Proposal Book [- D2K - CHIST-ERA Eligibility - YES]

Proposal Data

Acronym	CAMOMILE
Full name	Collaborative Annotation of multi-MOdal, multI-Lingual and multi-mEdia documents
Duration	36
Topic	D2K
Keywords	collaborative annotation; multilingual multimodal multimedia corpus ; Person Identification
Ranking	0

Coordinator contact point for the proposal

Name	Claude Barras
Institution	LIMSI/CNRS
Address	LIMSI-CNRS B.P. 133 F-91403 ORSAY cedex
Country	France
Phone	(33) 1 69 85 81 84
Email	Claude.Barras@limsi.fr

Consortium Partners

C/P	Institution	Contact	Other	Country	Legal Status	
C	LIMSI/CNRS	Claude Barras	Hervé Bredin ; Sophie Rosset ; Gilles Adda	France	Public research organisation	**Ind.Eff.:**82 **Ind.Cost:**446,247 **Ind.Bud.:**230,979
P	LIG/CNRS	Georges Quénot	Laurent Besacier;Jean-Pierre Chevallet; Philippe Mulhem; Benjamin Lecouteux	France	Public research organisation	**Ind.Eff.:**106 **Ind.Cost:**500,997 **Ind.Bud.:**249,999
P	IMMI/CNRS	Joseph-Jean Mariani	Martine Garnier-Rizet ; Gil Francopoulo	France	Public research organisation	**Ind.Eff.:**18 **Ind.Cost:**143,211 **Ind.Bud.:**120,605
P	Istanbul Technical University/Faculty of Computer and informatics	Hazim Kemal Ekenel	Cenk Demiroglu ; Zehra Cataltepe	Turkey	Public research organisation	**Ind.Eff.:**108 **Ind.Cost:**213,000 **Ind.Bud.:**213,000
P	Universitat Politècnica de Catalunya	Francisco Javier Hernando Pericàs	Jose Ramon Casas	Spain	Public research organisation	**Ind.Eff.:**81 **Ind.Cost:**195,000 **Ind.Bud.:**195,000
P	Centre De Recherche Public - Gabriel Lippmann	Benoît Otjacques	senior researcher	Luxembourg	Public research organisation	**Ind.Eff.:**36 **Ind.Cost:**308,840 **Ind.Bud.:**250,000

Abstract:

Human activity is constantly generating large volumes of heterogeneous data, in particular via the Web. These data can be collected and explored to gain new insights in social sciences, linguistics, economics, behavioural studies as well as artificial intelligence and computer sciences. In this regard, 3M (multimodal, multimedia, multilingual) data could be seen as a paradigm of sharing an object of study, human data, between many scientific domains. But, to be really useful, these data should be annotated, and available in very large amounts. Annotated data is useful for computer sciences which process human data with statistical-based machine learning methods, but also for social sciences which are more and more using the large corpora available to support new insights, in a way which was not imaginable few years ago. However, annotating data is costly as it involves a large amount of manual work, and in this regard 3M data, for which we need to annotate different modalities with different levels of abstraction is especially costly. Current annotation framework involves some local manual annotation, with the help sometimes of some automatic tools (mainly pre-segmentation). The proposal aims at developing a first prototype of collaborative annotation framework on 3M data, in which the manual annotation will be done remotely on many sites, while the final annotation will be localized on the main site. Furthermore, with the same principle, some systems devoted to automatic processing of the modalities (speech, vision) present in the multimedia data will help the transcription, by producing automatic annotations. These automatic annotations are done remotely in each expertise point, which will be then combined locally to produce a meaningful help to the annotators. In order to develop this new annotation concept, we will test it on a practical case study: the problem of person annotation (who is speaking?, who is seen?) in video, which needs collaboration of high level automatic systems dealing with different media (video, speech, audio tracks, OCR, ...). The quality of the annotated data will be evaluated through the task of person retrieval. This new way to envision the annotation process, should lead to some methodologies, tools, instruments and data that are useful for the whole scientific community who have interest in 3M annotated data; to support this will, all the work will be supervised by a committee which will contain representatives from the main international organizations dealing with language resources and evaluation.

