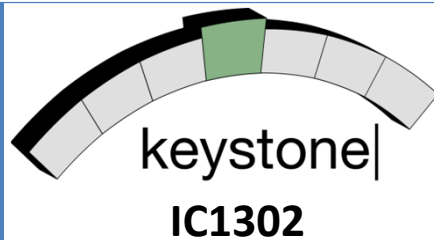
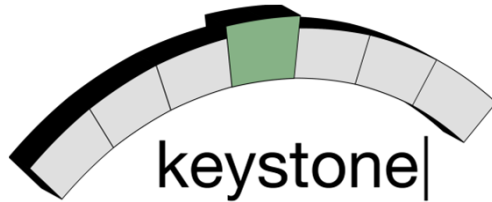


Brussels, 15 October 2013
1st MC Meeting



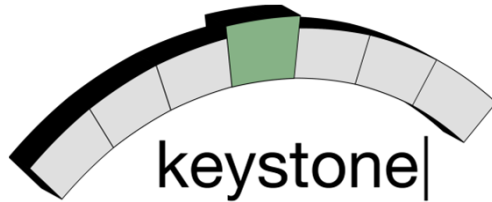
Workplan for the implementation
of the COST Action

Francesco Guerra
University of Modena and Reggio Emilia - Italy



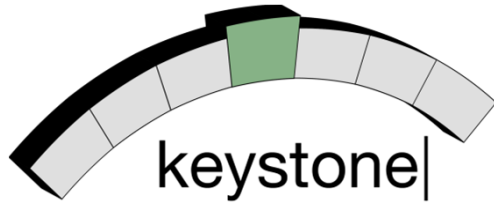
Outline

- ▶ Overview of the scientific topic
- ▶ Objectives and working programme
- ▶ Working method - organisation and management
- ▶ Distribution of tasks
- ▶ Promotion of gender balance and of Early Stage Researchers (ESR)
- ▶ Time-table

