



IN PARTNERSHIP WITH:
CNRS

**Institut national des sciences
appliquées de Rennes**

Université Rennes 1

Activity Report 2016

Project-Team LINKMEDIA

Creating and exploiting explicit links between
multimedia fragments

IN COLLABORATION WITH: Institut de recherche en informatique et systèmes aléatoires (IRISA)

RESEARCH CENTER
Rennes - Bretagne-Atlantique

THEME
**Vision, perception and multimedia
interpretation**

Table of contents

1. Members	1
2. Overall Objectives	2
2.1. Context	2
2.2. Scientific objectives	2
3. Research Program	3
3.1. Scientific background	3
3.2. Workplan	3
3.2.1. Unsupervised motif discovery	3
3.2.2. Description and structuring	4
3.2.3. Linking and collection data model	4
4. Application Domains	4
4.1. Asset management in the entertainment business	4
4.2. Multimedia Internet	5
4.3. Multiscreen TV	5
4.4. E-learning	5
5. Highlights of the Year	5
6. New Software and Platforms	5
7. New Results	6
7.1. Unsupervised motif and knowledge discovery	6
7.1.1. Multimodal person discovery in TV broadcasts	6
7.1.2. Efficient similarity self-join for near-duplicate video detection	6
7.1.3. Recommendation systems with matrix factorization	6
7.2. Multimedia content description and structuring	7
7.2.1. Hierarchical topic structuring	7
7.2.2. Multimedia-inspired descriptors for time series classification	7
7.2.3. Early time series classification	7
7.3. Content-based information retrieval	7
7.3.1. Bi-directional embeddings for cross-modal content matching	7
7.3.2. Intrinsic dimensions in language information retrieval	8
7.3.3. Evaluation of distributional thesauri	8
7.3.4. Scaling group testing similarity search	8
7.3.5. Large-scale similarity search using matrix factorization	9
7.4. Linking, navigation and analytics	9
7.4.1. Opinion similarity and target extraction	9
7.4.2. Reinforcement and fake detection in social networks	9
7.4.3. Multimodal video hyperlinking	10
7.4.4. User-centric evaluation of hyperlinked news content	10
7.5. Miscellaneous	10
7.5.1. Bidirectional GRUs in spoken dialog	10
7.5.2. Kernel principal components analysis with extreme learning machines	11
7.5.3. Pronunciation adaptation for spontaneous speech synthesis	11
7.6. Participation in benchmarking initiatives	11
8. Bilateral Contracts and Grants with Industry	11
9. Partnerships and Cooperations	11
9.1. Regional Initiatives	11
9.1.1. CominLabs Project Linking Media in Acceptable Hypergraphs (LIMAH)	11
9.1.2. CominLabs Project BigCLIN	12
9.2. National Initiatives	12
9.2.1. ANR Project IDFRAud	12

9.2.2. FUI 19 NexGenTV	12
9.3. European Initiatives	13
9.4. International Initiatives	13
9.4.1. Inria Associate Teams Not Involved in an Inria International Labs	13
9.4.2. Inria International Partners	13
9.4.3. Participation in Other International Programs	13
9.5. International Research Visitors	14
9.5.1. Visits of International Scientists	14
9.5.2. Visits to International Teams	14
10. Dissemination	14
10.1. Promoting Scientific Activities	14
10.1.1. Scientific Events Organisation	14
10.1.1.1. General Chair, Scientific Chair	14
10.1.1.2. Member of the Organizing Committees	15
10.1.2. Scientific Events Selection	15
10.1.2.1. Chair of Conference Program Committees	15
10.1.2.2. Member of Conference Program Committees	15
10.1.2.3. Reviewer	15
10.1.3. Journal	16
10.1.3.1. Member of the Editorial Boards	16
10.1.3.2. Reviewer - Reviewing Activities	16
10.1.4. Invited Talks	16
10.1.5. Leadership within the Scientific Community	17
10.1.6. Scientific Expertise	17
10.1.7. Research Administration	17
10.2. Teaching - Supervision - Juries	17
10.2.1. Teaching	17
10.2.2. Supervision	18
10.2.3. Juries	19
10.3. Popularization	19
11. Bibliography	19

Project-Team LINKMEDIA

Creation of the Project-Team: 2014 July 01

Keywords:

Computer Science and Digital Science:

- 3.3.2. - Data mining
- 3.3.3. - Big data analysis
- 5.3.3. - Pattern recognition
- 5.4.1. - Object recognition
- 5.4.3. - Content retrieval
- 5.7. - Audio modeling and processing
- 5.8. - Natural language processing
- 8.2. - Machine learning
- 8.4. - Natural language processing

Other Research Topics and Application Domains:

- 9. - Society and Knowledge

1. Members

Research Scientists

Guillaume Gravier [Team leader, CNRS, Senior Researcher, HDR]
Laurent Amsaleg [CNRS, Researcher, HDR]
Yannis Avrithis [Inria, Advanced Research position]
Vincent Claveau [CNRS, Researcher]
Teddy Furon [Inria, Researcher]

Faculty Members

Ewa Kijak [Univ. Rennes I, Associate Professor]
Simon Malinowski [Univ. Rennes I, Associate Professor]
Christian Raymond [INSA Rennes, Associate Professor]
Pascale Sébillot [INSA Rennes, Professor, HDR]

Engineers

Gabriel Sargent [CNRS]
Ronan Sicre [CNRS]
Arnaud Touboulic [CNRS, from Feb. 2016]

PhD Students

Raghavendran Balu [Inria, until Oct. 2016]
Rémi Bois [CNRS]
Ricardo Carlini Sperandio [PUC Minas]
Clément Dalloux [CNRS, from Dec. 2016]
Mikail Demirdelen [Univ. Rennes I, from Sep. 2016]
Ahmet Iscen [Inria]
Grégoire Jadi [Univ. Nantes]
Cédric Maigrot [Univ. Rennes I]
Oriane Simeoni [Inria, from Oct. 2016]
Vedran Vukotić [INSA Rennes]

Administrative Assistant

Aurélie Patier [Univ. Rennes I]

2. Overall Objectives

2.1. Context

Linked media appears today as a major challenge, with numerous potential applications in all areas of multimedia. The strong increase of ubiquitous access to the Internet and the resulting convergence of media on the network open countless opportunities for linked media and reinforce the key role of such a challenge. New applications centered on the notion of linked media are emerging today, such as second screen applications and recommendation services. However, because of the lack of adequate technology, linking related content is mostly deferred to human operators in current applications or to user behavior analysis, e.g., via collaborative filtering, thus indirectly considering the content. This fact severely limits the opportunities offered by a web of media, in terms of creativity, scalability, representativeness and completeness, thus negatively impacting the spread of linked media and the development of innovative services in the Internet of media.

Most of the research effort in automatic multimedia content analysis has been devoted so far to describing and indexing content on which core tasks around information retrieval and recommendation are built to develop multimedia applications. This general philosophy mostly reposes on a vision where documents are considered as isolated entities, i.e., as a basic unit which is indexed or analyzed regardless of other content items and of context. Considering documents in isolation has enabled key progress in content-based analysis and retrieval on a large scale: e.g., design of generic descriptors, efficient techniques for content-based analysis, fast retrieval methodology. But ignoring the links, implicit or explicit, between content items also appears as a rather strong assumption with direct consequences on algorithms and applications, both in terms of performance and in terms of possibilities.

2.2. Scientific objectives

LINKMEDIA investigates a number of key issues related to multimedia collections structured with explicit links: Can we discover what characterizes a collection and makes its coherence? Are there repeating motifs that create natural links and which deserve characterization and semantic interpretation? How to explicitly create links from pairwise distances? What structure should a linked collection have? How do we explain the semantic of a link? How explicit links can be used to improve information retrieval? To improve user experience? In this general framework, the global objective of LINKMEDIA is to develop the scientific, methodological and technological foundations facilitating or automating the creation, the description and the exploitation of multimedia collections structured with explicit links. In particular, we target a number of key contributions in the following areas:

- designing efficient methods dedicated to multimedia indexing and unsupervised motif discovery: efficiently comparing content items on a large scale and finding repeating motifs in an unsupervised manner are two key ingredients of multimedia linking based on a low-level representation of the content;
- improving techniques for structuring and semantic description: better description of multimedia content at a semantic—i.e., human interpretable—level, making explicit the implicit structure when it exists, is still required to make the most of multimedia data and to facilitate the creation of links to a precise target at a semantic level;
- designing and experimenting approaches to multimedia content linking and collection structuring: exploiting low-level and semantic content-based proximity to create explicit links within a collection requires specific methodology departing from pairwise comparison and must be confronted with real data;
- studying new paradigms for the exploitation of linked multimedia content as well as new usages: explicit links within media content collections change how such data is processed by machines and ultimately consumed by humans in ways that have yet to be invented and studied.

3. Research Program

3.1. Scientific background

LINKMEDIA is a multidisciplinary research team, with multimedia data as the main object of study. We are guided by the data and their specificity—semantically interpretable, heterogeneous and multimodal, available in large amounts, unstructured and disconnected—, as well as by the related problems and applications.

With multimedia data at the center, orienting our choices of methods and algorithms and serving as a basis for experimental validation, the team is directly contributing to the following scientific fields:

- multimedia: content-based analysis; multimodal processing and fusion; multimedia applications;
- computer vision: compact description of images; object and event detection;
- natural language processing: topic segmentation; information extraction;
- information retrieval: high-dimensional indexing; approximate k-nn search; efficient set comparison.