Relevance:

The proposal addresses the problem of creating a new annotation framework of 3M (multimodal, multimedia, multilingual) data with different levels of annotation. In this regard it is compliant with the general subject of D2K topic, as described in the call, to extract new knowledge from large amounts of heterogeneous, unstructured data, in order to elaborate models which, in turn, will be used to process other data: the annotation is essential to improve the modeling used in the automatic systems, but also to acquire new insights from the human and social sciences to whom these annotated data and tools will be made available. The data we want to process consists of videos from different countries and languages, most of them available on the web; this point is also mentioned in the call ("multilingual and multimedia data"). This proposal proposes a new way to envision the multimodal annotation process, and to deliver, at the end of the project, a pilot framework, but also tools and data; the core aspect of this annotation process is a shared format between the different annotation levels, which will be used by different manual and automatic annotation tools and instruments. In this regard, we address the point "6. Generic methodologies, tools and formats to ease the exchange of data and models" of the call, by developing a new methodology, tools, instruments and formats. At a more general level, the proposal clearly wants to challenge the actual annotation framework, in order to enable annotation of very large amounts of 3M data. This in turn will enable significant breakthroughs in the social sciences and computer sciences which share this same object of study. This ambitious project which is multidisciplinary in the sense that it leads to the set up of a common experimental framework for social sciences and computer sciences, is in line with the general CHIST-ERA principles as expressed in the 2011 call.

Proposal Book [- D2K - CHIST-ERA Eligibility - YES]

Proposal Data

Acronym	MUCKE
Full name	Multimedia And User Credibility Knowledge Extraction
Duration	36
Topic	D2K
Keywords	multimedia, large data, social network, user credibility, concept similarity
Ranking	0

Coordinator contact point for the proposal

Name	Allan Hanbury
Institution	Vienna University of Technology, Information And Software Engineering Group
Address	Favoriten Strasse 9-11/188, Vienna, 1040
Country	Austria
Phone	+43158801 188310
Email	hanbury@ifs.tuwien.ac.at

Consortium Partners

C/P	Institution	Contact	Other	Country	Legal Status	
C	Vienna University of Technology, Information And Software Engineering Group	Allan Hanbury	Mihai Lupu	Austria	Public research organisation	**Ind.Eff.:**72 **Ind.Cost:**409,269 **Ind.Bud.:**409,269
P	CEA LIST	Adrian Popescu		France	Public research organisation	**Ind.Eff.:**72 **Ind.Cost:**726,269 **Ind.Bud.:**304,004
P	Universitatea "Al. I. Cuza" Iasi, Facultatea de Informatica	Adrian Iftene	Diana Trandabat	Romania	Public research organisation	**Ind.Eff.:**72 **Ind.Cost:**270,798 **Ind.Bud.:**270,798
P	Bilkent University, Department of Computer Engineering	Pinar Duygulu-Sahin		Turkey	Private research organisation	**Ind.Eff.:**86 **Ind.Cost:**127,400 **Ind.Bud.:**127,400

Abstract:

Web3.0 has already appeared in the public vocabulary over 5 years ago. While its definition remains unclear, what has become clear in the last half decade is that the web has become a support for social media. Directly from cameras, phones, tablets or computers, users are pushing multimedia data towards their peers and the world at large. MUCKE addresses this stream of multimedia social data with new and reliable knowledge extraction models designed for multilingual and multimodal data shared on social networks. It departs from current knowledge

