific challenges in smart cities and we design and develop new methods and algorithms validated through experimental use cases in smart building, smart transportation and smart health-care. More specifically, we focus on multi-source real-time knowledge extraction (using NLP and other techniques for event discovery), visualization and learning-based prediction models.

Nowadays we are witnessing a new revolution in mobile computing and communication. Social and physical networks are being brought together to deal with the variety of social challenges in a number of areas. The main advantage of envisioning such systems is the ability to process data and information from a number of sources that contribute to creation and development of knowledge to operate and control the numerous systems and sub-systems that constitute these complex interdependent systems. The rise of internet, internet of things and social networks manifest in the form of social media provide new opportunities to create more effective and efficient monitoring and delivery of services that are evidence based and in a timely manner.

Our research results provide a framework that combines solutions and methodology from machine learning and event processing to propose efficient environments address that allow people,

organizations and companies to make smart decision and efficiently respond to urgent urban life situations (Thesis of Olivera Kotevska [PhD2*]). We made experimentation on real world examples (ex. crime data analytics in the Montgomery County, MD, USA and the city of New York). Moreover, we designed and developed a framework (Hadath) incorporating new algorithms and tools to address the challenges of real-time knowledge extraction and event discovery from multi-source structured and unstructured data (Tweet, micro-blogs, crowd sourcing, ...) and visualization through enriched maps (Thesis of Faizan Ur Rehman [PhD4]).

Evaluation of IR systems. Evaluation of information retrieval systems, when following the Cranfield paradigm [CLE60], requires test collection. Building such collection requires to define precise and realistic use cases, establish search scenarios including queries, contextual information and assessment of documents given the scenario. One way to collectively build such collections and share them with the community is to run evaluation campaigns. MRIM team has been very active in running such evaluation challenges, mainly in the context of the CLEF eHealth [C72, GCJ+14], CLEF CMC [C34] and TRECVID challenges [J18, O1*].

The CLEF eHealth challenge has been running since 2013 and aims at helping patients, their next-of-kins and medical professionals understanding health data. The challenge gathers evaluation tasks in information extraction, information management and information retrieval. Challenges and participation are documented in [C49, C48, C73, C76, C75, C74, C94, C21*, C96, C95, C109, C108, C105, C119]. These evaluation tasks have gathered communities and contribute to various publications in the domain [J10]. Focusing on the information retrieval task in particular, running it for several years has allowed to deeply analyze data sets and participants runs [J6, GKL14].

Similarly, the team has been involved in the organization of the CLEF CMC challenge, which purpose lies in improving micro-blog retrieval. As a use case, tasks were centered on micro-blogs surrounding cultural events. Data sets and tasks are described in [C39, C47, C55, C56, C34].

Along with these activities, the process of creation of a test collection and the validity of relevance assessment itself has been explored and questioned [C14*]. A comparison of various assessment methods and assessor types (hired professional assessors vs crowdsourcers) was conducted within the EIRAP project (on-going).

MRIM has co-organized with NIST the Semantic INdexing (SIN) from 2010 to 2015 included [J18, O1*] and Ad hoc Video Search (AVS) from 2016 to 2021 (planned) included [O1*] tasks at TRECVID³.

The SIN task at TRECVID is the counterpart for videos (876,527 video shots and 346 target concepts) of the popular ImageNet Large Scale Visual Recognition Challenge (ILSVRC) for still images with a number of significant differences beyond considering video shots instead of still images: contents are “from the wild” and target categories are not exclusive, highly imbalanced, and generally not the main focus of the video shots. Tens of international teams have participated to this task, which has been run for 6 consecutive years [J18, O1*], of which the 2014 [C101] and 2015 [C78] issues were within the period.

After the 6-year period, considering the progresses made, the SIN task was considered as less interesting and the focus moved to the AVS task [O1*], for which the SIN one was an enabler. The task consists in solving complex queries in natural language like “find shots of a person talking behind a podium wearing a suit outdoors during daytime” in a collection of 335,944 video shots (~600 hours). A first three-year cycle has been completed during the period (2016-2018) [C51, C36, C27] and a new one is planned starting this year⁴.

MRIM also participated to these evaluations [C104, C79] and organized joint participations at the level of the French community in the context of the IRIM group of the ISIS GDR [C99, C77, C53, C37, C28].

³<https://trecvid.nist.gov/>

⁴<https://www-nlpir.nist.gov/projects/tv2019/avs.html>

5) RÉFÉRENCES BIBLIOGRAPHIQUES

- [CLE60] C.W. CLEVERDON. The ASLIB cranfield research project on the comparative efficiency of indexing systems. *Aslib Proceedings*, 12(12):421–431, 1960.
- [GCJ⁺14] Lorraine Goeuriot, Wendy Chapman, Gareth JF Jones, Liadh Kelly, Johannes Leveling, and Sanna Salanterä. Building realistic potential patients queries for medical information retrieval evaluation. In *the LREC workshop on Building and Evaluating Resources for Health and Biomedical Text Processing.*, reykjavik, Iceland, 2014.
- [GF17] Bryce Goodman and Seth R. Flaxman. European union regulations on algorithmic decision-making and a "right to explanation". *AI Magazine*, 38(3):50–57, 2017.
- [GKL14] Lorraine Goeuriot, Liadh Kelly, and Johannes Leveling. An analysis of query difficulty for information retrieval in the medical domain. In *the 37th international ACM SIGIR conference*, Gold Coast, Australia, July 2014. ACM Press.

Note: in order to avoid duplicates, references to the publications from the MRIM team (e.g., [J1*]) are in the Annex on Scientific and Technological Results.

EVALUATION CAMPAIGN 2019-2020
VAGUE A

Name of the team : Multimedia Information Modeling and Retrieval
 Acronym : MRIM

Head of the team for the current contract : Georges Quénot
 Head of the team for the next contract : Georges Quénot

1.1 Production of knowledge and activities contributing to the influence and scientific attractiveness of the unit

1.1.1 Journal Articles

Scientific articles in English

Top 20%

- [J1*] Anuvabh Dutt, Georges Quénot, and Denis Pellerin. Coupled Ensembles of Neural Networks. *Neurocomputing*, April 2019.
- [J2*] Karam Abdulahhad, Catherine Berrut, Jean-Pierre Chevallet, and Gabriella Pasi. Modeling Information Retrieval by Formal Logic: A Survey. *ACM Computing Surveys*, 2018.
- [J3*] Catherine Berrut, Meriem Amina Zingla, Yaya Slimani, Philippe Mulhem, and Latiri Chiraz. Hybrid Query Expansion Model for Text and Microblog Information Retrieval. *Information Retrieval Journal*, 2018.
- [J4*] Olivera Kotevska, A. Gilad Kusne, Daniel V Samarov, Ahmed Lbath, and Abdella Battou. Dynamic Network Model for Smart City Data-Loss Resilience Case Study: City-to-City Network for Crime Analytics. *IEEE Access*, 5:20524–20535, 2017.
- [J5*] Azedine Boulmakoul, Lamia Karim, Adil Elbouziri, and Ahmed Lbath. A System Architecture for Heterogeneous Moving-Object Trajectory Metamodel Using Generic Sensors: Tracking Airport Security Case Study. *IEEE Systems Journal*, 9(1):283–291, March 2015.
- [J6] Lorraine Goeuriot, Gareth Jones, Liadh Kelly, Johannes Leveling, Mihai Lupu, Joao Palotti, and Guido Zuccon. An analysis of evaluation campaigns in ad-hoc medical information retrieval: CLEF eHealth 2013 and 2014. *Information Retrieval Journal*, 21(6):507 – 540, December 2018.
- [J7] Akhlaq Ahmad, Md. Abdur Rahman, Mohamed Ridza Wahiddin, Faizan

- Ur Rehman, Abdelmajid Khelil, and Ahmed Lbath. Context-aware services based on spatio-temporal zoning and crowdsourcing. *Behaviour and Information Technology*, 37(7):736–760, May 2018.
- [J8] Mohamed Nahri, Azedine Boulmakoul, Lamia Karim, and Ahmed Lbath. IoV distributed architecture for real-time traffic data analytics. *Procedia Computer Science*, 130:480–487, 2018.
- [J9] Maxime Portaz, Matthias Kohl, Jean-Pierre Chevallet, Georges Quénot, and Philippe Mulhem. Object Instance Identification with Fully Convolutional Networks. *Multimedia Tools and Applications*, pages <https://doi.org/10.1007/s11042-018-5798-7>, 2018.
- [J10] Hanna Suominen, Liadh Kelly, and Lorraine Goeuriot. Scholarly Influence of the Conference and Labs of the Evaluation Forum eHealth Initiative: Review and Bibliometric Study of the 2012 to 2017 Outcomes. *JMIR Research Protocols*, 7(7), 2018.
- [J11] Badredine Boulmakoul, Zineb Besri, Lamia Karim, Azedine Boulmakoul, and Ahmed Lbath. Galois's algebraic structure and bipartite graph spatio-structural analytics for urban public transportation system assessment. *Procedia Computer Science*, 109:172–179, 2017.
- [J12] Lamia Karim, Azedine Boulmakoul, Aziz Mabrouk, and Ahmed Lbath. Deploying Real Time Big Data Analytics in Cloud Ecosystem for Hazmat Stochastic Risk Trajectories. *Procedia Computer Science*, 109:180–187, 2017.
- [J13] Aziz Mabrouk, Azedine Boulmakoul, Lamia Karim, and Ahmed Lbath. Safest and shortest itineraries for transporting hazardous materials using split points of Voronoï spatial diagrams based on spatial modeling of vulnerable zones. *Procedia Computer Science*, 109:156–163, 2017.
- [J14] Meriem Mandar, Lamia Karim, Azedine Boulmakoul, and Ahmed Lbath. Triangular intuitionistic fuzzy number theory for driver-pedestrians interactions and risk exposure modeling. *Procedia Computer Science*, 109:148–155, 2017.
- [J15] Abdelkader Hamadi, Philippe Mulhem, and Georges Quénot. A comparative study for multiple visual concepts detection in images and videos. *Multimedia Tools and Applications*, 75(15):8973–8997, August 2016.
- [J16] Johann Poignant, Guillaume Fortier, Laurent Besacier, and Georges Quénot. Naming multi-modal clusters to identify persons in TV broadcast. *Multimedia Tools and Applications*, 75(15):8999–9023, August 2016. published online.
- [J17] Karam Abdulahhad, Jean-Pierre Chevallet, and Catherine Berrut. Logics, Lattices and Probability: The Missing Links to Information Retrieval. *The Computer Journal*, July 2016.
- [J18] George Awad, Cees G M Snoek, Alan F Smeaton, and Georges Quénot. TRECVID Semantic Indexing of Video: A 6-Year Retrospective. *ITE Transactions on Media Technology and Applications*, 4(2016):22, July 2016. Invited Paper.
- [J19] Azedine Boulmakoul, Lamia Karim, and Ahmed Lbath. T-Warehousing for hazardous materials transportation. *Ingénierie des Systèmes d'Information*, 2016.