LINKMEDIA also takes advantage of advances in the following fields, adapting recent developments to the multimedia area:

- signal processing: image processing; compression;
- machine learning: deep architectures; structured learning; adversarial learning;
- security: data encryption; differential privacy;
- data mining: time series mining and alignment; pattern discovery; knowledge extraction.

3.2. Workplan

Research activities in LINKMEDIA are organized along three major lines of research which build upon the scientific domains already mentioned.

3.2.1. *Unsupervised motif discovery*

As an alternative to supervised learning techniques, unsupervised approaches have emerged recently in multimedia with the goal of discovering directly patterns and events of interest from the data, in a totally unsupervised manner. In the absence of prior knowledge on what we are interested in, meaningfulness can be judged based on one of three main criteria: unexpectedness, saliency and recurrence. This last case posits that repeating patterns, known as motifs, are potentially meaningful, leading to recent work on the unsupervised discovery of motifs in multimedia data [54], [52], [53].

LINKMEDIA seeks to *develop unsupervised motif discovery approaches which are both accurate and scalable*. In particular, we consider the discovery of repeating objects in image collections and the discovery of repeated sequences in video and audio streams. Research activities are organized along the following lines:

- developing the scientific basis for scalable motif discovery: sparse histogram representations; efficient co-occurrence counting; geometry and time aware indexing schemes;
- designing and evaluating accurate and scalable motif discovery algorithms applied to a variety of multimedia content: exploiting efficient geometry or time aware matching functions; fast approximate dynamic time warping; symbolic representations of multimedia data, in conjunction with existing symbolic data mining approaches;
- developing methodology for the interpretation, exploitation and evaluation of motif discovery algorithms in various use-cases: image classification; video stream monitoring; transcript-free natural language processing (NLP) for spoken document.

3.2.2. Description and structuring

Content-based analysis has received a lot of attention from the early days of multimedia, with an extensive use of supervised machine learning for all modalities [56], [48]. Progress in large scale entity and event recognition in multimedia content has made available general purpose approaches able to learn from very large data sets and performing fairly decently in a large number of cases. Current solutions are however limited to simple, homogeneous, information and can hardly handle structured information such as hierarchical descriptions, tree-structured or nested concepts.

LINKMEDIA aims at *expanding techniques for multimedia content modeling, event detection and structure analysis*. The main transverse research lines that LINKMEDIA will develop are as follows:

- context-aware content description targeting (homogeneous) collections of multimedia data: latent variable discovery; deep feature learning; motif discovery;
- secure description to enable privacy and security aware multimedia content processing: leveraging encryption and obfuscation; exploring adversarial machine learning in a multimedia context; privacy-oriented image processing;
- multilevel modeling with a focus on probabilistic modeling of structured multimodal data: multiple kernels; structured machine learning; conditional random fields.

3.2.3. Linking and collection data model

Creating explicit links between media content items has been considered on different occasions, with the goal of seeking and discovering information by browsing, as opposed to information retrieval via ranked lists of relevant documents. Content-based link creation has been initially addressed in the hypertext community for well-structured texts [47] and was recently extended to multimedia content [57], [51], [50]. The problem of organizing collections with links remains mainly unsolved for large heterogeneous collections of unstructured documents, with many issues deserving attention: linking at a fine semantic grain; selecting relevant links; characterizing links; evaluating links; etc.

LINKMEDIA targets pioneering research on media linking by *developing scientific ground, methodology and technology for content-based media linking* directed to applications exploiting rich linked content such as navigation or recommendation. Contributions are concentrated along the following lines:

- algorithmic of linked media for content-based link authoring in multimedia collections: time-aware graph construction; multimodal hypergraphs; large scale k-nn graphs;
- link interpretation and characterization to provide links semantics for interpretability: text alignment; entity linking; intention vs. extension;
- linked media usage and evaluation: information retrieval; summarization; data models for navigation; link prediction.

4. Application Domains

4.1. Asset management in the entertainment business

Regardless of the ingestion and storage issues, media asset management—archiving, describing and retrieving multimedia content—has turned into a key factor and a huge business for content and service providers. Most content providers, with television channels at the forefront, rely on multimedia asset management systems to annotate, describe, archive and search for content. So do archivists such as the Institut National de l'Audiovisuel, the Nederlands Instituut voor Beeld en Geluid or the British Broadcast Corporation, as well as media monitoring companies, such as Yacast in France. Protecting copyrighted content is another aspect of media asset management.

4.2. Multimedia Internet

One of the most visible application domains of linked multimedia content is that of multimedia portals on the Internet. Search engines now offer many features for image and video search. Video sharing sites also feature search engines as well as recommendation capabilities. All news sites provide multimedia content with links between related items. News sites also implement content aggregation, enriching proprietary content with user-generated content and reactions from social networks. Most public search engines and Internet service providers offer news aggregation portals.

4.3. Multiscreen TV

The convergence between television and the Internet has accelerated significantly over the past few years, with the democratization of TV on-demand and replay services and the emergence of social TV services and multiscreen applications. These evolutions and the consequently ever growing number of innovative applications offer a unique playground for multimedia technologies. Recommendation plays a major role in connected TV. Enriching multimedia content, with explicit links targeting either multimedia material or knowledge databases, appears as a key feature in this context, at the core of rich TV and second screen applications.

4.4. E-learning

On-line courses are rapidly gaining interest with the recent movement for massive open on-line courses (MOOCs). Such courses usually aggregate multimedia material, such as a video of the course with handouts and potentially text books, exercises and other related resources. This setting is very similar to that of the media aggregation sites though in a different domain. Automatically analyzing and describing video and textual content, synchronizing all material available across modalities, creating and characterizing links between related material or between different courses are all necessary features for on-line courses authoring.

5. Highlights of the Year

5.1. Highlights of the Year

LINKMEDIA ranked first at the TRECVID 2016 Hyperlinking international benchmark [12].

LINKMEDIA is selected as the organizer of the IEEE Workshop on Information Forensics and Security in 2017.

LINKMEDIA deeply involved in the winning bid for the organization of the ACM Conf. on Multimedia in 2019.

6. New Software and Platforms

6.1. AllGO multimedia web services

Participants: Guillaume Gravier [correspondent], Gabriel Sargent.

Available at <http://allgo.irisa.fr>, the AllGO platform allows for the easy deployment of the technology developed in the team as web services. Based on the AllGO infrastructure, LINKMEDIA has continued making available a number of web services related to multimedia content analysis. In 2016, we revamped the multimedia-related web service offer, making all services interoperable and broadening the scope of services made available (silence detection, face detection, text-based fragmentation). We also made available **short videos that demonstrate the usefulness of our web services**.

7. New Results

7.1. Unsupervised motif and knowledge discovery

7.1.1. *Multimodal person discovery in TV broadcasts*

Participants: Guillaume Gravier, Gabriel Sargent, Ronan Sicre.

Work in collaboration with Silvio J. Guimarães, Gabriel B. de Fonseca and Izabela Lyon Freire, PUC Minas, in the framework of the Inria Associate Team MOTIF.

Pursuing efforts initiated in 2015 in the framework of the MediaEval benchmark on Multimodal Person Discovery, we investigated graph-based approaches to name the persons on screen and speaking in TV broadcasts with no prior information, leveraging text overlays, speech transcripts as well as face and voice comparison. We adopted a graph-based representation of speaking faces and investigated two tag-propagation approaches to associate overlays co-occurring with some speaking faces to other visually or audiovisually similar speaking faces. Given a video, we first build a graph from the detected speaking faces (nodes) and their audiovisual similarities (edges). Each node is associated to its co-occurring overlays (tags) when they exist. Then, we consider two tag-propagation approaches, respectively based on a random walk strategy and on Kruskal's minimum spanning tree algorithm for node clustering [28].

7.1.2. *Efficient similarity self-join for near-duplicate video detection*

Participants: Laurent Amsaleg, Guillaume Gravier.

Work in collaboration with Henrique B. da Silva, Silvio J. Guimarães, Zenilto do Patrocino Jr., PUC Minas, and Arnaldo de A. Araújo, UFMG, in the framework of the Inria Associate Team MOTIF.

The huge amount of redundant multimedia data, like video, has become a problem in terms of both space and copyright. Usually, the methods for identifying near-duplicate videos are neither adequate nor scalable to find pairs of similar videos. Similarity self-join operation could be an alternative to solve this problem in which all similar pairs of elements from a video dataset are retrieved. Methods for similarity self-join however exhibit poor performance when applied to high-dimensional data. In [33], we propose a new approximate method to compute similarity self-join in sub-quadratic time in order to solve the near-duplicate video detection problem. Our strategy is based on clustering techniques to find out groups of videos which are similar to each other.

7.1.3. *Recommendation systems with matrix factorization*

Participants: Raghavendran Balu, Teddy Furon.

Matrix factorization is a prominent technique for approximate matrix reconstruction and noise reduction. Its common appeal is attributed to its space efficiency and its ability to generalize with missing information. For these reasons, matrix factorization is central to collaborative filtering systems. In the real world, such systems must deal with million of users and items, and they are highly dynamic as new users and new items are constantly added. Factorization techniques, however, have difficulties to cope with such a demanding environment. Whereas they are well understood with static data, their ability to efficiently cope with new and dynamic data is limited. Scaling to extremely large numbers of users and items is also problematic. In [10], we propose to use the count sketching technique for representing the latent factors with extreme compactness, facilitating scaling.