extraction models, which are mainly quantitative, by giving a high importance to the quality of the processed data, in order to protect the user from an avalanche of equally topically relevant data. It does so using two central innovations: automatic user credibility estimation for multimedia streams and adaptive multimedia concept similarity. Credibility models for multimedia streams are a highly novel topic, which will be cast as a multimedia information fusion task and will constitute the main scientific contribution of the project. Adaptive multimedia concept similarity departs from existing models by creating a semantic representation of the underlying corpora and assigning a probabilistic framework to them. The utility of these two innovations will be demonstrated in an image retrieval system. Extensive evaluation will be performed in order to assess the reliability of the extracted knowledge against representative datasets. Additionally, a new, shared evaluation task focused on user credibility estimation will be proposed. The two core innovations rely on innovative text processing, image processing and fusion methods. Text processing will concentrate on tasks such as word sense disambiguation, concept recognition and anaphora resolution. Image processing will include parsimonious content description, large scale concept detection and detector robustness. Multimedia fusion will focus on a flexible combination of text and image modalities based on a probabilistic framework. All proposed methods will be designed to take advantage of the structural properties of the social networks. Particular focus will be placed on the proposition of scalable algorithms, which cope with large-scale, heterogeneous data. The consortium is formed of four partners, three universities and one research institute with complementary competences that cover the scientific domains associated to the project. Together, in MUCKE, they will introduce new models for processing noisy multimodal and multilingual data that will constitute the base for innovative services.

Relevance:

MUCKE responds to the main objective of the call by proposing novel methods for automatic and reliable new knowledge extraction from heterogeneous multimedia data found on social networks. A semantic and structured description of the data will be obtained through complex multimedia processing applied to large datasets. The main differences with respect to state-of-the art knowledge processing methods come from the central role given to the notion of credibility and from the modelling of raw data using a large set of unambiguous multimedia concepts. The obtained knowledge will be used to fuel the novel user credibility and concept similarity models developed. Thorough evaluation of the developed techniques will be performed on shared datasets and MUCKE will also propose a novel evaluation task related to user credibility. The project will put together partners' expertise from fields such as language processing, image processing, user modelling and information retrieval in order to make each field benefit from best practices from the other fields.

Proposal Book [- D2K - CHIST-ERA Eligibility - YES]

Proposal Data

Acronym	READERS
Full name	Readers: Evaluation And DEvelopment of Reading Systems
Duration	36
Topic	D2K
Keywords	Unsupervised Knowledge Acquisition, Machine Reading
Ranking	0

Coordinator contact point for the proposal

Name	Anselmo Peñas
Institution	Universidad Nacional de Educación a Distancia
Address	Juan del Rosal, 16; 28040 Madrid; Spain
Country	Spain
Phone	+34 91 398 7750
Email	anselmo@lsi.uned.es

Consortium Partners

C/P	Institution	Contact	Other	Country	Legal Status	
C	Universidad Nacional de Educación a Distancia	Anselmo Peñas		Spain	Public research organisation	**Ind.Eff.:**36 **Ind.Cost:**100,000 **Ind.Bud.:**100,000
P	University of Edinburgh	Maria Mirella Lapata		United Kingdom	Public research organisation	**Ind.Eff.:**36 **Ind.Cost:**476,934 **Ind.Bud.:**381,551
P	Synapse Développement	Dominique Laurent		France	Small and Medium Enterprises (SMEs)	**Ind.Eff.:**92 **Ind.Cost:**517,000 **Ind.Bud.:**232,650
P	UPV/EHU	Eneko Agirre	German Rigau	Spain	Public research organisation	**Ind.Eff.:**24 **Ind.Cost:**100,000 **Ind.Bud.:**100,000

Abstract:

The READERS project proposes new unsupervised computational models to automatically extract background knowledge after reading large amounts of unstructured text. This knowledge will be in the form of classes, categorized entities and predicates whose arguments are typified by probability distributions over classes. Classes themselves will be automatically organized into taxonomies related to the predicates in which they participate. In this way, new methods and models based on extensional definitions of concepts will be developed and deployed for the automatic creation of knowledge bases. Important, these will be closely related to textual representations and instrumental in enabling textual inferences. The extracted knowledge will be also linked to external

human-made resources such as Freebase, DBPedia and WordNet, and the knowledge bases will be interfaced with several engines for performing disambiguation, relation extraction, term expansion, and measuring relatedness. A key part of the project will be the development of a reading matching that will use all these resources and tools. . The purpose of our reading machine is to answer queries about a given text. Texts are never self-contained and their interpretation always requires recovering large amounts of background knowledge. Thus, the Machine Reading technology under development must incorporate not only language processing but also the recovery and use of large amounts of background knowledge. This Machine Reading technology will be evaluated through Multiple-Choice Reading Comprehension tests (MRC) developed by humans over unseen documents. MRC tests enable objective and reproducible evaluation experiments, and will be 100% reusable as benchmarks available for the international community. Interestingly, the industrial partner in charge of the Machine Reading system development will apply the reverse technology to automatically generate MRC tests for the automatic assessment of children's reading abilities. This reading machine will work with at least two languages, English and French. The support and coordination of an international evaluation campaign for Machine Reading in multiple languages (English, Spanish, French, German, Italian, Romanian, Bulgarian and Arabic) is part of the proposal. This evaluation campaign will serve to measure the progress in the development of the Machine Reading technology in a comparative/competitive environment. Evaluation exercises in specific domains such biomedicine will also provide a venue for technology transfer and allow us to assess the portability of the proposed technology.

Relevance:

The READERS project proposes new concepts and models for the challenging Machine Reading task. Core to the proposal is the development of unsupervised methods for the acquisition of background knowledge, its linking to existing knowledge bases, and the creation of new ones. As knowledge will be acquired continuously and incrementally, it will progressively enhance and accelerate the reading abilities of the machine. Methodologies, tools and formats will be substantially generic to deal with several languages and different sources of text. In the extreme, the project will explore also some unsupervised techniques to induce semantic roles without the need of parsers (enabling independence from language). This setting must serve to develop a deeper fundamental understanding of the knowledge acquisition process. In particular, how extensional definitions of concepts can be organized into knowledge bases useful for textual inference. Interdisciplinary collaboration not only among linguists and computer scientists but also with biomedical experts will be enhanced through the inclusion of the biomedical domain in the Machine Reading international evaluation campaigns coordinated by the project. The inclusion of an industrial partner with the role of integrating and developing the final system will ensure the transfer of technology from academic partners (Expected impact 3). In the long term, this research will serve to advance the state-of-the-art on reading technology and the ability to convert unstructured texts into machine operable information. The project will help develop a better understanding of the knowledge acquisition problem, allowing for a deeper machine understanding of text (Expected impact 1), and will accelerate the emergence of new applications on the boundaries of intelligent text processing (Expected impact 4).

Proposal Book [- D2K - CHIST-ERA Eligibility - YES]

Proposal Data

Acronym	REFRAME
Full name	Rethinking the Essence, Flexibility and Reusability of Advanced Model Exploitation
Duration	36
Topic	D2K
Keywords	machine learning, data mining, knowledge reuse, model exploitation, operating context
Ranking	0

Coordinator contact point for the proposal

Name	Prof Peter Flach
Institution	Intelligent Systems Lab, University of Bristol
Address	Merchant Venturers Building, Woodland Road, Bristol BS8 1UB
Country	United Kingdom
Phone	+44 117 9545162
Email	Peter.Flach@bristol.ac.uk