- [J20] Mateusz Budnik, Efrain-Leonardo Gutierrez-Gomez, Bahjat Safadi, Denis Pellerin, and Georges Quénot. Learned features versus engineered features for multimedia indexing. *Multimedia Tools and Applications*, 2016. published online.
- [J21] Olivera Kotevska, A Lbath, and S Bouzefrane. Toward a real-time framework in cloudlet-based architecture. *Tsinghua Science and Technology*, 21:80 – 88, 2016.
- [J22] Isaac Caicedo-Castro, Marie-Christine Fauvet, Ahmed Labath, and Helga Duarte-Amaya. Toward the highest effectiveness in text description-based service retrieval. *Revue des Sciences et Technologies de l'Information - Série Document Numérique*, 18(2-3):155 – 177, December 2015.
- [J23] Abdelkader Hamadi, Philippe Mulhem, and Georges Quénot. Extended conceptual feedback for semantic multimedia indexing. *Multimedia Tools and Applications*, 74(4):1225–1248, February 2015.
- [J24] Bahjat Safadi, Nadia Derbas, and Georges Quénot. Descriptor Optimization for Multimedia Indexing and Retrieval. *Multimedia Tools and Applications*, 74(4):1267–1290, February 2015.
- [J25] Johann Poignant, Laurent Besacier, and Georges Quénot. Unsupervised Speaker Identification in TV Broadcast Based on Written Names. *IEEE Transactions on Audio, Speech and Language Processing*, 23(1):57–68, January 2015.

Other articles (professional journals, etc.)

Top 20%

- [JF1*] Johann Poignant, Laurent Besacier, and Georges Quénot. Nommage non supervisé des personnes dans les émissions de télévision. Utilisation des noms écrits, des noms prononcés ou des deux ? *Revue des Sciences et Technologies de l'Information - Série Document Numérique*, 17(1):pp. 37–60, 2014.
- [JF2] Haifa Zargayouna, Catherine Roussey, and Jean-Pierre Chevallet. Recherche d'information sémantique : état des lieux. *Traitement Automatique des Langues*, 56(3), 2015.

1.1.2 Books

Management and coordination of scientific books Scientific book edition Management and coordination of scientific books / Scientific book edition in English or another foreign language

Top 20%

- [B1*] Bogdan Ionescu, Jenny Benois-Pineau, Tomas Piatrik, and Georges Quénot. *Fusion in Computer Vision: Understanding Complex Visual Content*. Springer international publishing, 2014.

- [B2] Massih-Reza Amini, Renaud Blanch, Marianne Clausel, Jean-Baptiste Durand, Éric Gaussier, Jérôme Malick, Christophe Picard, Vivien Quéma, and Georges Quénot. *Data Science*. Eyrolles, August 2018.

Book chapters in English or another foreign language

Top 20%

- [Ch1*] S. Tiberius Strat, A. Benoit, Hervé Bredin, Georges Quénot, and P. Lambert. Hierarchical Late Fusion for Concept Detection in Videos. In *Fusion in Computer Vision: Understanding Complex Visual Content*, pages 53–78. Springer international publishing, 2014.

Edited theses

Top 20%

- [PhD1*] Maxime Portaz. *Information Access in mobile environment for museum visits -Deep Neraul Networks for Instance and Gesture Recognition*. Theses, Université Grenoble Alpes, October 2018.
- [PhD2*] Olivera Kotevska. *Learning based event model for knowledge extraction and prediction system in the context of Smart City*. Theses, Université Grenoble Alpes, January 2018.
- [PhD3*] Mateusz Budnik. *Active and deep learning for multimedia*. Theses, l'Université de Grenoble-Alpes, February 2017.
- [PhD4] Faizan Ur Rehman. *Towards a Framework for Multiscale Social Event Extraction and Visualization*. Theses, Université Grenoble Alpes, December 2018.
- [PhD5] Mohannad Almasri. *Reducing Term Mismatch Probability by Exploiting Semantic Term Relations*. Theses, Université Grenoble Alpes, June 2017.
- [PhD6] Demeke Asres Ayele. *Knowledge Acquisition Framework from Unstructured Biomedical Knowledge Sources*. Theses, Université d'Addis Abeba, October 2016.
- [PhD7] Uyanga Sukhbaatar. *A model simulation for trip planning recommendation system in Tourism*. Theses, Université Grenoble Alpes, October 2016.
- [PhD8] Isaac Caicedo-Castro. *S3niffer : A text description-based service search system*. Theses, Université Grenoble Alpes, May 2015.
- [PhD9] Pathathai Na Lumpoon. *Toward a framework for automated service composition and execution: E-tourism Applications*. Theses, Université Grenoble Alpes, May 2015.
- [PhD10] Abdelkader Hamadi. *Using context for semantic indexing of images and videos*. Theses, Université Joseph Fourier Grenoble -1-, October 2014.

- [PhD11] Nadia Derbas. *Contributions for the concepts and events detection in videos documents*. Theses, Université de Grenoble, September 2014.
- [PhD12] Karam Abdulahhad. *Information Retrieval (IR) Modeling by Logic and Lattice. Application to Conceptual IR*. Theses, Université de Grenoble, May 2014.
- [PhD13] Kian Lam Tan. *Information Access in Cultural Heritage*. Theses, Université de Grenoble, April 2014.

1.1.3 Production in conferences / congresses and research seminars

Articles published in conference proceedings / congress

Top 20%

- [C1*] Catherine Berrut, Sourour Belhaj, and Latiri Chiraz. Combining Bilingual Lexicons Extracted from Comparable Corpora: the Complementary Approach between Word Embedding and Text Mining . In *DEXA*, Regensburg, Germany, September 2018.
- [C2*] Catherine Berrut, Sourour Belhadj, and Latiri Chiraz. Advanced Text Mining Methods For Bilingual Lexicon Extraction from specialized comparable corpora. In *CICLING*, Hanoi, Vietnam, March 2018.
- [C3*] Philippe Mulhem, Lorraine Goeuriot, Nayanika Dogra, and Nawal Ould Amer. TimeLine Illustration Based on Microblogs: When Diversification Meets Metadata Re-ranking. In *Experimental IR Meets Multilinguality, Multimodality, and Interaction - 8th International Conference of the CLEF Association, CLEF 2017*, Dublin, Ireland, September 2017.
- [C4*] Anuvabh Dutt, Denis Pellerin, and Georges Quénot. Improving Image Classification Using Coarse and Fine Labels. In *Proceedings of the 2017 ACM International Conference on Multimedia Retrieval, ICMR '17*, pages 438–442, Bucarest, Romania, June 2017. ACM.
- [C5*] Nawal Ould Amer, Philippe Mulhem, and Mathias Géry. Personalized Parsimonious Language Models for User Modeling in Social Bookmaking Systems. In *European Conference on Information Retrieval*, Aberdeen, United Kingdom, April 2017.
- [C6*] Jill-Jénn Vie, Fabrice Popineau, Françoise Tort, Benjamin Marteau, and Nathalie Denos. A Heuristic Method for Large-Scale Cognitive-Diagnostic Computerized Adaptive Testing. In *Learning @ scale (L@S'17)*, Proceedings of the Fourth (2017) ACM Conference on Learning @ Scale (L@S '17), Cambridge, MA, United States, April 2017. © ACM 2017. This is the author's version of the work. It is posted here for your personal use. Not for redistribution. The definitive Version of Record was published in ACM Digital Library, <http://dx.doi.org/10.1145/3051457.3054015>.
- [C7*] Lamia Karim, Azedine Boulmakoul, and Ahmed Lbath. Real time analytics of urban congestion trajectories on hadoop-mongoDB cloud ecosystem. In *the Second International Conference*, pages 1–11, Cambridge, United Kingdom, March 2017. ACM Press.