In [11], we discovered that sketching techniques implicitly provide differential privacy guarantees thanks to the inherent randomness of the data structure. Collaborative filtering is a popular technique for recommendation system due to its domain independence and reliance on user behavior data alone. But the possibility of identification of users based on these personal data raise privacy concerns. Differential privacy aims to minimize these identification risks by adding controlled noise with known characteristics. The addition of noise impacts the utility of the system and does not add any other value to the system other than enhanced privacy.

7.2. Multimedia content description and structuring

7.2.1. Hierarchical topic structuring

Participants: Guillaume Gravier, Pascale Sébillot.

In [37], we investigated the potential of a topical structure of text-like data that we recently proposed [55] in the context of summarization and anchor detection in video hyperlinking. This structure is produced by an algorithm that exploits temporal distributions of words through word burst analysis to generate a hierarchy of topically focused fragments. The obtained hierarchy aims at filtering out non-critical content, retaining only the salient information at various levels of detail. For the tasks we choose to evaluate the structure on, the loss of important information is highly damaging. We show that the structure can actually improve the results of summarization or at least maintain state-of-the-art results, while for anchor detection it leads us to the best precision in the context of the Search and Anchoring in Video Archives task at MediaEval. The experiments were carried on written text and a more challenging corpus containing automatic transcripts of TV shows.

7.2.2. Multimedia-inspired descriptors for time series classification

Participant: Simon Malinowski.

The SIFT framework has shown to be effective in the image classification context. Recently, we designed a bag-of-words approach based on an adaptation of this framework to time series classification. It relies on two steps: SIFT-based features are first extracted and quantized into words; histograms of occurrences of each word are then fed into a classifier. In [38], we investigated techniques to improve the performance of bag-of-temporal-SIFT-words: dense extraction of keypoints and different normalizations of Bag-of-Words histograms. Extensive experiments have shown that our method significantly outperforms nearly all tested standalone baseline classifiers on publicly available UCR datasets. In [23], we also investigate the use of convolutional neural networks (CNN) for time series classification. Such networks have been widely used in many domains like computer vision and speech recognition, but only a little for time series classification. We have designed a convolutional neural network that consists of two convolutional layers. One drawback with CNN is that they need a lot of training data to be efficient. We propose two ways to circumvent this problem: designing data-augmentation techniques and learning the network in a semi-supervised way using training time series from different datasets. These techniques are experimentally evaluated on a benchmark of time series datasets.

7.2.3. Early time series classification

Participant: Simon Malinowski.

In time series classification, two antagonist notions are at stake. On the one hand, in most cases, the sooner the time series is classified, the higher the reward. On the other hand, an early classification is more likely to be erroneous. Most of the early classification methods have been designed to take a decision as soon as a sufficient level of reliability is reached. However, in many applications, delaying the decision with no guarantee that the reliability threshold will be met in the future can be costly. Recently, a framework dedicated to optimizing the trade-off between classification accuracy and the cost of delaying the decision was proposed, together with an algorithm that decides online the optimal time instant to classify an incoming time series. On top of this framework, we have built in [29] two different early classification algorithms that optimize the trade-off between decision accuracy and the cost of delaying the decision. These algorithms are non-myopic in the sense that, even when classification is delayed, they can provide an estimate of when the optimal classification time is likely to occur. Our experiments on real datasets demonstrate that the proposed approaches are more robust than existing methods.

7.3. Content-based information retrieval

7.3.1. Bi-directional embeddings for cross-modal content matching

Participants: Guillaume Gravier, Christian Raymond, Vedran Vukotić.

Common approaches to problems involving multiple modalities (classification, retrieval, hyperlinking, etc.) are early fusion of the initial modalities and crossmodal translation from one modality to the other. Recently, deep neural networks, especially deep autoencoders, have proven promising both for crossmodal translation and for early fusion via multimodal embedding. In [31], we propose a flexible cross-modal deep neural network architecture for multimodal and crossmodal representation. By tying the weights of two deep neural networks, symmetry is enforced in central hidden layers thus yielding a multimodal representation space common to the two original representation spaces. The proposed architecture is evaluated in multimodal query expansion and multimodal retrieval tasks within the context of video hyperlinking. In [32], we extend the approach, focusing on the evaluation of a good single-modal continuous representations both for textual and for visual information. word2vec and paragraph vectors are evaluated for representing collections of words, such as parts of automatic transcripts and multiple visual concepts, while different deep convolutional neural networks are evaluated for directly embedding visual information, avoiding the creation of visual concepts. We evaluate methods for multimodal fusion and crossmodal translation, with different single-modal pairs, in the task of video hyperlinking.

7.3.2. *Intrinsic dimensions in language information retrieval*

Participant: Vincent Claveau.

Examining the properties of representation spaces for documents or words in information retrieval (IR) brings precious insights to help the retrieval process. Recently, several authors have studied the real dimensionality of the datasets, called intrinsic dimensionality, in specific parts of these spaces. In [34], we propose to revisit this notion through a coefficient called α in the specific case of IR and to study its use in IR tasks. More precisely, we show how to estimate α from IR similarities and to use it in representation spaces used for documents and words. Indeed, we prove that α may be used to characterize difficult queries. We moreover show that this intrinsic dimensionality notion, applied to words, can help to choose terms to use for query expansion.

7.3.3. *Evaluation of distributional thesauri*

Participants: Vincent Claveau, Ewa Kijak.

With the success of word embedding methods, all the fields of distributional semantics have experienced a renewed interest. Beside the famous word2vec, recent studies have presented efficient techniques to build distributional thesaurus, including our work on information retrieval (IR) tools and concepts to build a thesaurus [14]. In [13], we address the problem of the evaluation of such thesauri or embedding models. Several evaluation scenarii are considered: direct evaluation through reference lexicons and specially crafted datasets, and indirect evaluation through a third party tasks, namely lexical substitution and Information Retrieval. Through several experiments, we first show that the recent techniques for building distributional thesaurus outperform the word2vec approach, whatever the evaluation scenario. We also highlight the differences between the evaluation scenarii, which may lead to very different conclusions when comparing distributional models. Last, we study the effect of some parameters of the distributional models on these various evaluation scenarii.

7.3.4. *Scaling group testing similarity search*

Participants: Laurent Amsaleg, Ahmet Iscen, Teddy Furon.

The large dimensionality of modern image feature vectors, up to thousands of dimensions, is challenging high dimensional indexing techniques. Traditional approaches fail at returning good quality results within a response time that is usable in practice. However, similarity search techniques inspired by the group testing framework have recently been proposed in an attempt to specifically defeat the curse of dimensionality. Yet, group testing does not scale and fails at indexing very large collections of images because its internal procedures analyze an excessively large fraction of the indexed data collection. In [16], we identify these difficulties and proposes extensions to the group testing framework for similarity searches that allow to handle larger collections of feature vectors. We demonstrate that it can return high quality results much faster compared to state-of-the-art group testing strategies when indexing truly high-dimensional features that are indeed hardly indexable with traditional indexing approaches.

We also discovered that group testing helps in enforcing security and privacy in identification. We detail a particular scheme based on embedding and group testing. Whereas embedding poorly protects the data when used alone, the group testing approach makes it much harder to reconstruct the data when combined with embedding. Even when curious server and user collude to disclose the secret parameters, they cannot accurately recover the data. Our approach reduces as well the complexity of the search and the required storage space. We show the interest of our work in a benchmark biometrics dataset [17], where we verify our theoretical analysis with real data.

7.3.5. *Large-scale similarity search using matrix factorization*

Participants: Ahmet Iscen, Teddy Furon.

Work in collaboration with Michael Rabbat, McGill University, Montréal.

We consider the image retrieval problem of finding the images in a dataset that are most similar to a query image. Our goal is to reduce the number of vector operations and memory for performing a search without sacrificing accuracy of the returned images. In [18], we adopt a group testing formulation and design the decoding architecture using either dictionary learning or eigendecomposition. The latter is a plausible option for small-to-medium sized problems with high-dimensional global image descriptors, whereas dictionary learning is applicable in large-scale scenarios. Experiments with standard image search benchmarks, including the Yahoo100M dataset comprising 100 million images, show that our method gives comparable (and sometimes better) accuracy compared to exhaustive search while requiring only 10 % of the vector operations and memory. Moreover, for the same search complexity, our method gives significantly better accuracy compared to approaches based on dimensionality reduction or locality sensitive hashing.

7.4. Linking, navigation and analytics

7.4.1. *Opinion similarity and target extraction*

Participants: Vincent Claveau, Grégoire Jadi.

Work in collaboration with Laura Monceaux and Béatrice Daille, LINA, Nantes.

In [19], we propose to evaluate the lexical similarity information provided by word representations against several opinion resources using traditional information retrieval tools. Word representation have been used to build and to extend opinion resources, such as lexicon and ontology, and their performance have been evaluated on sentiment analysis tasks. We question this method by measuring the correlation between the sentiment proximity provided by opinion resources and the semantic similarity provided by word representations using different correlation coefficients. We also compare the neighbors found in word representations and list of similar opinion words. Our results show that the proximity of words in state-of-the-art word representations is not very effective to build sentiment similarity.

In [20], we present the development of an opinion target extraction system in English and transpose it to the French language. In addition, we realize an analysis of the features and their effectiveness in English and French which suggest that it is possible to build an opinion target extraction system independant of the domain. Finally, we propose a comparative study of the errors of our systems in both English and French and propose several solutions to these problems.