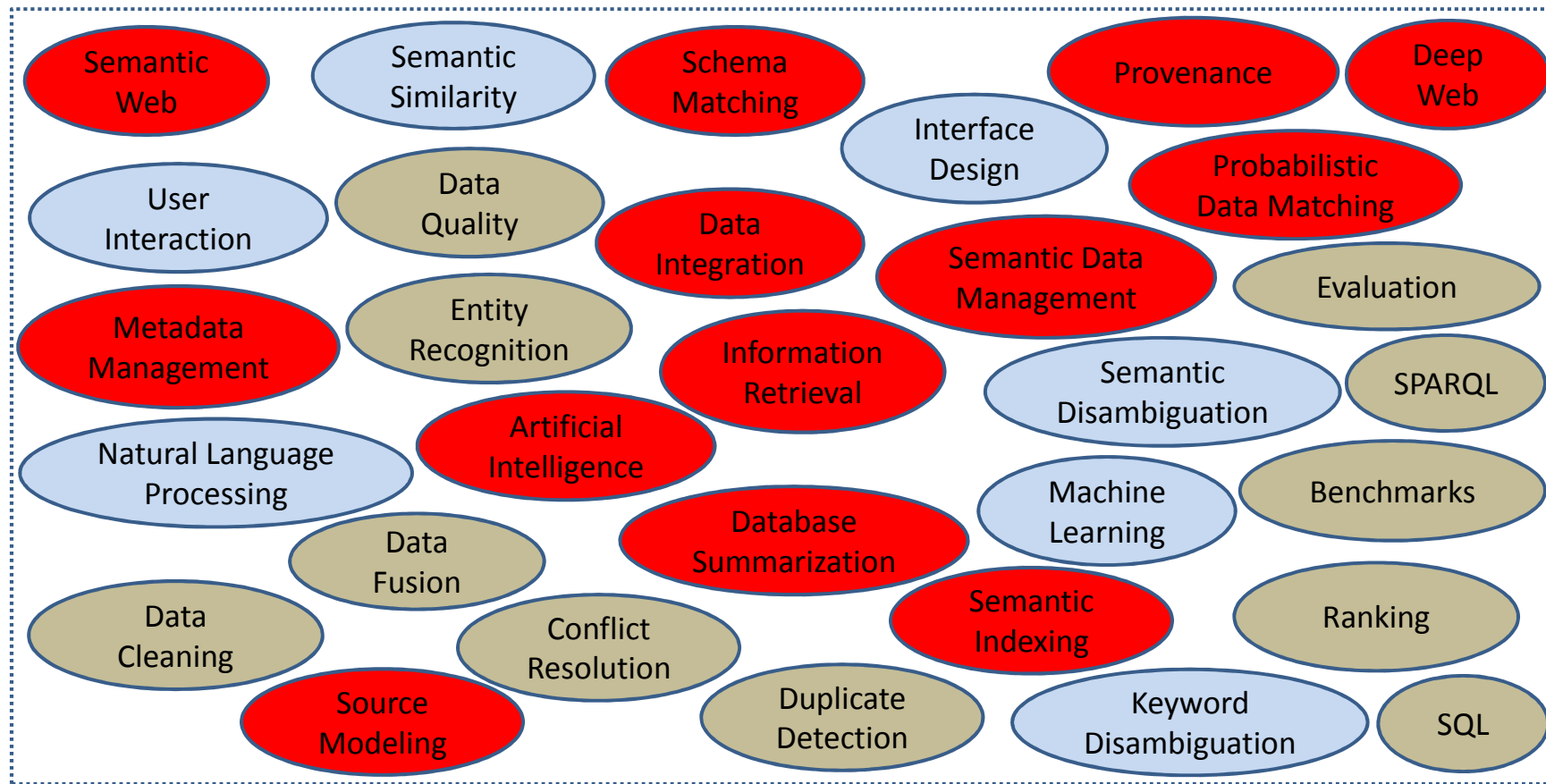


Overview

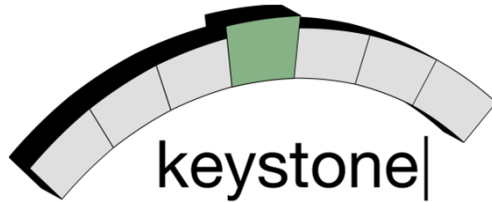
- ▶ Keyword-based search is the de-facto standard for finding information
 - ▶ Existing approaches for structured data sources suffer from limitations
 - ▶ multi-source scenarios
 - ▶ sources with frequent updates
 - ▶ large amount of data
 - ▶ Lack of benchmarks / evaluation framework
- ▶ Lot of research effort in the past into supporting the availability and interoperability among sources
 - ▶ techniques for publishing, semantically enriching, matching and integrating structured data
 - ▶ at the database level
 - ▶ at the Web level
 - ▶ Typical scenarios: open data, big data, linked data, ...
- ▶ Most efforts have not focused on techniques for searching and retrieving data from these sources
 - ▶ Need for the **KEYSTONE ACTION!**



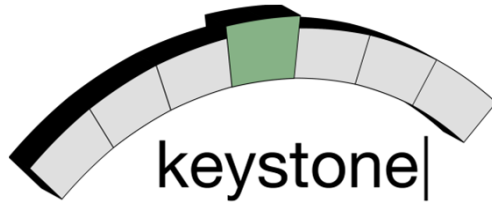
The multi-faceted nature of the problem



- ▶ KEYSTONE key concepts from the research perspective

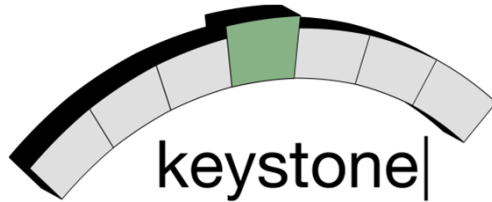


- ▶ Launch a **cooperative network** to enable **research activity** and **technology transfer** in the area of keyword-based search over structured data sources
- ▶ Provide **users** with keyword-based **search capabilities** for structured data sources as they currently do with documents
- ▶ Interested people research in **fields** related to semantic data management, the Semantic Web, information retrieval, artificial intelligence, machine learning and natural language processing, etc.