Consortium Partners

C/P	Institution	Contact	Other	Country	Legal Status	
C	Intelligent Systems Lab, University of Bristol	Prof Peter Flach	Dr Oliver Ray, Prof Nello Cristianini, Dr Nanlin Jin	United Kingdom	Public research organisation	**Ind.Eff.:**77 **Ind.Cost:**687,398 **Ind.Bud.:**466,160
P	University of Strasbourg	Dr Nicolas Lachiche	Dr Agnès Braud, Prof Pierre Gançarski, Dr Anne Puissant	France	Public research organisation	**Ind.Eff.:**96 **Ind.Cost:**435,094 **Ind.Bud.:**289,417
P	Technical University of Valencia	Dr Jose Hernandez-Orallo	Dr M Jose Ramirez, Dr Cesar Ferri	Spain	Public research organisation	**Ind.Eff.:**84 **Ind.Cost:**478,776 **Ind.Bud.:**183,000

Abstract:

Reuse of learnt knowledge is of critical importance in the majority of knowledge-intensive application areas, particularly because the operating context can be expected to vary from training to deployment. In machine learning this is most commonly studied in relation to variations in class and cost skew in classification. While this is evidently useful in many practical situations, there is a clear and pressing need to generalise the notion of operating context beyond the narrow framework of skew-sensitive classification. This project aims to address the challenge of redesigning the entire data-to-knowledge (D2K) pipeline in order to take account of a significantly generalised notion of operating context. We will develop an innovative and principled approach to knowledge reuse which will allow a range of known machine learning and data mining techniques to deal with common contextual changes, including: (i) changes in data representation; (ii) the availability of new background knowledge; (iii) predictions required at a different aggregation level; and (iv) models to be applied to a different subgroup or distribution. The approach is based around the new notion of model reframing, which can be applied to inputs

(features), outputs (predictions) or parts of models (patterns), in this way generalising, integrating and broadening the more traditional and diverse notions of model adjustment in machine learning and data mining. The ultimate goal of the project is to provide a much better understanding of the issues involved in the generation and deployment of a model for different contexts, as well as the development of tools which ease the extraction, reuse, exchange and adaptation of knowledge for a wide spectrum of operating contexts. The project will focus on three complex domain areas: geographical applications with spatio-temporal data, smart use of energy (resource production and consumption), and human genomics (genotype-phenotype relation analysis). These three demanding domains will ground the project by means of challenge problems and allow us to experimentally validate our methodologies, tools and algorithms.

Relevance:

The project connects to all topics of the call, but significantly addresses the following ones in particular. Topic 3 (Generic models and systems for processing highly heterogeneous data, especially involving different levels or scales): heterogeneous, multi-level and multi-scale data across operating contexts is at the heart of the project. Topic 4 (Systems able to know when they don't know and dynamically cope with unpredicted input data): model reframing is specifically designed to deal with dynamic contexts that change in unpredictable ways. Topic 6 (Generic methodologies, tools and formats to ease the exchange of data and models): the possibilities for model reuse and exchange will be significantly extended by the reframing techniques developed in this project.

Proposal Book [- D2K - CHIST-ERA Eligibility - YES]

Proposal Data

Acronym	uComp
Full name	Embedded Human Computation for Knowledge Extraction and Evaluation
Duration	36
Topic	D2K
Keywords	human computation, knowledge resource acquisition, open evaluation methods, natural language processing, ontology engineering, heterogeneous web data
Ranking	0

Coordinator contact point for the proposal

Name	Dr Wilhelmus Peters
Institution	The University of Sheffield / Department of Computer Science
Address	Computer Science, Regent Court, 211 Portobello, Sheffield, S1 4DP
Country	United Kingdom
Phone	+44 114 222 1901
Email	w.peters@dcs.shef.ac.uk