7.4.2. *Reinformation and fake detection in social networks*

Participants: Vincent Claveau, Ewa Kijak, Cédric Maigrot.

Traditional media are increasingly present on social networks, but these usual sources of information are confronted with other sources called reinformation sources. These last ones sometimes tend to distort the information relayed to match their ideologies, rendering it partially or totally false. In [25], we conduct a study pursuing two goals: first, we present a corpus containing Facebook messages issued from both types of media sources; secondly, we propose some experiments in order to automatically detect reinformation messages. In particular, we investigate the influence of shallow features versus features more specifically describing the message content. We also developed a multi-modal hoax detection system composed of text, source, and image analysis [24]. As hoax can be very diverse, we want to analyze several modalities to better detect them. This system is applied in the context of the Verifying Multimedia Use task of MediaEval 2016. Experiments show the performance of each separated modality as well as their combination.

7.4.3. *Multimodal video hyperlinking*

Participants: Rémi Bois, Guillaume Gravier, Christian Raymond, Pascale Sébillot, Ronan Sicre, Vedran Vukotić.

Pursuing previous work on video hyperlinking and recent advances in multimodal content matching [32], we benchmarked a full video hyperlinking system in the framework of the TRECVID international benchmark [12]. The video hyperlinking task aims at proposing a set of video segments, called targets, to complement a query video segment defined as anchor. The 2016 edition of the task encouraged participants to use multiple modalities. In this context, we chose to submit four runs in order to assess the pros and cons of using two modalities instead of a single one and how crossmodality differs from multimodality in terms of relevance. The crossmodal run performed best and obtained the best precision at rank 5 among participants. In parallel, we also demonstrated that, in this framework, multimodal and crossmodal approaches offer significantly more diversity in the set of target proposed than classical information retrieval based approaches where all modalities are combined. We compared bidirectional multimodal embeddings [31] with multimodal LDA approaches as experimented last year in TRECVID [49]. The former offers more accurate matching, the latter exhibiting slightly more diversity.

7.4.4. *User-centric evaluation of hyperlinked news content*

Participants: Rémi Bois, Guillaume Gravier, Pascale Sébillot, Arnaud Touboulic.

Work in collaboration with Éric Jamet, Martin Ragot and Maxime Robert, CRPCC, Rennes.

Following our study of professional user needs in multimedia news analytics [15], we developed a prototype news analytics interface that facilitates the exploration of collections of multimedia documents by journalists. The application, based on standard web technology, enriches classical functionalities for this type of applications (e.g., keyword highlights, named entity detection, keyword search, etc.) with navigation-based functionalities. The latter exploit a graph-based organization of the collection, established from content-based similarity graphs on which community detection is performed along with basic link characterization. We performed usage tests on students in journalism and on journalists where each user was asked to write a synthesis article on a given topic. Preliminary results indicate that the graph-based navigation improves the completeness of the synthesis by exposing users to more content than with a standard search engine.

7.5. Miscellaneous

In parallel with mainstream research activities, LINKMEDIA has a number of contributions in other domains based on the expertise of the team members.

7.5.1. *Bidirectional GRUs in spoken dialog*

Participants: Christian Raymond, Vedran Vukotić.

Recurrent neural networks recently became a very popular choice for spoken language understanding (SLU) problems. They however represent a big family of different architectures that can furthermore be combined to form more complex neural networks. In [30], we compare different recurrent networks, such as simple recurrent neural networks, long short-term memory networks, gated memory units and their bidirectional versions, on the popular ATIS dataset and on MEDIA, a more complex French dataset. Additionally, we propose a novel method where information about the presence of relevant word classes in the dialog history is combined with a bidirectional gated recurrent unit (GRU).

7.5.2. *Kernel principal components analysis with extreme learning machines*

Participant: Christian Raymond.

Work in collaboration with M'Sila University, Algeria.

Nowadays, wind power and precise forecasting are of great importance for the development of modern electrical grids. In [26], we investigate a prediction system for time series based on kernel principal component analysis (KPCA) and extreme learning machine (ELM). Comparison with standard dimensionality reduction techniques show that the reduction of the original input space affects positively the prediction output.

7.5.3. *Pronunciation adaptation for spontaneous speech synthesis*

Work in collaboration with Gwénoél Lecorvé and Damien Lolive, IRISA, Rennes.

In [36], we present a new pronunciation adaptation method which adapts canonical pronunciations to a spontaneous style. This is a key task in text-to-speech as those pronunciation variants bring expressiveness to synthetic speech, thus enabling new potential applications. The strength of the method is to solely rely on linguistic features and to consider a probabilistic machine learning framework, namely conditional random fields, to produce the adapted pronunciations.

7.6. Participation in benchmarking initiatives

- Video hyperlinking, TRECVID
- Search and anchoring, Mediaeval Multimedia International Benchmark
- Multimodal person discovery in broadcast TV, Mediaeval Multimedia International Benchmark
- DeFT 2015 text-mining challenge

8. Bilateral Contracts and Grants with Industry

8.1. Bilateral Contracts with Industry

Teddy Furon spent 20 % of his time during 6 months to transfer research result to IRT B-com

Ph. D. contract with Alcatel-Lucent Bell Labs (Raghavendran Balu) in the framework of the joint Inria-Alcatel Lucent lab.

9. Partnerships and Cooperations

9.1. Regional Initiatives

9.1.1. *CominLabs Project Linking Media in Acceptable Hypergraphs (LIMAH)*

Participants: Rémi Bois, Vincent Claveau, Guillaume Gravier, Grégoire Jadi, Pascale Sébillot, Arnaud Touboulic.

Duration: 4 years, started in April 2014

Partners: Telecom Bretagne (IODE), Univ. Rennes II (CRPCC, PREFics), Univ. Nantes (LINA/TAL)

URL: <http://limah.irisa.fr>

LIMAH aims at exploring hypergraph structures for multimedia collections, instantiating actual links reflecting particular content-based proximity—similar content, thematic proximity, opinion expressed, answer to a question, etc. Exploiting and developing further techniques targeting pairwise comparison of multimedia contents from an NLP perspective, LIMAH addresses two key issues: How to automatically build from a collection of documents an hypergraph, i.e., a graph combining edges of different natures, which provides exploitable links in selected use cases? How collections with explicit links modify usage of multimedia data in all aspects, from a technology point of view as well as from a user point of view? LIMAH studies hypergraph authoring and acceptability taking a multidisciplinary approach mixing ICT, law, information and communication science as well as cognitive and ergonomics psychology.

9.1.2. *CominLabs Project BigCLIN*

Participants: Vincent Claveau, Ewa Kijak, Clément Dalloux.

Duration: 3 years, started in September 2016

Partners: STL-CNRS, Inserm/CHU Rennes, Inria Cidre URL: <http://www.bigclin.cominlabs.ueb.eu>

Data collected or produced during clinical care process can be exploited at different levels and across different domains. Yet, a well-known challenge for secondary use of health big data is that much of detailed patient information is embedded in narrative text, mostly stored as unstructured data. The project proposes to address the essential needs when reusing unstructured clinical data at a large scale. We propose to develop new clinical records representation relying on fine-grained semantic annotation thanks to new NLP tools dedicated to French clinical narratives. To efficiently map this added semantic information to existing structured data for further analysis at big scale, the project also addresses distributed systems issues: scalability, management of uncertain data and privacy, stream processing at runtime, etc.

9.2. National Initiatives

9.2.1. *ANR Project IDFRAud*

Participant: Teddy Furon.

Duration: 3 years, started in Feb. 2015

Partners: AriadNext, IRCGN, École Nationale Supérieure de Police

The IDFRAud project consists in proposing an automatic solution for ID analysis and integrity verification. Our ID analysis goes through three processes: classification, text extraction and ID verification. The three processes rely on a set of rules that are externalized in formal manner in order to allow easy management and evolving capabilities. This leads us to the ID knowledge management module. Finally, IDFRAud addresses the forensic link detection problem and to propose an automatic analysis engine that can be continuously applied on the detected fraud ID database. Cluster analysis methods are used to discover relations between false IDs in their multidimensional feature space. This pattern extraction module will be coupled with a suitable visualization mechanism in order to facilitate the comprehension and the analysis of extracted groups of inter-linked fraud cases.

9.2.2. *FUI 19 NexGenTV*

Participants: Vincent Claveau, Guillaume Gravier, Ewa Kijak, Gabriel Sargent, Ronan Sicre.

Duration: 2.5 years, started in May 2015

Partners: Eurecom, Avisto Telecom, Wildmoka, Envivio-Ericsson

Television is undergoing a revolution, moving from the TV screen to multiple screens. Today's user watches TV and, at the same time, browses the web on a tablet, sends SMS, posts comments on social networks, searches for complementary information on the program, etc. Facing this situation, NexGen-TV aims at developing a generic solution for the enrichment, the linking and the retrieval of video content targeting the cost-cutting edition of second screen and multiscreen applications for broadcast TV. The main outcome of the project will be a software platform to aggregate and distribute video content via a second-screen edition interface connected to social media. The curation interface will primarily make use of multimedia and social media content segmentation, description, linking and retrieval. Multiscreen applications will be developed on various domains, e.g., sports, news.