- [C8*] Faizan Ur Rehman, Imad Afyouni, Ahmed Lbath, Ahmad Sohaib, Saleh Basalamah, and Mohamed Mokbel. Building Multi-Resolution Event-Enriched Maps From Social Data. In *EDBT 2017*, Venice, Italy, 2017.
- [C9*] Shuja Jamil, Sohaib Khan, Anas Basalamah, and Ahmed Lbath. Classifying smart-phone screen ON/OFF state based on wifi probe patterns. In *the 2016 ACM International Joint Conference*, pages 301–304, Heidelberg, Germany, September 2016. ACM Press.
- [C10*] Philippe Mulhem, Nawal Ould Amer, and Mathias Géry. Axiomatic Term-Based Personalized Query Expansion Using Bookmarking System. In Sven Hartmann and Hui Ma, editors, *Database and Expert Systems Applications - 27th International Conference, DEXA 2016*, volume 9828 of *Lecture Notes in Computer Science*, pages 235 – 243, Porto, Portugal, September 2016. Springer.
- [C11*] Nawal Ould Amer, Philippe Mulhem, and Mathias Géry. Toward Word Embedding for Personalized Information Retrieval. In *Neu-IR: The SIGIR 2016 Workshop on Neural Information Retrieval*, volume abs/1606.06991, Pisa, Italy, July 2016.
- [C12*] Joao Palotti, Lorraine Goeuriot, Guido Zuccon, and Allan Hanbury. Ranking Health Web Pages with Relevance and Understandability. In *the 39th International ACM SIGIR conference*, Pisa, Italy, July 2016. ACM Press.
- [C13*] Mohannad Almasri, Catherine Berrut, and Jean-Pierre Chevallet. A Comparison of Deep Learning Based Query Expansion with Pseudo-Relevance Feedback and Mutual Information. In *Conférence ECIR*, volume 42, pages 369 – 715, Padoue, Italy, March 2016.
- [C14*] Joao Palotti, Guido Zuccon, Johannes Bernhardt, Allan Hanbury, and Lorraine Goeuriot. Assessors agreement: A case study across assessor type, payment levels, query variations and relevance dimensions. In *International conference of the cross-language evaluation forum for European languages*, Evora, Portugal, 2016.
- [C15*] Azedine Boulmakoul, Lamia Karim, Adil Elbouziri, Faizan Ur Rehman, Ahmed Lbath, Imad Afyouni, Abdullah Murad, Md. Abdur Rahman, Bilal Sadiq, Akhlaq Ahmad, and Saleh Basalamah. Semantic multimedia-enhanced spatio-temporal queries in a crowdsourced environment. In *the 23rd SIGSPATIAL International Conference*, volume 9, pages 1–4, Seattle, Washington, November 2015. ACM Press.
- [C16*] Faizan Ur Rehman, Ahmed Lbath, Abdullah Murad, Md. Abdur Rahman, Bilal Sadiq, Akhlaq Ahmad, Ahmad Ahmad, and Saleh Basalamah. A Semantic Geo-Tagged Multimedia-Based Routing in a Crowdsourced Big Data Environment. In *the 23rd ACM international conference*, pages 759–760, Brisbane, Australia, October 2015. ACM Press.
- [C17*] Bilal Sadiq, Md. Abdur Rahman, Abdullah Murad, Muhammad Shahid, Faizan Ur Rehman, Ahmed Lbath, Akhlaq Ahmad, and Ahmad Ahmad. Crowdsourced Multimedia Enhanced Spatio-temporal Constraint Based on-Demand Social Network for Group Mobility. In *the 23rd ACM international conference*, pages 761–762, Brisbane, Australia, October 2015. ACM Press.
- [C18*] Shuja Jamil, Anas Basalamah, Ahmed Lbath, and Moustafa Youssef. Hybrid participatory sensing for analyzing group dynamics in the largest annual religious gathering. In *the 2015 ACM International Joint Conference*, pages 547–558, Osaka, Japan, September 2015. ACM Press.

- [C19*] Meriem Amina Zingla, Latiri Chiraz, Yahya Slimani, and Catherine Berrut. Statistical and Semantic Approaches for Tweet Contextualization. In *19th International Conference on Knowledge Based and Intelligent Information and Engineering Systems*, Singapour, Singapore, September 2015. National University of Singapore Institute of System Science (NUS-ISS).
- [C20*] Olivera Kotevska, Ahmed Lbath, and Samia Bouzefrane. Toward a Framework in Cloudlet-Based Architecture for a Real-Time Predictional Model. In *2015 IEEE 12th Intl Conf on Ubiquitous Intelligence and Computing and 2015 IEEE 12th Intl Conf on Autonomic and Trusted Computing and 2015 IEEE 15th Intl Conf on Scalable Computing and Communications and Its Associated Workshops (UIC-ATC-ScalCom)*, pages 1016–1021, Beijing, China, August 2015. IEEE.
- [C21*] Lorraine Goeriot, Liadh Kelly, Hanna Suominen, Leif Hanlen, Aurelie Neveol, Cyril Grouin, Joao Palotti, and Guido Zuccon. Overview of the CLEF eHealth Evaluation Lab 2015. In *CLEF 2015*, Toulouse, France, 2015.
- [C22*] P. Na Lumpoon, Marie-Christine Fauvet, and Ahmed Lbath. Toward a framework for automated service composition and execution. In *8th International Conference on Software, Knowledge, Information Management and Applications (SKIMA 2014)*, Dhala, Bangladesh, December 2014.
- [C23*] Nadia Derbas and Georges Quénot. Joint Audio-Visual Words for Violent Scenes Detection in Movies. In *International Conference on Multimedia Retrieval*, Glasgow, United Kingdom, 2014.
- [C24*] Abdelkader Hamadi, Philippe Mulhem, and Georges Quénot. Infrequent concept pairs detection in multimedia documents. In *ACM International Conference on Multimedia Retrieval*, pages 2081–2084, Glasgow, Scotland, United Kingdom, 2014. ACM.
- [C25] Lucas Albarede, Francis Jambon, and Philippe Mulhem. Exploration de l'apport de l'analyse des perceptions oculaires : étude préliminaire pour le bouclage de pertinence. In *CONFérence en Recherche d'Informations et Applications - CORIA 2019, 16th French Information Retrieval Conference*, Lyon, France, March 2019.
- [C26] Philippe Mulhem, Lydie Du Bousquet, and Sara Lakah. Quelques pas vers l'Honnêteté et l'Explicabilité de moteurs de recherche sur le Web. In *CONFérence en Recherche d'Information et Applications*, Villeurbanne, France, March 2019.
- [C27] George Awad, Asad A Butt, Keith Curtis, Yooyoung Lee, Jonathan Fiscus, Afzad Godil, David Joy, Andrew Delgado, Alan F. Smeaton, Yvette Graham, Wessel Kraaij, Georges Quénot, Joao Magalhaes, David Semedo, and Saverio Blasi. TRECVID 2018: Benchmarking Video Activity Detection, Video Captioning and Matching, Video Storytelling Linking and Video Search. In *Proceedings of TRECVID 2018*, Gaithersburg, MD, United States, November 2018.
- [C28] Boris Mansencal, Jenny Benois-Pineau, Hervé Bredin, and Georges Quénot. IRIM at TRECVID 2018: Instance Search. In *Proceedings of TRECVID*, Gaithersburg, MD, United States, November 2018.
- [C29] Jean-Marc Francony, Philippe Mulhem, Florence Andraocola, Lorraine Goeriot, and Georges Quénot. Détection de détournements coordonnés de hashtags et messages-images pour l'analyse des fils d'actualités sur les réseaux sociaux. In

Modèles et analyse des réseaux : approches mathématiques et informatiques, Avignon, France, October 2018.

- [C30] Georges Quénot, Anuvabh Dutt, and Denis Pellerin. Coupled Ensembles of Neural Networks. In *International Conference on Content-Based Multimedia Indexing (CBMI 2018)*, La Rochelle, France, September 2018.
- [C31] Jambon Francis, Philippe Mulhem, and Lucas Albarede. Toward Eye Gaze Enhanced Information Retrieval Relevance Feedback. page 23, June 2018. Poster.
- [C32] Nayanika Dogra, Philippe Mulhem, Lorraine Goeuriot, and Massih-Reza Amini. Corpus d'entraînement sur les plongements de mots pour la recherche de microblogs culturels. In *CONFérence en Recherche d'Informations et Applications - CO-RIA 2018*, Rennes, France, May 2018.
- [C33] Jibril FREJ, Philippe Mulhem, Didier Schwab, and Jean-Pierre Chevallet. Combining Subword information and Language model for Information Retrieval. In *15e CONFérence en Recherche d'Information et Applications*, Rennes, France, May 2018.
- [C34] Lorraine Goeuriot, Josiane Mothe, Philippe Mulhem, and Sanjuan Eric. Building Evaluation Datasets for Cultural Microblog Retrieval. In *Proceedings of the Eleventh International Conference on Language Resources and Evaluation, LREC 2018*, Miyazaki, Japan, May 2018.
- [C35] Anuvabh Dutt, Denis Pellerin, and Georges Quénot. Coupled Ensembles of Neural Networks. In *Sixth International Conference on Learning Representations - Workshop Track (ICLR 2018)*, Vancouver, Canada, April 2018.
- [C36] George Awad, Asad A Butt, Jonathan Fiscus, David Joy, Andrew Delgado, Willie McClinton, Martial Michel, Alan F. Smeaton, Yvette Graham, Wessel Kraaij, Georges Quénot, Maria Eskevich, Roeland Ordelman, Gareth Jones, and Benoit Huet. TRECVID 2017: Evaluating Ad-hoc and Instance Video Search, Events Detection, Video Captioning, and Hyperlinking. In *TREC Video Retrieval Evaluation (TRECVID)*, Gaithersburg, MD, United States, November 2017.
- [C37] Boris Mansencal, Jenny Benois-Pineau, Hervé Bredin, Alexandre Benoit, Nicolas Voiron, Patrick Lambert, Hervé Le Borgne, Adrian Popescu, Alexandru Ginsca, and Georges Quénot. IRIM at TRECVID 2017: Instance Search. In *TRECVID workshop 2017*, Gaithersburg, Maryland, United States, November 2017.
- [C38] Maxime Portaz, Matthias Kohl, Georges Quénot, and Jean-Pierre Chevallet. Fully Convolutional Network and Region Proposal for Instance Identification with Ego-centric Vision. In *IEEE International Conference on Computer Vision Workshop (ICCVW)*, Venice, Italy, October 2017. IEEE.
- [C39] Liana Ermakova, Lorraine Goeuriot, Josiane Mothe, Philippe Mulhem, Jian-Yun Nie, and Eric Sanjuan. CLEF 2017 Microblog Cultural Contextualization Lab Overview. In *International Conference of the Cross-Language Evaluation Forum for European Languages (CLEF 2017)*, volume 10456, pages pp. 304–314, Dublin, Ireland, September 2017.
- [C40] Adil Sheikh, Emad Felemban, Ahmad Alhindi, Atif Naseer, Mukhtar Ghaleb, and Ahmed Lbath. OpToGen - A genetic algorithm based framework for optimal topology generation for linear networks. In *2017 13th IEEE International Conference on Intelligent Computer Communication and Processing (ICCP)*, pages 255–262, Cluj-Napoca, Romania, September 2017. IEEE.