Consortium Partners

C/P	Institution	Contact	Other	Country	Legal Status	
C	The University of Sheffield / Department of Computer Science	Dr Wilhelmus Peters	Dr Mark Hepple, Dr Kalina Bontcheva	United Kingdom	Public research organisation	**Ind.Eff.:**57 **Ind.Cost:**553,348 **Ind.Bud.:**442,675
P	MODUL University Vienna / Department of New Media Technology	Prof Dr Arno Scharl	Dr Marta Sabou	Austria	Private research organisation	**Ind.Eff.:**65 **Ind.Cost:**451,954 **Ind.Bud.:**319,516
P	Vienna University of Economics and Business / Research Institute for Computational Methods	Prof Dr Kurt Hornik	Dr Gerhard Wohlgenannt	Austria	Public research organisation	**Ind.Eff.:**68 **Ind.Cost:**440,409 **Ind.Bud.:**297,486
P	LIMSI-CNRS / Man-Machine Communication Dept	Dr Patrick Paroubek	Dr Xavier Tannier	France	Public research organisation	**Ind.Eff.:**56 **Ind.Cost:**535,277 **Ind.Bud.:**188,165

Abstract:

The rapid growth and fragmented character of social media and publicly available structured data challenges established approaches to knowledge extraction. Many algorithms fail when they encounter noisy, multilingual and

contradictory input. Efforts to increase the reliability and scalability of these algorithms face a lack of suitable training data and gold standards. Given that humans excel at interpreting contradictory and context-dependent evidence, the uComp project will address the above mentioned shortcomings by merging collective human intelligence and automated methods in a symbiotic fashion. The project will build upon the emerging field of Human Computation (HC) in the tradition of games with a purpose and crowdsourcing marketplaces. It will advance the field of Web Science by developing a scalable and generic HC framework for knowledge extraction and evaluation, delegating the most challenging tasks to large communities of users and continuously learning from their feedback to optimise automated methods as part of an iterative process. A major contribution is the proposed foundational research on Embedded Human Computation (EHC), which will advance and integrate the currently disjoint research fields of human and machine computation. EHC goes beyond mere data collection and embeds the HC paradigm into adaptive knowledge extraction workflows. An open evaluation campaign will validate the accuracy and scalability of EHC to acquire factual and affective knowledge. In addition to novel evaluation methods, uComp will also provide shared datasets and benchmark EHC against established knowledge processing frameworks. While uComp methods will be generic and evaluated across domains, climate change was chosen as the main use case for its challenging nature, subject to fluctuating and often conflicting interpretations. Collaborating with international organisations such as EEA, NOAA and NASA will increase impact, provide a rich stream of input data, attract and retain a critical mass of users, and promote the adoption of EHC among a wide range of stakeholders.

Relevance:

The inter-disciplinary uComp project will study new concepts and models based on Human Computation (HC) to support robust and scalable knowledge extraction from large, noisy, and multilingual content such as social media and Linked Open Data in line with the D2K call and its target outcomes (TO). When encountering challenging tasks (TO1) or contradictory, complex and multilingual data (TO2), automated knowledge extraction methods will gain the ability to cope in an adaptive manner, even when faced with highly dynamic or unpredictable input (TO4). This will be achieved by tapping into collective knowledge and embedding HC into complex knowledge processing workflows. The result is a novel, generic and reusable approach (TO3) to knowledge extraction and evaluation. uComp will stimulate efforts to create sizable public training data and evaluation resources, a largely neglected but vital task to guide ontology matching, the extension of Linked Open Data repositories, and the extraction of multilingual affective resources (TO5). uComp will leverage existing open source research platforms and support standards to promote its take-up, reuse and sustainability (TO6).

Proposal Book [- D2K - CHIST-ERA Eligibility - YES]

Proposal Data

Acronym	Visual Sense (ViSen)
Full name	Tagging visual data with semantic descriptions
Duration	36
Topic	D2K
Keywords	Image and video annotation, image search,
Ranking	0

Coordinator contact point for the proposal

Name	Krystian Mikolajczyk
Institution	University of Surrey/Department of Electronic Engineering
Address	University of Surrey , GU27XH, Guildford, UK
Country	United Kingdom
Phone	00441483683959
Email	K.Mikolajczyk@surrey.ac.uk