9.3. European Initiatives

9.3.1. Collaborations with Major European Organizations

Big Data Value Association (BDVA)

LINKMEDIA is a co-founder and co-leader of the media group (TF7) within BDVA

9.4. International Initiatives

9.4.1. Inria Associate Teams Not Involved in an Inria International Labs

9.4.1.1. MOTIF

Title: Unsupervised motif discovery in multimedia content

International Partner (Institution - Laboratory - Researcher):

Pontifícia Universidade Católica de Minas Gerais (Brazil) - Audio-Visual Information Processing Laboratory (VIPLAB) - Silvio Jamil Guimarães

Universidade Federal Minas Gerais, Brasil - NPDI - Arnaldo Albuquerque de Araújo

Duration: 2014–2016

MOTIF aims at studying various approaches to unsupervised motif discovery in multimedia sequences, i.e., to the discovery of repeated sequences with no prior knowledge on the sequences. On the one hand, we will develop symbolic approaches inspired from work on bioinformatics to motif discovery in the multimedia context, investigating symbolic representations of multimedia data and adaptation of existing symbolic motif discovery algorithms. On the other hand, we will further develop cross modal clustering approaches to repeated sequence discovery in video data, building upon previous work.

9.4.2. Inria International Partners

9.4.2.1. Informal International Partners

- National Institute for Informatics, Japan
- University of Amsterdam, The Netherlands
- Czech Technical University, Czech Republic
- Katholieke Universiteit Leuven, Belgium

9.4.3. Participation in Other International Programs

- PICS CNRS MM-Analytics
 - Title: Fouille, visualisation et exploration multidimensionnelle de contenus multimédia ; Multi-Dimensional Multimedia Browsing, Mining, Analytics (num 6382).
 - International Partner (Institution - Laboratory - Researcher):
 - Reykjavík University, Iceland - Björn Þór Jónsson
 - Jan. 2014 – Dec. 2016

- CNRS – CONFAP FIGTEM
 - Title: Fine-grained text-mining for clinical trials
 - International Partner (Institution - Laboratory - Researcher): Pontifícia Universidade Católica do Paraná - Health Informatics dept, Claudia Moro

FIGTEM aims at developing natural language processing methods, including information extraction and indexing, dedicated to the clinical trial domain. The goal is to populate a formal representation of patients (via their electronic patient records) and clinical trial data in different languages (French, English, Portuguese).
 - Jan. 2016 – Dec. 2018

9.5. International Research Visitors

9.5.1. Visits of International Scientists

Claudia Moro, Lucas Oliveira

Date: Oct. 2016 (1 week)

Institution: Pontifícia Universidade Católica do Paraná - Health Informatics dept

Giorgos Toliás

Date: Sept. 2016 (1 week)

Institution: Czech Technical University, Czech Republic

9.5.1.1. Internships

Gabriel B. de Fonseca

Date: Nov. 2016 - Jan. 2017

Institution: PUC Minas, Brazil

9.5.2. Visits to International Teams

Vincent Claveau

Date: 7-17 December 2016

Institution: Health Informatics dept, Pontifícia Universidade Católica do Paraná, Curitiba, Brazil

Vincent Claveau

Date: 7-13 May 2016

Institution: OLST, Univ. of Montreal, Canada

Guillaume Gravier, Simon Malinowski

Date: Jul. 2016 (1 week)

Institution: PUC Minas, Brazil

Ahmet Iscen

Date: Apr. 2016 - May 2016

Institution: Czech Technical University, Czech Republic

Vedran Vukotić

Date: Sep. 2016 - Dec. 2016

Institution: TU Delft, The Netherlands

10. Dissemination

10.1. Promoting Scientific Activities

10.1.1. Scientific Events Organisation

10.1.1.1. General Chair, Scientific Chair

Laurent Amsaleg was the technical program chair of the Intl. Conf. on Similarity Search and Applications 2016.

Laurent Amsaleg and Vincent Claveau, with X. Tannier (LIMSI-CNRS), D. Nouvel (Inalco) organized a workshop about Computational Journalism. More than 50 persons (computer scientists, sociologists, and news/media compan fellows) attended. Program available at <http://compjournalism2016.irisa.fr>.

Pascale Sébillot is a member of the permanent steering committee of Conf. Francophone en Traitement Automatique des Langues Naturelles.

10.1.1.2. Member of the Organizing Committees

Simon Malinowski was in the organizing committee of the Workshop on Advanced Analysis and Learning on Temporal Data (AALTD 16) held in Riva del Garda, colocated with ECML/PKDD 2016.

Guillaume Gravier was co-chair of a special session at The Speaker and Language Recognition Workshop Odyssey 2016.

10.1.2. Scientific Events Selection

10.1.2.1. Chair of Conference Program Committees

Vincent Claveau was area chair of the French-speaking conference Traitement Automatique du Langage Naturel 2016.

Guillaume Gravier was technical program chair of the International Workshop on Content-Based Multimedia Indexing 2016.

Guillaume Gravier was area chair of the ACM Intl. Conf. on Multimedia 2016.

Guillaume Gravier was technical program chair of the French-speaking international conference Journées d'Études sur la Parole 2016.

Pascale Sébillot was area chair of the French-speaking conference Traitement Automatique du Langage Naturel 2016.

10.1.2.2. Member of Conference Program Committees

Laurent Amsaleg was a PC member of: Intl. Conf. on Similarity Search and Applications; ACM Intl. Conf. on Multimedia Retrieval; IEEE Intl. Conf. on Multimedia and Exhibition.

Yannis Avrithis was a PC member of: IEEE Conference on Computer Vision and Pattern Recognition; European Conference on Computer Vision; British Machine Vision Conference; ACM Multimedia Conference on Multimedia; European Signal Processing Conference; Intl. Conference on Multimedia and Expo; IEEE Intl. Conference in Image Processing; IEEE Intl. Conference on Acoustics, Speech, and Signal Processing; Intl. Workshop on Content-Based Multimedia Indexing.

Vincent Claveau was a PC member of: ACL demonstration track; Workshop on Content-Based Multimedia Indexing; Conf. Traitement Automatique du Langage Naturel; CORIA; Language Resources and Evaluation Conf.

Teddy Furon was a PC member of ACM Information Hiding and Multimedia Security.

Guillaume Gravier was a PC member of: ACM Intl. Conf. on Multimedia; IEEE Intl. Conf. on Multimedia and Exhibition; Annual Conf. of the Intl. Speech Communication Association; IEEE Intl. Workshop on Multimedia Signal Processing; European Conf. on Information Retrieval; Conf. on Knowledge and Information Management; IEEE Intl. Conf. on Acoustics, Speech and Signal Processing; Intl. Conf. on Statistical Language and Speech Processing; Language Resources and Evaluation Conf.

Ewa Kijak was a PC member of Intl. Workshop on Content-Based Multimedia Indexing.

Pascale Sébillot was a PC member of: ACM International Conference on Information and Knowledge Management; European Conference on Information Retrieval; Conférence nationale en intelligence artificielle; Conf. Traitement Automatique des Langues Naturelles; Colloque international sur l'écrit et le document

10.1.2.3. Reviewer

Ewa Kijak reviewed for ACM Conf. on Multimedia.

Pascale Sébillot was a scientific committee member of: Language Resources and Evaluation Conference

10.1.3. Journal

10.1.3.1. Member of the Editorial Boards

Vincent Claveau is member of the editorial board of the journal *Traitement Automatique des Langues*.

Vincent Claveau was guest editor of the special issue on Information Retrieval and Natural Language Processing of the journal *Traitement Automatique des Langues* (published in 2016).

Guillaume Gravier was appointed associate editor of *IEEE Trans. on Multimedia* in Sep. 2016.

Guillaume Gravier is guest editor of the special issue of *Multimedia Tools and Applications* on content-based multimedia indexing (to appear in 2017).

Guillaume Gravier was editor of the Working Notes Proc. of the MediaEval Multimedia Benchmark.

Pascale Sébillot is editor of the Journal *Traitement Automatique des Langues*.

Pascale Sébillot is member of the editorial board of the Journal *Traitement Automatique des Langues*.

10.1.3.2. Reviewer - Reviewing Activities

Laurent Amsaleg was a reviewer for: *IEEE Trans. on Circuits and Systems for Video Technology*, *IEEE Trans. on Multimedia*, *Pattern Recognition Letters*, *IEEE Trans. on Information Forensics & Security*, *IEEE Trans. on Image Processing*.

Yannis Avrithis was a reviewer for *Multimedia Tools and Applications*, Elsevier *Neurocomputing*.

Vincent Claveau reviewed for *Multimedia Tools and Applications*, Elsevier *Knowledge-Based Systems*, Elsevier *Information Sciences*, *Traitement Automatique des Langues*.

Teddy Furon is a reviewer for *IEEE Trans. on Information Forensics and Security*, Elsevier *Digital Signal Processing*, *EURASIP Journal of Information Security*, *IEEE Trans. on Image Processing*, *IEEE Trans. on Signal Processing*.

Guillaume Gravier reviewed for *Computer Speech and Language*.

Pascale Sébillot was member of the reading committee for several issues of the Journal *Traitement Automatique des Langues*.

10.1.4. Invited Talks

Yannis Avrithis gave an invited talk at Pattern Recognition and Computer Vision Colloquium on March 31 in Prague, <http://cmp.felk.cvut.cz/cmp/events/colloquium-2016.03.31>.

Vincent Claveau gave an invited talk at Congrès de l'ACFAS, colloque Documents et ressources pour leur traitement : un couplage crucial, Montreal.