- [C41] Anuvabh Dutt, Denis Pellerin, and Georges Quénot. Improving Hierarchical Image Classification with Merged CNN Architectures. In *Proceedings of the 15th International Workshop on Content-Based Multimedia Indexing, CBMI '17*, pages 31:1–31:7, Florence, Italy, June 2017. ACM.
- [C42] Soumaya Guesmi, Chiraz Trabelsi, Catherine Berrut, and Chiraz Latiri. Détection de communautés multi-relationnelles dans les réseaux sociaux hétérogènes. In *Conférence CORIA*, Marseille, France, March 2017.
- [C43] Philippe Mulhem, Nawal Ould Amer, and Mathias Géry. Variations axiomatiques pour la recherche d'information personnalisée. In *Conférence en Recherche d'Informations et Applications*, Marseille, France, March 2017.
- [C44] Nawal Ould Amer, Philippe Mulhem, and Mathias Géry. Modèles de Documents Parcimonieux basés sur les annotations et les " word embeddings " -Application à la personnalisation. In *CORIA 2017 - Conférence en Recherche d'Informations et Applications*, Marseille, France, March 2017.
- [C45] Maxime Portaz, Johann Poignant, Budnik Mateusz, Philippe Mulhem, Jean-Pierre Chevallet, and Lorraine Goeuriot. Construction et évaluation d'un corpus pour la recherche d'instances d'images muséales. In *Conférence en Recherche d'Informations et Applications - CORIA 2017, 14th French Information Retrieval Conference*, pages 17–34, Marseille, France, March 2017.
- [C46] Seydou S. Doumbia, Lorraine Goeuriot, and Marie-Christine Fauvet. Quels problèmes pour la recherche d'information médicale au Mali ? In *RJCRI - CORIA*, Marseille, France, 2017.
- [C47] Lorraine Goeuriot, Philippe Mulhem, and Eric Sanjuan. CLEF 2017 MC2 search and time line tasks overview. In *CLEF 2017 Online Working Notes*, Dublin, Ireland, 2017.
- [C48] Joao Palotti, Guido Zuccon, Jimmy Jimmy, Mihai Lupu, Lorraine Goeuriot, Liadh Kelly, and Allan Hanbury. Clef 2017 task overview: The IR task at the ehealth evaluation lab. In *Working Notes of Conference and Labs of the Evaluation (CLEF) Forum*, Dublin, Ireland, 2017.
- [C49] Joao Palotti, Guido Zuccon, Jimmy Jimmy, Pavel Pecina, Mihai Lupu, Lorraine Goeuriot, Liadh Kelly, and Allan Hanbury. CLEF 2017 Task Overview: The IR Task at the eHealth Evaluation Lab - Evaluating Retrieval Methods for Consumer Health Search. In *CLEF (Working Notes), volume 1866 of CEUR Workshop Proceedings*, Dublin, Ireland, 2017.
- [C50] Olivera Kotevska, Ahmed Lbath, and Judith Gelernter. Event model to facilitate data sharing among services. In *IEEE 3rd World Forum on Internet of Things (WF-IoT), Internet of Things (WF-IoT), 2016 IEEE 3rd World Forum on*, pages 577 – 584, Reston, VA, United States, December 2016.
- [C51] George Awad, Jonathan Fiscus, David Joy, Martial Michel, Alan F. Smeaton, Wessel Kraaij, Maria Eskevich, Robin Aly, Roeland Ordelman, Marc Ritter, Georges Quénot, Gareth Jones, Benoit Huet, and Martha Larson. TRECVID 2016: Evaluating Video Search, Video Event Detection, Localization, and Hyperlinking. In *TREC Video Retrieval Evaluation (TRECVID)*, Gaithersburg, MD, United States, November 2016.

- [C52] Seyyed Hadi Hashemi, Nawal Ould Amer, and Jaap Kamps. Neural Endorsement Based Contextual Suggestion. In *Proceedings of The Twenty-Fifth Text REtrieval Conference, TREC 2016*, Gaithersburg, United States, November 2016.
- [C53] Boris Mansencal, Jenny Benois-Pineau, Hervé Bredin, Alexandre Benoit, Nicolas Voiron, Patrick Lambert, and Georges Quénot. IRIM at TRECVID 2016: Instance Search. In *TRECVID workshop 2016*, TRECVID workshop proceedings, Gaithersburg, Maryland, United States, November 2016. National Institute of Standards and Technology (NIST).
- [C54] Nayanika Dogra, Philippe Mulhem, Nawal Ould Amer, and Lorraine Goeriot. LIG at CLEF 2016 Cultural Microblog Contextualization: TimeLine Illustration based on Microblogs. In *CLEF 2016*, volume 1609 of *LEF 2016 Working Notes*, pages 1201–1206, Evora, Portugal, September 2016.
- [C55] Liana Ermakova, Lorraine Goeriot, Josiane Mothe, Philippe Mulhem, Jian-Yun Nie, and Eric Sanjuan. Cultural micro-blog Contextualization 2016 Workshop Overview: data and pilot tasks. In *CLEF 2016*, volume 1609 of *CLEF 2016 Working Notes*, pages 1197–1200, Evora, Portugal, September 2016.
- [C56] Lorraine Goeriot, Josiane Mothe, Philippe Mulhem, Fionn Murtagh, and Sanjuan Eric. Overview of the CLEF 2016 Cultural Micro-blog Contextualization Workshop. In Springer International Publishing, editor, *International Conference of the Cross-Language Evaluation Forum for European Languages*, volume Lecture Notes in Computer Science, volume 9822 of *CLEF 2016: Experimental IR Meets Multilinguality, Multimodality, and Interaction*, pages 371–378, Evora, Portugal, September 2016.
- [C57] Olivera Kotevska, Sarala Padi, and Ahmed Lbath. Automatic Categorization of Social Sensor Data. In *EUSPN/ICTH*, volume 98, pages 596 – 603, London, United Kingdom, September 2016.
- [C58] Nawal Ould-Amer, Philippe Mulhem, Mathias Géry, and Karam Abdulahhad. Word Embedding for Social Book Suggestion. In Krisztian Balog, Linda Cappellato, Nicola Ferro, and Craig Macdonald, editors, *Clef 2016 Conference*, volume 1609, Evora, Portugal, September 2016.
- [C59] Maxime Portaz, Mateusz Budnik, Philippe Mulhem, and Johann Poignant. MRIM-LIG at ImageCLEF 2016 Scalable Concept Image Annotation Task. In *Clef 2016 Conference*, volume 1609 of *CLEF 2016 Working Notes*, pages 363–370, Evora, Portugal, September 2016.
- [C60] Nawal Ould Amer. Enhancing Personalized Document Ranking using Social Information. In *ACM UMAP '16. Proceedings of the 2016 Conference on User Modeling Adaptation and Personalization*, Halifax, Canada, July 2016.
- [C61] Mateusz Budnik, Laurent Besacier, Ali Khodabakhsh, and Cenk Demiroglu. Deep complementary features for speaker identification in TV broadcast data. In *Odyssey Workshop 2016*, Bilbao, Spain, June 2016.
- [C62] Bahjat Safadi, Philippe Mulhem, Georges Quénot, and Jean-Pierre Chevallet. Lifelog Semantic Annotation using deep visual features and metadata-derived descriptors. In *14th International Workshop on Content-Based Multimedia Indexing (CBMI)*, Content-Based Multimedia Indexing (CBMI), 2016 14th International Workshop on, pages 1 – 6, Bucarest, Romania, June 2016.