Consortium Partners

C/P	Institution	Contact	Other	Country	Legal Status	
C	University of Surrey/Department of Electronic Engineering	Krystian Mikolajczyk	Fei Yan	United Kingdom	Public research organisation	**Ind.Eff.:**43 **Ind.Cost:**497,183 **Ind.Bud.:**397,746
P	Institut de Robòtica i Informàtica Industrial	Francesc Moreno-Noguer	Ariadna Quattoni, Xavier Carreras	Spain	Public research organisation	**Ind.Eff.:**72 **Ind.Cost:**200,000 **Ind.Bud.:**200,000
P	Ecole Centrale de Lyon	Emmanuel Dellandréa	Liming Chen, Alexandre Saidi	France	Public research organisation	**Ind.Eff.:**108 **Ind.Cost:**329,318 **Ind.Bud.:**329,318
P	University of Sheffield, Department of Computer Science	Prof Robert Gaizauskas	Paul D Clough, Ahmet Aker	United Kingdom	Public research organisation	**Ind.Eff.:**44 **Ind.Cost:**458,764 **Ind.Bud.:**367,011

Abstract:

Today a typical Web document will contain a mix of visual and textual content. Most traditional tools for search and retrieval can successfully handle textual content, but are not prepared to handle hetereogeneous documents. The new type of content demands the development of new efficient tools for search and retrieval. The visual sense project aims at mining automatically the semantic content of visual data to enable "machine reading" of images. In recent years, we have witnessed significant advances in the automatic recognition of visual concepts (VCR). These advances allowed for the creation of systems that can automatically generate keyword-based image annotations. The goal of this project is to move a step forward and predict semantic image representations that can be used to

generate more informative sentence-based image annotations. Thus, facilitating search and browsing of large multi-modal collections. More specifically, the project targets three case studies, namely image annotation, re-ranking for image search, and automatic image illustration of articles. It will address the following key open research challenges: 1. To develop methods that can predict a semantic representation of visual content. This representation will go beyond the detection of objects and scenes and will also recognize a wide range of object relations. 2. To extend state-of-the-art natural language techniques to the tasks of mining large collections of multi-modal documents and generating image captions using both semantic representations of visual content and object/scene type models derived from semantic representations of the multi-modal documents. 3. To develop learning algorithms that can exploit available multi-modal data to discover mappings between visual and textual content. These algorithms should be able to leverage 'weakly' annotated data and be robust to large amounts of noise. For this purpose, the current project will build on expertise from multiple disciplines, including computer vision, machine learning and natural language processing (NLP), and gathers four research groups from University of Surrey (Surrey, UK), Institut de Robòtica i Informàtica Industrial (IRI, Spain) , Ecole Centrale de Lyon (ECL, France), and University of Sheffield (Sheffield, UK) having each well established and complementary expertise in their respective areas of research.

Relevance:

The visual sense project aims at mining the semantic content of visual data and involves interdisciplinary expertise, including the one from computer vision, machine learning and natural language processing. Specifically, the current project answers the focus of the CHIST-ERA call 2011 on data to new knowledge (D2K) in the following way: Target outcome No.1. The automatic generation of sentence-based image annotations rather than keywords-based annotations provides a more accurate and informative description of visual content and enables machine reading of images and videos. These new tools will allow more accurate search, browsing and exploitation of large multimedia document collections. Target outcome No.2. Visual content on the web presents itself embedded in diverse textual contexts. The textual context itself can take a variety of forms such as: 1) articles in which images are embedded, 2) keyword tags describing image content 3) highly descriptive image captions and 4) image comments loosely related to the visual content . Making sense of such data is a challenging task but solving it will allow us to untap a great source of new knowledge. Target outcome No.3. The main focus of the Visual Sense project is the development of machine learning methods for knowledge and information extraction from large collections of visual content. The tools developed in this project will have a variety of applications. To demonstrate it, we will focus on 1) the generation of descriptive image captions, 2) re-ranking for improved image search and 3) automatic illustration of article with image.