Vincent Claveau gave an invited talk at Séminaire D2K du labex Digicosme, Paris.

Vincent Claveau gave an invited talk at Séminaire commun RALI-OLST, Montréal.

Guillaume Gravier gave an invited talk at the ACM Multimedia 2016 TPC workshop.

Guillaume Gravier was invited as a panelist at the Intl. Workshop on Content-Based Multimedia Indexing.

Guillaume Gravier gave an invited talk at the Journées de Statistique de Rennes (JSTAR).

Guillaume Gravier animated a round-table on 'Jeux de données et données en jeu : l'accès des données à des fins de recherche' at the Rencontres du Numérique organized by the National Research Agency (ANR).

Simon Malinowski gave an invited talk on time series classification at PUC Minas, Brazil.

Pascale Sébillot gave an invited tutorial (8h) about NLP at Institut de l'information scientifique et technique (INiST) Nancy.

10.1.5. Leadership within the Scientific Community

Laurent Amsaleg was nominated to be part of the Steering Committee of SISAP for a 2016–2020 term.

Vincent Claveau is finance head of the Association pour la Recherche d'Informations et ses Applications (ARIA).

Guillaume Gravier is president of the Association Francophone de la Communication Parlée (AFCP), French-speaking branch of the Intl. Speech Communication Association.

Guillaume Gravier is co-founder and general chair of the ISCA SIG Speech, Language and Audio in Multimedia.

Guillaume Gravier is member of the Community Council of the Mediaeval Multimedia Evaluation series.

Guillaume Gravier is the technical representative of Inria in the cPPP Big Data Value Association, actively working on technical aspects of data analytics.

10.1.6. Scientific Expertise

Laurent Amsaleg was an Expert Evaluator for the Luxembourg National Research Fund, 2016.

Vincent Claveau was an expert for the FNRS (Belgian funding agency).

Vincent Claveau was an expert for GIFAS - Groupement des Industries Françaises Aéronautiques et Spatiales.

Vincent Claveau was an expert for the Programme Hubert Curien (funding program for bilateral projects).

Teddy Furon is scientific adviser for the company LAMARK.

Guillaume Gravier was vice-president of the Scientific Evaluation Committee of the National Research Agency for the theme 'Interaction, Robotic, Content, Automatic' up to July 2016.

Since September 2016, Guillaume Gravier is president of the Scientific Evaluation Committee of the National Research Agency for the theme 'Knowledge, data, content, big data, HPC, simulation'.

10.1.7. Research Administration

Vincent Claveau is deputy head of the GdR MaDICS, a CNRS inter-lab initiative to promote research about Big Data and Data Science.

Guillaume Gravier is a member of the Board of the technology cluster Images & Réseaux.

Guillaume Gravier is a member of the Board of the Comité des Projets of Inria - Rennes Bretagne Atlantique.

Pascale Sébillot is a member of the Conseil National des Universités 27th section (computer science).

Pascale Sébillot is the director of the Computer Science Laboratory, INSA Rennes.

Pascale Sébillot is the deputy director of the Scientific Advisory Committee of IRISA UMR 6074.

Pascale Sébillot is a member of the theses advisory committee of the Matisse doctoral school.

10.2. Teaching - Supervision - Juries

10.2.1. Teaching

For researchers, all activities are given. For professors and assistant professors, only courses at the M. Sc. level are listed.

Licence: Teddy Furon, Probabilities, 40h, L1, Agrocampus Rennes, France

Licence: Guillaume Gravier, Databases, 30h, L2, INSA Rennes, France

Licence: Guillaume Gravier, Probability and statistics, 16h, L3, INSA Rennes, France

Licence: Guillaume Gravier, Natural Language Processing, 12h, L3, INSA Rennes, France

Master: Laurent Amsaleg, Multidimensional indexing, 13h, M2, University Rennes 1, France

Master: Vincent Claveau, Information retrieval, 7h, M2, ENSAI, Rennes, France

Master: Vincent Claveau, Data-Based Knowledge Acquisition: Symbolic Methods, 20h, M1, INSA de Rennes, France

Master: Vincent Claveau, Text Mining, 36h, M2, Univ. Rennes 1, France

Master: Vincent Claveau, Machine Learning for symbolic and sequential data, 7h, M2, Univ. Rennes 1, France

Master: Vincent Claveau, Information Retrieval, 15h, M2, ENSSAT, France

Master: Vincent Claveau, Information Retrieval, 13h, M2, Univ. Rennes 1, France

Master: Teddy Furon, Rare events, 20h, M2, Insa Rennes, France

Master: Guillaume Gravier, Data analysis and probabilistic modeling, 30h, M2, University Rennes 1, France

Master: Ewa Kijak, Image processing, 64h, M1, ESIR, France

Master: Ewa Kijak, Supervised learning, 15h, M2R, University Rennes 1, France

Master: Ewa Kijak, Supervised learning, 45h, M1, ESIR, France

Master: Ewa Kijak, Statistical data mining, 14h, M2, University Rennes 1, France

Master: Ewa Kijak, Indexing and multimedia databases, 15h, M2, ENSSAT, France

Master: Ewa Kijak, Computer vision, 15h, M2, ESIR, France

Master: Simon Malinowski, Short-term time series prediction, 29h, M1, Univ. Rennes 1

Master: Simon Malinowski, Supervised Learning, 10h, M2, Univ. Rennes 1

Master: Pascale Sébillot, Data-Based Knowledge Acquisition: Symbolic Methods, 14h, M1, INSA Rennes, France

Master: Pascale Sébillot, From multimedia documents structuring to multimedia collections structuring, 2h, M2 Language Industry, Université Grenoble-Alpes, France

Master: Pascale Sébillot, Advanced Databases and Modern Information Systems, 70h, M2, INSA Rennes, France

Master: Pascale Sébillot, Logic Programming, 12h, M1, INSA Rennes, France

10.2.2. Supervision

PhD: Raghavendran Balu, Privacy-preserving data aggregation and service personalization using highly-scalable data indexing techniques, defended Nov. 2016, Teddy Furon and Laurent Amsaleg

PhD: Petra Bosilj, Content based image indexing and retrieval using hierarchical image representations, defended Jan. 2016, Ewa Kijak and Sebastien Lefèvre (with OBELIX, IRISA team)

PhD: Thomas Gaillat, Reference in interlanguage: The case of this and that, defended Jun. 2016, Nicolas Ballier and Pascale Sébillot (with Université Paris Diderot)

PhD in progress: Rémi Bois, Navigable directed multimedia hypergraphs: construction and exploitation, started October 2014, Guillaume Gravier and Pascale Sébillot

PhD in progress: Ricardo Carlini Sperandio, Unsupervised motif mining in multimedia time series, started August 2015, Laurent Amsaleg and Guillaume Gravier

PhD in progress: Ahmet Iscen, Continuous memories for representing sets of vectors and image collections, started September 2014, Teddy Furon

PhD in progress: Grégoire Jadi, Opinion mining in multimedia data, started October 2014, Vincent Claveau, Béatrice Daille (LINA, Nantes) and Laura Monceaux (LINA, Nantes)

PhD in progress: Raheel Kareem Qader, Phonology modeling for emotional speech synthesis, started January 2014, Gwénolé Lecorvé and Pascale Sébillot (with EXPRESSION, IRISA Team)

PhD in progress: Mathieu Laroze, Active learning on adaptive representations for object detection in high-resolution imaging, started June 2016, Romain Dambreville, Chloe Friguet, Ewa Kijak and Sebastien Lefevre (with OBELIX, IRISA team)

PhD in progress: Cédric Maigrot, Detecting fake information on social networks, started October 2015, Laurent Amsaleg, Vincent Claveau and Ewa Kijak

PhD in progress: Vedran Vukotič, Deep neural architectures for automatic representation learning from multimedia multimodal data, started October 2014, Guillaume Gravier and Christian Raymond

PhD in progress: Oriane Simeoni, Invariance and supervision in visual learning, started Oct. 2016, Yannis Avrithis and Guillaume Gravier

PhD in progress: Mikail Demirdelen, User-adapted multi-document multimedia synthesis, started Oct. 2016, Guillaume Gravier and Pascale Sébillot

PhD in progress: Clément Dalloux, Clinical Text Mining and Indexing, started Dec. 2016, Olivier Dameron and Vincent Claveau (with DYLISS projet-team)

10.2.3. Juries

Vincent Claveau

PhD, reviewer, Gabriel Bernier-Colborne, Université de Montréal, Canada

Teddy Furon

PhD, reviewer, Thijs Laarhoven, Eindhoven University

PhD, jury member, Julien Lolive, Télécom-Bretagne Brest

Guillaume Gravier

PhD, reviewer and president, Xavier Bost, Université d'Avignon et des Pays du Vaucluse

Ewa Kijak

PhD, Hassan Wehbe, Université Toulouse III

Pascale Sébillot

PhD, reviewer, Jean-Philippe Fauconnier Biard, Université Toulouse 3

PhD, member, Abdessalam Boucekif, Université du Maine

PhD, member, Sana Ben Nasr, Université de Rennes 1

HDR, member, Olivier Ferret, Université Paris Sud 11

10.3. Popularization

Vincent Claveau

An article entitled 'Vélos, bicyclettes et moteur de recherche', was published by ACFAS on its website

An article entitled 'Détecter l'intox sur Twitter' was published in a special issue of the Dossiers de l'Université de Rennes 1