- [C63] Bahjat Safadi, Philippe Mulhem, Georges Quénot, and Jean-Pierre Chevallet. LIG-MRIM at NTCIR-12 Lifelog Semantic Access Task. In *12th NTCIR Conference on Evaluation of Information Access Technologies*, Proceedings of the 12th NTCIR Conference on Evaluation of Information Access Technologies, Tokyo, Japan, June 2016.
- [C64] Lamia Karim, Azedine Boulmakoul, and Ahmed Lbath. Near real-time big data analytics for NFC-enabled logistics trajectories. In *2016 3rd International Conference on Logistics Operations Management (GOL)*, pages 1–7, Fez, Morocco, May 2016. IEEE.
- [C65] Johann Poignant, Mateusz Budnik, Hervé Bredin, Claude Barras, Mickael Stefas, Pierrick Bruneau, Gilles Adda, Laurent Besacier, Hazim Ekenel, Gil Francopoulo, Javier Hernando, Joseph Mariani, Ramon Morros, Georges Quénot, Sophie Rosset, and Thomas Tamisier. The CAMOMILE Collaborative Annotation Platform for Multi-modal, Multi-lingual and Multi-media Documents. In *LREC 2016 Conference*, Portoroz, Slovenia, May 2016.
- [C66] Mateusz Budnik, Laurent Besacier, Ali Khodabakhsh, and Cenk Demiroglu. OCR-aided person annotation and label propagation for speaker modeling in TV shows. In *IEEE ICASSP 2016*, Shanghai, China, March 2016.
- [C67] Maxime Portaz, Philippe Mulhem, and Jean-Pierre Chevallet. Étude préliminaire à la recherche de photographies muséales en mobilité. In *CORIA 2016 Conférence en Recherche d'Information et Applications 2016*, pages 335–344, Toulouse, France, March 2016.
- [C68] Adil Sheikh, Ahmed Lbath, Ehsan Warriach, Shahbaz Awan, Sheikh Sheikh, and Emad Felemban. A software platform for smart data-driven intelligent transport applications. In *2016 IEEE International Conference on Pervasive Computing and Communication Workshops (PerCom Workshops)*, pages 1–6, Sydney, Australia, March 2016. IEEE.
- [C69] Meriem Amina Zingla, Chiraz Latiri, Yahya Slimani, Catherine Berrut, and Philippe Mulhem. Tweet Contextualization Based on Wikipedia and Dbpedia. In *CONFERENCE en Recherche d'Information et Applications (CORIA), Semaine du Document NUMérique et de la Recherche d'Information. Conférence CORIA.*, pages 545–560, Toulouse, France, March 2016.
- [C70] Alexander Beloborodov and Lorraine Goeuriot. Improving Health Consumer Search with Contextual Information. In *2nd SIGIR workshop on Medical Information Retrieval (MedIR).*, pisa, Italy, 2016.
- [C71] Julie Budaher, Mohannad Almasri, and Lorraine Goeuriot. Comparison of Several Word embedding Sources for Medical Information Retrieval. In *CLEF 2016 (online working notes).*, Evora, Portugal, 2016.
- [C72] Lorraine Goeuriot, Liadh Kelly, Guido Zuccon, and Joao Palotti. Building Evaluation Datasets for Consumer-Oriented Information Retrieval. In *the Tenth International Conference on Language Resources and Evaluation (LREC 2016).*, portoroz, Slovenia, 2016.
- [C73] Liadh Kelly, Lorraine Goeuriot, Hanna Suominen, Aurelie Neveol, Joao Palotti, and Guido Zuccon. Overview of the CLEF eHealth Evaluation Lab 2016. In *Fuhr N. et al. (eds) Experimental IR Meets Multilinguality, Multimodality, and Interaction. CLEF 2016*, Evora, Portugal, 2016.

- [C74] Aurelie Neveol, Lorraine Goeuriot, Liadh Kelly, K Cohen, Cyril Grouin, Thierry Hamon, Thomas Lavergne, Grégoire Rey, Aude Robert, Xavier Tannier, and Pierre Zweigenbaum. Clinical information extraction at the clef ehealth evaluation lab 2016. In *CLEF 2016 (online working notes)*., Evora, Portugal, 2016.
- [C75] Hanna Suominen, Liyuan ZHOU, Lorraine Goeuriot, and Liadh Kelly. Task 1 of the CLEF eHealth Evaluation Lab 2016: Handover Information Extraction. In *CLEF (Working Notes)*, Evora, Portugal, 2016.
- [C76] Guido Zuccon, Joao Palotti, Lorraine Goeuriot, Liadh Kelly, Mihai Lupu, Pavel Pecina, Henning Mueller, Julie Budaher, and Anthony Deacon. The IR task at the clef ehealth evaluation lab 2016: User-centred health information retrieval. In *CLEF 2016 (online working notes)*., evora, Portugal, 2016.
- [C77] Hervé Le Borgne, Philippe Gosselin, David Picard, Miriam Redi, Bernard Mérialdo, Boris Mansencal, Jenny Benois-Pineau, Stéphane Ayache, Abdelkader Hamadi, Bahjat Safadi, Nadia Derbas, Mateusz Budnik, Georges Quénot, Boyang Gao, Chao Zhu, Yuxing Tang, Emmanuel Dellandrea, Charles-Edmond Bichot, Liming Chen, Alexandre Benoit, Patrick Lambert, and Tiberius Strat. IRIM at TRECVID 2015: Semantic Indexing. In *Proceedings of TRECVID*, Gaithersburg, MD, United States, November 2015.
- [C78] Paul Over, George Awad, Jon Fiscus, Martial Michel, David Joy, Alan F Smeaton, Wessel Kraaij, Georges Quénot, Roeland Ordelman, and Robin Aly. TRECVID 2015 – An Overview of the Goals, Tasks, Data, Evaluation Mechanisms, and Metrics. In *TREC Video Retrieval Evaluation (TRECVID)*, Gaithersburg, MD, United States, November 2015.
- [C79] Bahjat Safadi, Nadia Derbas, Abdelkader Hamadi, Mateusz Budnik, Philippe Mulhem, and Georges Quénot. LIG at TRECVID 2015: Semantic Indexing. In *Proceedings of TRECVID*, Gaithersburg, MD, United States, November 2015.
- [C80] Meriem Amina Zingla, Chiraz Latiri, Yahya Slimani, and Catherine Berrut. Wikipedia-based Semantic Approach for Tweet Contextualization. In *KDDA - Knowledge Discovery and Data Analysis*, Alger, Algeria, November 2015.
- [C81] Sourour Belhaj, Latiri Chiraz, Yahya Slimani, and Catherine Berrut. Query Expansion based on Association Rules measured by the Entropic Implication Intensity. In *IDEAL (International Conference on Intelligent Data Engineering and Automated Learning)*, Wroclaw, Poland, October 2015.
- [C82] Mateusz Budnik, Bahjat Safadi, Laurent Besacier, Georges Quénot, Ali Khodabakhsh, and Cenk Demiroglu. LIG at MediaEval 2015 Multimodal Person Discovery in Broadcast TV Task. In *MediaEval 2015 Workshop*, Wurzen, Germany, September 2015.
- [C83] Nawal Ould-Amer, Philippe Mulhem, and Mathias Géry. LIG at CLEF 2015 SBS Lab. In *Conference and Labs of the Evaluation Forum*, Toulouse, France, September 2015.
- [C84] Emre Demir, Zehra Cataltepe, Umit Ekmekci, Mateusz Budnik, and Laurent Besacier. Unsupervised Active Learning For Video Annotation. In *ICML Active Learning Workshop 2015*, Lille, France, July 2015.

- [C85] Mateusz Budnik, Efrain Leonardo Gutierrez Gomez, Bahjat Safadi, and Georges Quénot. Learned features versus engineered features for semantic video indexing. In *13th International Workshop on Content-Based Multimedia Indexing (CBMI)*, Prague, Czech Republic, June 2015.
- [C86] Abdelkader Hamadi, Philippe Mulhem, and Georges Quénot. Temporal rescore vs. temporal descriptors for semantic indexing of videos. In *13th International Workshop on Content-Based Multimedia Indexing (CBMI)*, pages 1–4, Prague, Czech Republic, June 2015.
- [C87] Bahjat Safadi and Georges Quénot. A factorized model for multiple SVM and multi-label classification for large scale multimedia indexing. In *13th International Workshop on Content-Based Multimedia Indexing (CBMI)*, pages 1–6, Prague, Czech Republic, June 2015.
- [C88] Philippe Mulhem and Jean-Pierre Chevalet. Contrainte de correspondance Document-Document pour la RI. Application à la Divergence de Kullback-Leibler. In *CORIA 2015 Conférence sur la Recherche d'Information et ses Applications*, pages 157–172, Paris, France, March 2015. ARIA.
- [C89] Nawal Ould Amer, Philippe Mulhem, and Mathias Géry. Recherche de conversations dans les réseaux sociaux : modélisation et expérimentations sur Twitter. In *Conférence en Recherche d'Informations et Applications - 12th French Information Retrieval Conference*, Paris, France, March 2015.
- [C90] Adil Amjad, Ahmed Lbath, Ehsan Ullah, and Emad Felemban. A Predictive Data Reliability Method for Wireless Sensor Network Applications. In *ICA3PP 2015 International Conference on Algorithms and Architectures for Parallel Processing*, Zhangjiajie, China, 2015.
- [C91] Isaac Caicedo, Marie-Christine Fauvet, Helga Duarte, and Ahmed Lbath. Applying a family of IR models to text description-based service retrieval. In *Conférence en Recherche d'Information et Applications*, Conférence en Recherche d'Information et Applications, Paris, France, 2015.
- [C92] Isaac Caicedo-Castro, Marie-Christine Fauvet, Ahmed Lbath, and Helga Duarte-Amaya. Applying a family of IR models to text description-based service retrieval. In *CORIA, PARIS*, France, 2015.
- [C93] Sanjay Kamath, Lorraine Goeriot, and Marie-Christine Fauvet. Processing natural language queries to disambiguate named entities and extract users' goals: application to e-Tourism. In *RJCRI - CORIA*, Toulouse, France, 2015.
- [C94] Aurelie Neveol, Cyril Grouin, Xavier Tannier, Thierry Hamon, Liadh Kelly, Lorraine Goeriot, and Pierre Zweigenbaum. CLEF eHealth Evaluation Lab 2015 Task 1b: Clinical Named Entity Recognition. In *CLEF 2015 Working notes.*, Toulouse, France, 2015.
- [C95] Joao Palotti, Guido Zuccon, Lorraine Goeriot, Liadh Kelly, Allan Hanbury, Gareth Jones, Mihai Lupu, and Pavel Pecina. CLEF eHealth Evaluation Lab 2015, Task 2: Retrieving Information About Medical Symptoms. In *CLEF 2015 Working notes.*, toulouse, France, 2015.
- [C96] Hanna Suominen, Leif Hanlen, Lorraine Goeriot, Liadh Kelly, and Gareth J. F. Jones. Task 1a of the CLEF eHealth Evaluation Lab 2015: Clinical Speech Recognition. In *CLEF 2015 Working notes.*, Toulouse, France, 2015.