11. Bibliography

Publications of the year

Doctoral Dissertations and Habilitation Theses

- [1] R. BALU. *Privacy-aware and scalable recommender systems using sketching techniques*, Université Rennes 1, November 2016, <https://tel.archives-ouvertes.fr/tel-01430156>

Articles in International Peer-Reviewed Journals

- [2] P. BOSILJ, E. APTOULA, S. LEFÈVRE, E. KIJAK. *Retrieval of Remote Sensing Images with Pattern Spectra Descriptors*, in "ISPRS International Journal of Geo-Information", 2016 [DOI : 10.3390/IJGI5120228], <https://hal.archives-ouvertes.fr/hal-01397883>
- [3] P. BOSILJ, M. H. WILKINSON, E. KIJAK, S. LEFÈVRE. *Local 2D pattern spectra as connected region descriptors*, in "Mathematical Morphology - Theory and Applications", 2016 [DOI : 10.1515/MATHM-2016-0011], <https://hal.archives-ouvertes.fr/hal-01320009>
- [4] M. JAIN, H. JÉGOU, P. BOUTHEMY. *Improved Motion Description for Action Classification*, in "Frontiers in information and communication technologies", January 2016 [DOI : 10.3389/FICT.2015.00028], <https://hal.inria.fr/hal-01401833>
- [5] G. SARGENT, F. BIMBOT, E. VINCENT. *Estimating the structural segmentation of popular music pieces under regularity constraints*, in "IEEE/ACM Transactions on Audio, Speech, and Language Processing", 2017, <https://hal.inria.fr/hal-01403210>
- [6] L. WENG, L. AMSALEG, T. FURON. *Privacy-Preserving Outsourced Media Search*, in "IEEE Transactions on Knowledge and Data Engineering", July 2016, vol. 28, n^o 10 [DOI : 10.1109/TKDE.2016.2587258], <https://hal.inria.fr/hal-01391444>
- [7] C. E. DOS SANTOS, E. E. KIJAK, G. GRAVIER, W. ROBSON SCHWARTZ. *Partial Least Squares for Face Hashing*, in "Neurocomputing", 2016, vol. 213, pp. 34-47 [DOI : 10.1016/J.NEUCOM.2016.02.083], <https://hal.archives-ouvertes.fr/hal-01399660>

Articles in National Peer-Reviewed Journals

- [8] V. CLAVEAU, J.-Y. NIE. *Recherche d'information et traitement automatique des langues : collaboration, synergie et convergence : Introduction au numéro spécial sur la recherche d'information et le traitement automatique des langues*, in "Traitement Automatique des Langues", August 2016, vol. 56, n^o 3, <https://hal.archives-ouvertes.fr/hal-01394789>

Articles in Non Peer-Reviewed Journals

- [9] I. BOGDAN, H. MÜLLER, G. GRAVIER, Y. KOMPATSIARIS. *2016 14th International Workshop on Content-Based Multimedia Indexing: IEEE Multimedia Conference Report*, in "IEEE MultiMedia", 2016, <https://hal.archives-ouvertes.fr/hal-01399468>

International Conferences with Proceedings

- [10] R. BALU, T. FURON, L. AMSALEG. *Sketching techniques for very large matrix factorization*, in "ECIR 2016 - 38th European Conference on Information Retrieval", Padoue, Italy, Proceedings of the European conference on Information Retrieval, 2016, <https://hal.inria.fr/hal-01249621>
- [11] R. BALU, T. FURON. *Differentially Private Matrix Factorization using Sketching Techniques*, in "ACM Workshop on Information Hiding and Multimedia Security", Vigo, Spain, June 2016 [DOI : 10.1145/2909827.2930793], <https://hal.archives-ouvertes.fr/hal-01317596>

- [12] R. BOIS, V. VUKOTIĆ, R. SICRE, C. RAYMOND, G. GRAVIER, P. SÉBILLOT. *IRISA at TRECVID2016: Crossmodality, Multimodality and Monomodality for Video Hyperlinking*, in "Working Notes of the TRECVID 2016 Workshop", Gaithersburg, United States, 2016, <https://hal.archives-ouvertes.fr/hal-01400275>
- [13] V. CLAVEAU, E. KIJAK. *Direct vs. indirect evaluation of distributional thesauri*, in "International Conference on Computational Linguistics, COLING", Osaka, Japan, December 2016, <https://hal.archives-ouvertes.fr/hal-01394739>
- [14] V. CLAVEAU, E. KIJAK. *Distributional Thesauri for Information Retrieval and vice versa*, in "Language and Resource Conference, LREC", Portoroz, Slovenia, May 2016, <https://hal.archives-ouvertes.fr/hal-01394770>
- [15] G. GRAVIER, M. RAGOT, L. AMSALEG, R. BOIS, G. JADI, E. JAMET, L. MONCEAUX, P. SÉBILLOT. *Shaping-Up Multimedia Analytics: Needs and Expectations of Media Professionals*, in "22nd International Conference on Multimedia Modelling, Special Session Perspectives on Multimedia Analytics", Miami, United States, January 2016, <https://hal.inria.fr/hal-01214829>
- [16] A. ISCEN, L. AMSALEG, T. FURON. *Scaling Group Testing Similarity Search*, in "ACM International Conference on Multimedia Retrieval 2016", New York, United States, June 2016 [DOI : 10.1145/2911996.2912010], <https://hal.inria.fr/hal-01304287>
- [17] A. ISCEN, T. FURON. *Group Testing for Identification with Privacy*, in "ACM Workshop on Information Hiding and Multimedia Security", Vigo, Spain, June 2016, <https://hal.inria.fr/hal-01309032>
- [18] A. ISCEN, M. RABBAT, T. FURON. *Efficient Large-Scale Similarity Search Using Matrix Factorization*, in "2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)", Las Vegas, United States, June 2016, <https://hal.inria.fr/hal-01294736>
- [19] G. JADI, V. CLAVEAU, B. DAILLE, L. MONCEAUX-CACHARD. *Evaluating Lexical Similarity to build Sentiment Similarity*, in "Language and Resource Conference, LREC", portoroz, Slovenia, May 2016, <https://hal.archives-ouvertes.fr/hal-01394768>
- [20] G. JADI, L. MONCEAUX, V. CLAVEAU, B. DAILLE. *Opinion Target Expression extraction : from English to French*, in "Traitement Automatique des Langues Naturelles, JEP-TALN-RECITAL", Paris, France, July 2016, <https://hal.archives-ouvertes.fr/hal-01397188>
- [21] B. Þ. JÓNSSON, L. AMSALEG, H. LEJSEK. *SSD Technology Enables Dynamic Maintenance of Persistent High-Dimensional Indexes*, in "ACM International Conference on Multimedia Retrieval 2016", New York, United States, June 2016 [DOI : 10.1145/2911996.2912065], <https://hal.inria.fr/hal-01305651>
- [22] B. Þ. JÓNSSON, M. WORRING, J. ZAHÁLKA, S. RUDINAC, L. AMSALEG. *Ten Research Questions for Scalable Multimedia Analytics*, in "22nd International Conference on Multimedia Modelling, Special Session Perspectives on Multimedia Analytics", Miami, United States, January 2016, <https://hal.inria.fr/hal-01214847>
- [23] A. LE GUENNEC, S. MALINOWSKI, R. TAVENARD. *Data Augmentation for Time Series Classification using Convolutional Neural Networks*, in "ECML/PKDD Workshop on Advanced Analytics and Learning on Temporal Data", Riva Del Garda, Italy, September 2016, <https://halshs.archives-ouvertes.fr/halshs-01357973>