- [C97] Faizan Ur Rehman, Ahmed Lbath, Bilal Sadiq, Md. Abdur Rahman, Abdullah Murad, and Imad AFYOUNI. A constraint-aware optimized path recommender in a crowdsourced environment. In *2015 IEEE/ACS 12th International Conference of Computer Systems and Applications (AICCSA)*, Marrakech, Morocco, 2015.
- [C98] Akhlaq Ahmad, Md. Abdur Rahman, Faizan Ur Rehman, Ahmed Lbath, Imad Afyouni, Abdelmajid Khelil, Syed Osama Hussain, Bilal Sadiq, and Mohamed Ridza Wahiddin. A framework for crowd-sourced data collection and context-aware services in Hajj and Umrah. In *2014 IEEE/ACS 11th International Conference on Computer Systems and Applications (AICCSA)*, pages 405–412, Doha, Qatar, November 2014. IEEE.
- [C99] Nicolas Ballas, Benjamin Labbé, Hervé Le Borgne, Philippe Gosselin, David Picard, Miriam Redi, Bernard Merialdo, Boris Mansencal, Jenny Benois-Pineau, Stéphane Ayache, Abdelkader Hamadi, Bahjat Safadi, Nadia Derbas, Mateusz Budnik, Georges Quénot, Boyang Gao, Chao Zhu, Yuxing Tang, Emmanuel Dellandrea, Charles-Edmond Bichot, Liming Chen, Alexandre Benoit, Patrick Lambert, and Tiberius Strat. IRIM at TRECVID 2014: Semantic Indexing and Instance Search. In *Proceedings of TRECVID*, Orlando, United States, November 2014.
- [C100] Lei Mou and Ahmed Lbath. A grid-based location privacy-preserving method for LBS users. In *the 1st ACM SIGSPATIAL International Workshop*, pages 1–4, Dallas/Fort Worth, United States, November 2014. ACM Press.
- [C101] Paul Over, Jon Fiscus, Greg Sanders, David Joy, Martial Michel, George Awad, Alan F. Smeaton, Wessel Kraaij, and Georges Quénot. TRECVID 2014 – An Overview of the Goals, Tasks, Data, Evaluation Mechanisms and Metrics. In *Proceedings of TRECVID*, page 52, Orlando, Florida, United States, November 2014.
- [C102] Ahmad Qamar, Imad Afyouni, Md. Abdur Rahman, Faizan Ur Rehman, Delwar Hussain, Saleh Basalamah, and Ahmed Lbath. A GIS-based serious game interface for therapy monitoring. In *the 22nd ACM SIGSPATIAL International Conference*, pages 589–592, Dallas, United States, November 2014. ACM Press.
- [C103] Faizan Ur Rehman, Ahmed Lbath, Md. Abdur Rahman, Saleh Basalamah, Imad Afyouni, Akhlaq Ahmad, and Syed Osama Hussain. Toward dynamic path recommender system based on social network data. In *the 7th ACM SIGSPATIAL International Workshop*, pages 64–69, Dallas/Fort Worth, United States, November 2014. ACM Press.
- [C104] Bahjat Safadi, Nadia Derbas, Abdelkader Hamadi, Mateusz Budnik, Philippe Mulhem, and Georges Quénot. LIG at TRECVID 2014: Semantic Indexing. In *Proceedings of TRECVID*, Orlando, United States, November 2014.
- [C105] Lorraine Goeriot, Liadh Kelly, Wei Li, Joao Palotti, Pavel Pecina, Guido Zuccon, Allan Hanbury, Gareth J. F. Jones, and Henning Müller. ShARe/CLEF eHealth Evaluation Lab 2014, Task 3: User-centred health information retrieval. In *Proceedings of CLEF 2014*, Sheffield, United Kingdom, September 2014.
- [C106] Shuja Jamil, Ahmed Lbath, and Anas Basalamah. Crowdsensing traces using bluetooth low energy (BLE) proximity tags. In *the 2014 ACM International Joint Conference*, pages 71–74, Seattle, United States, September 2014. ACM Press.
- [C107] Budnik Mateusz, Johann Poignant, Laurent Besacier, and Georges Quénot. Active Selection with Label Propagation for Minimizing Human Effort in Speaker Annotation of TV Shows. In *Workshop on Speech, Language and Audio in Multimedia (SLAM 2014)*, page 5 p., Penang, Malaysia, September 2014.

- [C108] Danielle L. Mowery, Sumithra Velupillai, Brett R. South, Lee Christensen, David Martinez, Liadh Kelly, Lorraine Goeriot, Noemie Elhadad, Sameer Pradhan, Guergana Savova, and Wendy Chapman. Task 2: ShARe/CLEF eHealth Evaluation Lab 2014. In *Proceedings of CLEF 2014*, Sheffield, United Kingdom, September 2014.
- [C109] Hanna Suominen, Tobias Schreck, Gondy Leroy, Harry Hochheiser, Lorraine Goeriot, Liadh Kelly, Danielle L. Mowery, Jaume Nualart, Gabriela Ferraro, and Daniel Keim. Task 1 of the CLEF eHealth Evaluation Lab 2014 Visual-Interactive Search and Exploration of eHealth Data. In *Proceedings of CLEF 2014*, Sheffield, United Kingdom, September 2014.
- [C110] Lei Mou, Ahmed Lbath, and Marie-Christine Fauvet. Towards a mobile application framework with preservation of location privacy: Application to m-Tourism. In *Colloque International de Géomatique*, Orléans, France, July 2014.
- [C111] Mateusz Budnik, Johann Poignant, Laurent Besacier, and Georges Quénot. Automatic propagation of manual annotations for multimodal person identification in TV shows. In *12th International Workshop on Content-Based Multimedia Indexing (CBMI)*, Klagenfurt, Austria, June 2014.
- [C112] Mohannad Almasri, Jean-Pierre Chevallet, and Catherine Berrut. Exploiting Wikipedia Structure for Short Query Expansion in Cultural Heritage. In *CORIA*, Nancy, France, 2014.
- [C113] Mohannad Almasri, Jean-Pierre Chevallet, Catherine Berrut, Philippe Mulhem, and kiam lam tan. Integrating semantic term relations into information retrieval systems based on language models. In *Proc. of Asia Information Retrieval Society Conference (AIRS)*, Kuching, Malaysia, 2014.
- [C114] Mohannad Almasri, Jean-Pierre Chevallet, Catherine Berrut, Philippe Mulhem, and kiam lam tan. Integrating terms hierarchy into Dirichlet Language model. In *6e Atelier Recherche d'Information SEMantique (RISE)*, Nancy, France, 2014.
- [C115] Azedine Boulmakoul and Ahmed Lbath. Near Real-Time Space-Time-Path Warehousing for Hazardous Materials Transportation Trajectories. In *Conference on Advanced Decisional Systems*, Hammamet, Tunisia, 2014.
- [C116] Nadia Derbas and Georges Quénot. Mots audio-visuels joints pour la détection de scènes violentes dans les vidéos. In *CORIA*, pages 63–77, Nancy, France, 2014.
- [C117] Abdelkader Hamadi, Philippe Mulhem, and Georges Quénot. Annotation de vidéos par paires rares de concepts. In *CORIA*, Nancy, France, 2014.
- [C118] Abdelkader Hamadi, Georges Quénot, and Philippe Mulhem. Annotation of still images by multiple visual concepts. In *12th International Workshop on Content-Based Multimedia Indexing*, Klagenfurt, Austria, 2014.
- [C119] Liadh Kelly, Lorraine Goeriot, Hanna Suominen, Tobias Schreck, Gondy Leroy, Danielle L. Mowery, Sumithra Velupillai, Wendy W. Chapman, Guido Zuccon, and Joao Palotti. Overview of the ShARe/CLEF eHealth Evaluation Lab 2014. In *CLEF 2014*, sheffield, United Kingdom, 2014.
- [C120] Philippe Mulhem and Jean-Pierre Chevallet. Correspondances compatibles avec les fichiers inverses pour la recherche d'information. In *CORIA*, pages –, Nancy, France, 2014.

Other products presented in symposia / congress and research seminars

Top 20%

[O1*] George Awad, Duy-Dinh Le, Chong-Wah Ngo, Vinh-Tiep Nguyen, Georges Quénot, Cees G M Snoek, and Shin'ichi Satoh. Tutorial – Video Indexing, Search, Detection, and Description with Focus on TRECVID. In *International Conference on Multimedia Retrieval*, Bucharest, Romania, June 2017. ACM Press.

[O2] M.-C Fauvet, I.-B Caicedo-Castro, P Na-Lumpoon, and Ahmed Lbath. A Framework for Discovering and Automatically Composing Services. In *ICSOC - demo session*, GOA, India, 2015.

[O3] Marie-Christine Fauvet, Sanjay Kamath, Isaac-Bernardo Caicedo-Castro, Pathathai Na-Lumpoon, Ahmed Lbath, and Lorraine Gœuriot. Offering Context-Aware Personalised Services for Mobile Users. In *ICSOC 2015 (demo paper)*, Goa, India, 2015.

1.1.4 Tools and products

Softwares

- Coupled ensembles of neural networks: https://github.com/vabh/coupled_ensembles.
- BIOSOFT: microbiology laboratory automation tool, transfer to the Biokubes startup via SATT Linksium (Technology Transfer Accelerator).

Databases

The team has been involved in the creation of several data collections:

- TRECVID Semantic Indexing (SIN, 2010-2015) and Ad hoc Video Search (AVS, 2016-2021)¹
- Annotated collection of the GUIMUTEIC project²
- Annotated collection of the CLICIDE project³
- CLEF eHealth 2014-2018
- CLEF CMC 2016-2017

1.1.5 Editorial activities

Participation in editorial committees (books, collections, etc.)

- Georges Quénot, editor of the special issue of *Multimedia Tools and Applications (MTAP) on Content-Based Multimedia Indexing (CBMI) 2014*.
- *Journal of the Association for Information Science and Technology (JASIST)* (Lorraine Gœuriot)
- *Journal of Information Systems* (Lorraine Gœuriot)

¹ <https://www-nlpir.nist.gov/projects/tv2019/avs.html>

² <http://mrim.imag.fr/Guimuteic/index.html>

³ <http://mrim.imag.fr/Clicide/index.html>

1.1.6 Reviewing activities

Reviewing of articles

Members of the team have reviewed approximately 600 articles.

- International journals : C-ACM, Information Processing letters, JASIST, JIR, JIS, JMIR, Journal of Cognitive Computation, J. on Big Data and Security Challenges in IoTs and 5G Networks for Smart cities, etc
- International Conferences: ACM SIGIR, ACM SAC, ACM SIGSPATIAL, AIRE, ANT, BMC, BPM, CBMI, CIKM, CLEF, DATABASES, DEXA, ECIR, FDIA, ICME, ICSOC, IEEE AICCSA, IJMI, Int. Conf. on e-Business, Int. Conf. on Information Systems and Economic Intelligence, IPM, ISM, KDIR, MTAP, PLOS ONE, SKIMA, WISE, etc
- National journals : Doc numérique, Revue I3
- National conferences : Conférence Francophone sur les Architectures Logicielles (CAL), CORIA, INFORSID, AIME, TAL-RI, etc

Grant evaluation (public or charities)

Evaluation of national or international projects:

- Expertise of ANR projects (22)
- Member et vice-présidente of ANR commitee (2017, 2018, 2019),
- ANR : mid-term evaluation of projets (4)
- International projects evaluation : universidad austral de Chile, Fonds FQRNT Canada, Stic AMSUD, NIST/ITL.
- Evaluation for DIGITEO_DIGICOSME 2014 (1)
- Project evaluation for CIMI Labex (1)
- AAAS expert in USA (14).

Reviewing of research institutes

- Catherine Berrut has been reviewing 2 French research institutes as an expert with Hcéres (in 2015 and 2016).

Participation in institutional committees and juries (CNRS, Inserm, etc.)

- Catherine Berrut: HCERES scientific advisor since 2016.
- Ahmed Lbath: Expert at the ministère de l'enseignement supérieur de de la recherche (2007-2018).

1.1.7 Academic research grants

European (ERC, H2020, etc.) and international (NSF, JSPS, NIH, World Bank, FAO, etc.) grants - partnership

- European Union's Erasmus Mundus project "Sustainable e-tourism, Erasmus Mundus Action 2 programme". 2011-2014, 2M€ for the European partners (450 K€ for UJF and MRIM).

Other European grants - coordination

- EIRAP: ELIAS Subvention de l'ESF (European Science Foundation)- University of Amsterdam, 10/06/2016-31/12/2016 (UGA) Lorraine Gœuriot 5 000,00 €

Other European grants - partnership

National public grants (ANR, PHRC, FUI, INCA, etc.) - coordination

- Grant from French Embassy in USA, 01/01/2016-31/12/2017 (UGA) Ahmed LBATH 7 500 €
- PHC RIMS-FD with Tunisia (2014 à 2016), and LIRIS, LORIA labs, Catherine BERRUT (27000 €).

National public grants (ANR, PHRC, FUI, INCA, etc.) - partnership

- CAMOMILE-ANR 2011-CNRS, Collaborative annotation of multimodal, multilingual and multimedia documents 01/10/2012-30/11/2016, Georges Quénot, 249 999 €
- PHC ShootMyMind with Vietnam, Taiwan and Japon (2015 à 2017), Jean-Pierre Chevallet.
- GUIMUTEIC - FUI - RRA - FEDER, 04/12/2014-30/11/2018 (UGA), EEC commission through Région Rhône Alpes, Jean-Pierre CHEVALLET, 445 865 €
- QCOMPERE Consortium Quaero pour la Reconnaissance Multimodale des Personnes, 01/11/2010-31/08/2014, Université Joseph Fourier ANR MRIM-GETALP 81 000 €

Local grants (collectivités territoriales) - coordination

- Projet RESPIR ARC 2014, Région Auvergne-Rhône-Alpes (RARA), Philippe MULHEM 91 000 €

Local grants (collectivités territoriales) - partnership

PIA (labex, equipex etc.) grants - partnership

- PERSYVAL-DeCoRe-UGA Equipe Action - Deep Convolutional and Recurrent networks for image, speech, and text 01/06/2016-30/11/2019, (UGA) Georges Quénot, Laurent Besacier, Jakob Verbeek, Denis Pellerin, 250 000 €

Grants from foundations and charities (ARC, FMR, FRM, etc.) - partnership

- AYUSHI-CDC-UGA, Fonds national pour la Société Numérique, 01/01/2012-30/06/2014, Université Grenoble Alpes (UGA), Nathalie Denos, 49 286 €

1.1.8 Visiting senior scientists and post-doc

Foreign visiting scientists (15)

- Pr Saleh Basalamah, UQU University, KSA
- Pr. Nopasit Chokpitak, Chiang Mai University, Thailand
- Pr. Fabio Crestani, University of Lugano, Switzerland
- Pr. Marlon Dumas, University of Tartu, Estonia
- Dr Chris Greer, NIST, USA
- Dr Edouard Griffor, NIST, USA
- Pr. Chiraz Latiri, U. La Manouba, Tunisia

- Dr Francisco Lopez Orozc, de l'université Autonome de Ciudad Juarez, Mexico
- Dr Teresa Lynn, ADAPT Research Centre, Dublin City University
- Dr Mend-Amar Majig, University of Mongolia, Mongolia
- Dr Ram Sriram, NIST, USA
- Dr Pradorn Sureephong, Chiang Mai UNiversity, Thailand
- Dr. Kian Lam Tan, Sultan Idris Education University, Malaysia
- Dr. Chiraz Trabelsi, ISAMM, Tunisia
- Dr Santichai Wicha, ChiangRai University, Thailand
- Pr. Guido Zuccon, Queensland University of Technology, Australia

1.1.9 Scientific recognition

Prizes and/or distinctions

- Georges Quénot : Excellent Paper Award (Niwa & Takayanagi Award) for the paper [J18] in ITE Transactions on Media Technologies and Applications.
- Philippe Mulhem : Best Paper Award at CORIA 2019 [C26].
- Lorraine Gœuriot : Best Paper Award at CLEF 2016 [C14*].
- Ahmed Lbath: Best Demo Award at ACMsigspatial 2014 and Best Best Presentation Award at IEEE Big Data 2015.
- Georges Quénot : Outstanding Reviewer Award at ICMR 2017.
- Marie-Christine Fauvet : Palmes Académiques (2017).

Chair of learned and scientific societies

- Catherine Berrut: Membre du Conseil d'administration, du Conseil scientifique et trésorière de la Société Informatique de France,
- Catherine Berrut, Georges Quénot, Philippe, Lorraine Gœuriot, Jean-Pierre Chevallet: membre du comité d'orientation de ARIA (Association Francophone de Recherche d'Information et Applications),
- Ahmed Lbath: member of the executive committee of ACMsigspatial and the steering committee of AICCSA (ACS/IEEE International Conference on Computer Systems and Applications).
- Lorraine Gœuriot, membre de l'Information Retrieval Specialist Group (IRSG)

Invitations to meetings and symposia

- Ahmed Lbath: Invited speaker, University of Purdue, USA, 2014, Universities of Chiang Mai and Chiang Rai, Thailand oct. 2016.
- Ahmed Lbath: Keynote speaker at AICSSA16, IEEE ICDS17 and INTIS17
- Lorraine Gœuriot: Invited speaker BLAH 5, Japan, 2019.
- Georges Quénot: invited lectures at EARIA (École d'Automne en Recherche d'Information) 2014 and 2016.