- [24] C. MAIGROT, V. CLAVEAU, E. KIJAK, R. SICRE. *MediaEval 2016: A multimodal system for the Verifying Multimedia Use task*, in "MediaEval 2016: "Verfying Multimedia Use" task", Hilversum, Netherlands, October 2016 [DOI : 10.1145/1235], <https://hal.archives-ouvertes.fr/hal-01394785>
- [25] C. MAIGROT, E. KIJAK, V. CLAVEAU. *Médias traditionnels, médias sociaux : caractériser la réinformation*, in "TALN 2016 - 23ème Conférence sur le Traitement Automatique des Langues Naturelles", Paris, France, July 2016, <https://hal.inria.fr/hal-01349871>
- [26] H. MEZAACHE, H. BOUZGOU, C. RAYMOND. *Kernel Principal Components Analysis with Extreme Learning Machines for Wind Speed Prediction*, in "Seventh International Renewable Energy Congress, IREC2016", Hammamet, Tunisia, March 2016, <https://hal.inria.fr/hal-01394000>
- [27] A. ROYER, G. GRAVIER, V. CLAVEAU. *Audio word similarity for clustering with zero resources based on iterative HMM classification*, in "International Conference on Acoustics, Speech and Signal Processing, ICASSP", Shanghai, China, March 2016, pp. 5340 - 5344 [DOI : 10.1109/ICASSP.2016.7472697], <https://hal.archives-ouvertes.fr/hal-01394757>
- [28] G. SARGENT, G. BARBOSA DE FONSECA, I. LYON FREIRE, R. SICRE, Z. K. G. DO PATROCÍNIO JR., S. J. GUIMARÃES, G. GRAVIER. *PUC Minas and IRISA at Multimodal Person Discovery*, in "Working Notes Proceedings of the MediaEval Workshop", Hilversum, Netherlands, CEUR-WS.org, October 2016, vol. 1739, <https://hal.archives-ouvertes.fr/hal-01400261>
- [29] R. TAVENARD, S. MALINOWSKI. *Cost-Aware Early Classification of Time Series*, in "European Conference on Machine Learning and Principles and Practice of Knowledge Discovery", Riva del Garda, Italy, September 2016, <https://halshs.archives-ouvertes.fr/halshs-01339007>
- [30] V. VUKOTIC, C. RAYMOND, G. GRAVIER. *A step beyond local observations with a dialog aware bidirectional GRU network for Spoken Language Understanding*, in "Interspeech", San Francisco, United States, September 2016, <https://hal.inria.fr/hal-01351733>
- [31] V. VUKOTIC, C. RAYMOND, G. GRAVIER. *Bidirectional Joint Representation Learning with Symmetrical Deep Neural Networks for Multimodal and Crossmodal Applications*, in "ICMR", New York, United States, ACM, June 2016, <https://hal.inria.fr/hal-01314302>
- [32] V. VUKOTIĆ, C. RAYMOND, G. GRAVIER. *Multimodal and Crossmodal Representation Learning from Textual and Visual Features with Bidirectional Deep Neural Networks for Video Hyperlinking*, in "ACM Multimedia 2016 Workshop: Vision and Language Integration Meets Multimedia Fusion (iV&L-MM'16)", Amsterdam, Netherlands, ACM Multimedia, October 2016, <https://hal.inria.fr/hal-01374727>
- [33] H. B. DA SILVA, Z. PATROCINO JR., G. GRAVIER, L. AMSALEG, A. DE A. ARAÚJO, S. J. F. GUIMARÃES. *Near-Duplicate Video Detection Based on an Approximate Similarity Self-Join Strategy*, in "14th International Workshop on Content-based Multimedia Indexing", bucares, Romania, June 2016, <https://hal.inria.fr/hal-01305691>

National Conferences with Proceedings

- [34] V. CLAVEAU. *Dimensionnalité intrinsèque dans les espaces de représentation des termes et des documents*, in "Conférence en Recherche d'Information et Applications, CORIA", Toulouse, France, March 2016, <https://hal.archives-ouvertes.fr/hal-01394749>

- [35] C. MAIGROT, S. BRINGAY, J. AZÉ. *Concept drift vs suicide: comment l'un peut prévenir l'autre ?*, in "EGC: Extraction et Gestion des Connaissances", Reims, France, January 2016, <https://hal.inria.fr/hal-01283889>
- [36] R. QADER, G. LECORVÉ, D. LOLIVE, P. SÉBILLOT. *Adaptation de la prononciation pour la synthèse de la parole spontanée en utilisant des informations linguistiques*, in "Journées d'Études sur la Parole", Paris, France, July 2016, <https://hal.inria.fr/hal-01321361>
- [37] A. SIMON, G. GRAVIER, P. SÉBILLOT. *Évaluation d'une nouvelle structuration thématique hiérarchique des textes dans un cadre de résumé automatique et de détection d'ancres au sein de vidéos*, in "Conférence sur le Traitement Automatique des Langues Naturelles", Paris, France, 2016, pp. 139-152, <https://hal.archives-ouvertes.fr/hal-01399670>

Scientific Books (or Scientific Book chapters)

- [38] A. BAILLY, S. MALINOWSKI, R. TAVENARD, L. CHAPEL, T. GUYET. *Dense Bag-of-Temporal-SIFT-Words for Time Series Classification*, in "Advanced Analysis and Learning on Temporal Data", Springer, 2016, <https://hal.archives-ouvertes.fr/hal-01252726>
- [39] P. BAS, T. FURON, F. CAYRE, G. DOËRR, B. MATHON, W.-S. GAN, C.-C. J. KUO, T. F. ZHENG, M. BARNI (editors) *Watermarking Security: Fundamentals, Secure Designs and Attacks*, Springer Briefs, Springer, January 2016, <https://hal.archives-ouvertes.fr/hal-01283066>
- [40] T. FURON. *Watermarking security*, in "Information hiding", Information security and privacy series, Artech House, 2016, <https://hal.archives-ouvertes.fr/hal-01275880>

Books or Proceedings Editing

- [41] L. AMSALEG, G. Þ. GUÐMUNDSSON, C. GURRIN, B. Þ. JÓNSSON, S. SATOH (editors). *Proceedings of the 23rd International Conference on Multimedia Modeling, MMM 2017, Reykjavik, Iceland, January 4-6, 2017*, MultiMedia Modeling, Springer International Publishing, Reykjavík, Iceland, January 2017, vol. 10132, n^o 0302-9743 [DOI : 10.1007/978-3-319-51811-4], <https://hal.archives-ouvertes.fr/hal-01430555>
- [42] L. AMSALEG, M. E. HOULE, E. SCHUBERT (editors). *Similarity Search and Applications: 9th International Conference, SISAP 2016, Tokyo, Japan, October 24-26, 2016, Proceedings*, Springer International Publishing, Tokyo, Japan, October 2016, vol. 9939, XVII, 339 p. [DOI : 10.1007/978-3-319-46759-7], <https://hal.inria.fr/hal-01391426>
- [43] V. CLAVEAU, J. NIE (editors). *Recherche d'information et traitement automatique des langues*, Traitement Automatique des Langues, TAL, ATALA, August 2016, vol. 56, n^o 3, <https://hal.archives-ouvertes.fr/hal-01394788>
- [44] G. GRAVIER, D. CLAIRE-HÉLÈNE, H. BREDIN, I. BOGDAN, B. CHRISTINA, E. DELLANDRÉA, C. JAEYONG, M. A. RIEGLER, S. RICHARD, I. SZÖKE, G. J. F. JONES, M. LARSON (editors). *Working Notes Proceedings of the MediaEval Workshop*, 2016, <https://hal.archives-ouvertes.fr/hal-01399674>

Research Reports

- [45] G. Þ. GUÐMUNDSSON, L. AMSALEG, B. T. JÓNSSON, M. J. FRANKLIN. *Towards Engineering a Web-Scale Multimedia Service: A Case Study Using Spark*, Inria Rennes Bretagne Atlantique ; Reykjavik University ; UC Berkeley, December 2016, <https://hal.inria.fr/hal-01416089>

- [46] G. SARGENT, F. BIMBOT, E. VINCENT. *Supplementary material to the article: Estimating the structural segmentation of popular music pieces under regularity constraints*, IRISA-Inria, Campus de Beaulieu, 35042 Rennes cedex ; Inria Nancy, équipe Multispeech, September 2016, <https://hal.inria.fr/hal-01368683>

References in notes

- [47] M. AGOSTI, J. ALLAN. *Special issue on methods and tools for the automatic construction of hypertext*, in "Information Processing and Retrieval", 1997, vol. 33, n^o 2
- [48] C. A. BHATT, M. S. KANKANHALLI. *Multimedia data mining: State of the art and challenges*, in "Multimedia Tools and Applications", 2011, vol. 51, n^o 1, pp. 35–76
- [49] R. BOIS, A. SIMON, R. SICRE, G. GRAVIER, P. SÉBILLOT, E. MORIN. *IRISA at TrecVid2015: Leveraging Multimodal LDA for Video Hyperlinking*, in "Proc. of the TREC Vid Workshop", 2015
- [50] O. CHUM, J. MATAS. *Large-Scale Discovery of Spatially Related Images*, in "Pattern Analysis and Machine Intelligence", 2010, vol. 32, n^o 2, pp. 371-377
- [51] M. ESKEVICH, G. J. JONES, S. CHEN, R. ALY, R. ORDELMAN, D. NADEEM, C. GUINAUDEAU, G. GRAVIER, P. SÉBILLOT, T. DE NIES, P. DEBEVERE, R. V. DE WALLE, P. GALUŠCKOVÁ, P. PECINA, M. LARSON. *Multimedia Information Seeking through Search And Hyperlinking*, in "Intl. Conf. on Multimedia Retrieval", 2013
- [52] C. HERLEY. *ARGOS: Automatically extracting Repeating Objects from multimedia Streams*, in "IEEE Trans. on Multimedia", Feb. 2006, vol. 8, n^o 1, pp. 115–129
- [53] A. PARK, J. R. GLASS. *Unsupervised pattern discovery in speech*, in "IEEE Transaction on Acoustic, Speech and Language Processing", Jan. 2008, vol. 16, n^o 1, pp. 186–197
- [54] K. M. PUA, J. GAUCH, S. E. GUCH, J. Z. MIADOWICZ. *Real time repeated video sequence identification*, in "Computer Vision and Image Understanding", 2004, vol. 93, n^o 2004, pp. 310–327
- [55] A. SIMON, P. SÉBILLOT, G. GRAVIER. *Hierarchical topic structuring: from dense segmentation to topically focused fragments via burst analysis*, in "Recent Advances on Natural Language Processing", Hissar, Bulgaria, 2015, <https://hal.archives-ouvertes.fr/hal-01186443>
- [56] C. G. M. SNOEK, M. WORRING. *Multimodal Video Indexing: A Review of the State-of-the-Art*, in "Multimedia Tools and Applications", 2005, vol. 25, n^o 1, pp. 5–35
- [57] O. DE ROOIJ, M. WORRING. *Browsing Video Along Multiple Threads*, in "IEEE Transactions on Multimedia", 2010, vol. 12, n^o 2, pp. 121–130