- Catherine Berrut: invited lecture at École d'Automne en Recherche d'Information : Fondements et Applications 2014 Tunisie
- Catherine Berrut and Philippe Mulhem: invited lectures at Journées Analyse des réseaux sociaux : de la fouille à la RI et la Recommandation ARS Tunisie 2017

Members' long-term visits abroad

- Marie-Christine Fauvet: visits at Queensland University of Technology, Australia, 2015.
- Ahmed Lbath: visits at NIST, Washington, USA, 2014-2019.

1.1.10 Scientific animation

Organisations of meetings and symposia

- Lorraine Gœuriot and Georges Quénot: General Chairs of ECIR 2018. The whole MRIM team was part of the organizing committee.
- Lorraine Gœuriot and Catherine Berrut: Organization of ACM SIGIR 2019 (communication).
- Philippe Mulhem: General Chair of EARIA 2014.
- Philippe Mulhem: Scientific Chair of Journée Masterclass at CLEF 2018.
- Lorraine Gœuriot: organization of the MEDIR workshop at SIGIR 2016.

Scientific and steering committees

- Catherine Berrut: member of the ESSIR steering committee,
- Catherine Berrut: membre du comité pour ECIR TTA 2019 (best paper des 10 dernières années),
- Lorraine Gœuriot, member of the CLEF steering committee
- Georges Quénot: president of the CBMI Conference Steering Committee
- Georges Quénot: member of the TRECVID challenge Steering Committee⁴

1.2 Interaction of the unit with the non-academic world, impacts on economy, society, culture or health

1.2.1 Socio-economic interactions

Industrial and R&D contracts

- BIOKUBES AUTOMATION-SATT-UGA, CONVENTION DE MATURATION TECHNO-FLASH N°CMTF180005 - Identifiant : 170005I, 01/11/2018-1/10/2019, (UGA) SATT (Société d'Accélération de Transfert de Technologies, Ahmed LBATH 69 000.00 €
- BIOSOFT, Ahmed Lbath, 1/11/2017-31/10/2018, 56 K€ (LSI) + 10K€ (Linksium)
- CLICIDE, Un Clic un guide 03/02/2012-13/10/2014, CNRS, Pôle de compétitivité IMAG-INOVE, Philippe MULHEM, 46 432,00 €

⁴<https://trecvid.nist.gov/contact.html>

- MILLIWORKS-SATT-UGA, Convention d'incubation, 01/02/2016-31/07/2016, (UGA) SATT (Société d'Accélération de Transfert de Technologies), J.P.CHEVALLET, 30 000,00 €
- GLOBE VIP 7125 00 MRIM, Entreprise, Contrats de recherche industriels, 21/12/2017-21/05/2018, FLORALIS JP Chevallet 7 500.00 €

We have participated to the GUIMUTEIC industrial project, leaded by the Ophrys company. This project aim to develop a new museum tour guides that localize the user in the museum and recognize the artifact in front of the visitor. User then can interact with this tour guide using gesture, for example, to start the audio description related to this artifact. Our contribution in this project is a new approach of deep neural network trained to recognize object instances instead of class. The problem to solve is to enable the training a deep network using very few image example of the museum artifact. Moreover, our work has emphasis the building of the smallest possible network in order to perform the object and gesture recognition using limited resources of a smart phone.

We have realized technology transfers with industry. On the theme of Mobile and Cultural Heritage, we have work with the the GlobeVIP from Lyon. This company is specialized in mobile information access using image. We have transfer to this company the library MIDL (Mobile Image Detection Library) that is used to quickly recognize simple object using extra low computing and storage resources. The technique used is the description of images given key-points associated with bit string describing the key-points. Our contribution is a dedicated data structure based on Locality Sensitive Binary Hashing that retrieve stored relevant key-points with limited smart phone resources. This library is written in C++ programming language.

The logical model of K. Abdulahhad [PhD12] has been transferred to the Milliworcks company. This software called "Semantic Information Retrieval Coping with Logical Expressions" (SIRCLE) indexes typed documents and exploit the matching function proposed in the PhD thesis based on a logic model. The software transfer process has been founded by a local funding structure (Grenoble Linksiium, SATT TechTransfert & Création Startups) for the set up of a Java version of this system called jsIRCLE.

Start-ups

We are currently under a transfer process to a start-up compagny, of a system that index set of multi-structured data set, link to web pages, with knowledge representation that solve precise and complex queries.

We are also under another transfer process funded by the SATT Linksiium (Technology Transfer Accelerator) in order to set up a startup named Biokubes. Biokubes is a startup focused on the development and commercialization of microbiology laboratory automation. The Biokubes platform combines advanced instrumentation and bioinformatics to provide modular, partial or full automation handling, management, incubation and analysis of microbial samples. The conducted work consist of the design and the development of a software named Biosoft that will: i) improve the existing software tool based on the analysis at the current validation stage, ii) and design and develop a framework for data collection and machine learning based system to improve the image and video analysis. Biosoft will incorporate a generic approach in order to interconnect a plurality of lab instruments.

1.2.2 Expertise

Consulting

- Jean-Pierre Chevallet and Georges Quénot: consulting for GlobeVIP (Company).

1.2.3 Public outreach

Radio broadcasts, TV shows, magazines and newspaper

- Lorraine Gœuriot: presentation at the Pint of Science festival, 2016⁵.

Journal articles, interviews, book edition, videos, other popularization outputs, debates on science and society, etc.

1.3 Involvement of the unit and of each team in training through research

1.3.1 Educational outputs

Books

Georges Quénot: Data Science [B2], course at the master level.

1.3.2 Scientific productions (articles, books, etc.) from theses

Scientific productions (articles, books, etc.) from theses

120.

Mean number of publications per student (Biology & Science and technology only)

5.

1.3.3 Training

Habilitated (HDR) scientists (6)

Catherine Berrut, Marie-Christine Fauvet, Ahmed Lbath, Georges Quénot, Jean-Pierre Chevallet and Philippe Mulhem.

HDR obtained during the period

None.

PhD students (21)

We had 21 PhD students over the period, of which 14 of them have defended (see below). The other 7 are listed in the team members section.

⁵<https://www.echosciences-grenoble.fr/evenements/pint-of-science-2016>

PhD students benefiting from a specific doctoral contract

Defended PhDs

14 PhDs were defended over the period.

| Name | Year | Situation in March 2019 |
|----------------------|------|--|
| Demeke Ayele | 2014 | Professeur en Éthiopie. |
| Kian Lam Tan | 2014 | Maître de Conférence en Malaisie. |
| Karam Abdulahhad | 2014 | Post-doc GESIS - Leibniz institute for the Social Sciences, Allemagne. |
| Nadia Derbas | 2014 | Ingénieur de Recherche chez VISEO. |
| Abdelkader Hamadi | 2014 | Maître de Conférence à l'Université de Mostaghanem en Algérie. |
| Isaac Caicedo-Castro | 2015 | Maître de Conférence en Colombie. |
| Pathathai Na Lumpoon | 2015 | Associate professor à l'université de Chiang Mai, Vice doyen |
| Uyanga Sukhbaatar | 2016 | Assistant Professor à l'université nationale de Mongolie. |
| Mou Lei | 2016 | Professeur en Chine. |
| Mateusz Budnik | 2017 | Post-doc à l'IRISA Rennes. |
| Mohannad Almasri | 2017 | Post-doc au LIG Grenoble, équipe SLIDE. |
| Olivera Kotevska | 2018 | Post-doc au laboratoire ITL /NIST à Washington DC, USA. |
| Maxime Portaz | 2018 | Ingénieur chez Qwant |
| Faizan Ur Rehman | 2018 | Assistant Professor à ABHA University en Arabie Saoudite. |

Mean PhD duration

The mean PhD duration is 44 months, which is comparable to the average duration at the MSTII doctoral school (42 months).

Internships (M1, M2)

30+ master or engineer students made an internship in the MRIM team over the period.

People in charge for a mention or a master's degree course (3)

- Georges Quénot: responsable of the MOSIG M2 option (8 courses) Artificial Intelligence and the Web (AIW)⁶. MOSIG is a two-year European Standard (LMD) master program where courseware is offered in English.
- Jean-Pierre Chevallet, Philippe Mulhem and Georges Quénot: MOSIG M2 course on Information access and retrieval (IAR)⁷
- Georges Quénot: M2 GI course on Documents multimédia : description et recherche automatique⁸

⁶<https://master-informatique.univ-grenoble-alpes.fr/main-menu/academic-program/master-of-science-mosig/artificial-intelligence-and-the-web/>

⁷<http://formations.univ-grenoble-alpes.fr/fr/catalogue/master-XB/sciences-technologies-sante-STS/master-informatique-program-master-informatique/parcours-master-of-science-in-informatics-at-grenoble-mosig-subprogram-master-of-science-information-access-and-retrieval-IGDF9Z8M.html>

⁸<http://formations.univ-grenoble-alpes.fr/fr/catalogue/master-XB/sciences-technologies-sante-STS/master-informatique-program-master-informatique/parcours-genie-informatique-alternance-2e-annee-subprogram-parcours-genie-informatique-documents-multimedia-description-et-recherche-automatique-IFQV03ZG.html>