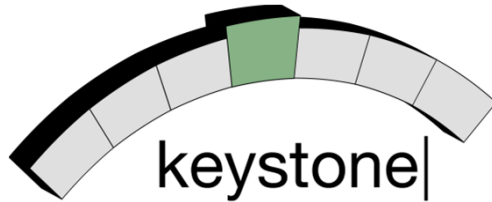
Objectives & Drivers

1. Promote the research involving European units
 1. Critical analysis of the emergent techniques
 2. Conceptual map of open / closed issues
 3. Scenarios, showcases and benchmarks for evaluations and comparisons
 4. Writing / Proposing proposals to European Call
2. Facilitate the transfer of knowledge and technology
 1. Scientific papers, newsletters and discussion groups
 2. Meetings, training schools and academic / industrial events
3. Build a critical mass of research activities
 1. Promotion participation in the network
4. Support the involvement of early stage researchers and gender balance
 1. Idea sharing among ESRs and Senior Researchers
 2. Lectures for Master Thesis and PhD
 3. Equal participation in the positions of responsibility (gender balance)
 4. Supporting ESR participation in the Action coordination roles



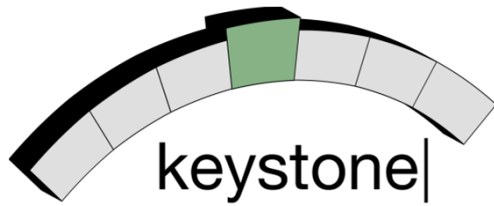
Promoted Activities

- ▶ **Activity 1: Setup of a common repository of knowledge**, supporting the development of an open annotated bibliography of the relevant literature and a collection of the materials produced by the Action
- ▶ **Activity 2: Organization of Working Group meetings** for sharing the research achievements involving external audience from related disciplines
- ▶ **Activity 3: Organization of open workshops** for the dissemination and the discussion of the outcomes achieved within the Working Groups to a larger audience
- ▶ **Activity 4: Organization of training and disseminations activities** including tutorials, summer schools and industrial events
- ▶ **Activity 5: Exchanges of early stage researchers** for visiting other partners, performing joint researches and sharing methodologies
- ▶ **Activity 6: Visiting activities for senior researchers** to coordinate KEYSTONE, identifying the relevant open issues in the field, the promising research directions, and the methodologies to be exploited



Tools supported by COST for achieving KEYSTONE goals

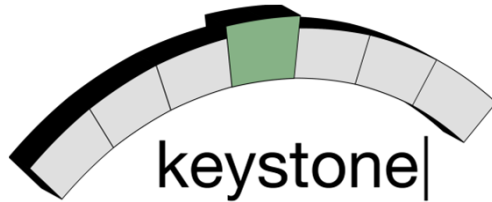
- ▶ Website
- ▶ STSMs
- ▶ Meetings
- ▶ Training Schools
[not scheduled in KEYSTONE first year]



KEYSTONE website

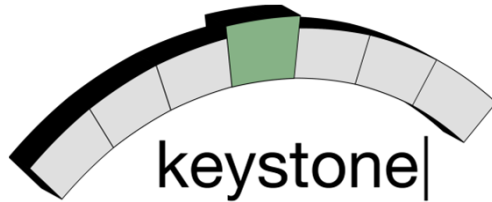


Website



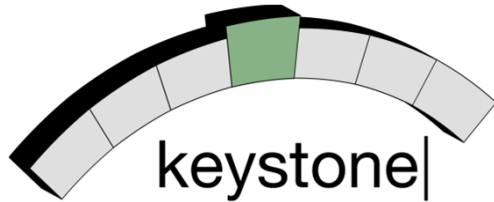
KEYSTONE website

- ▶ Creation of a list of profiles of researchers and practitioners working in the field
- ▶ Creation of a map of the hot research topics
- ▶ Providing a platform for matching «supply with demand»
 - ▶ Offers for internships / STSM
 - ▶ Proposals for joint researches
 - ▶ Knowledge and technology transfer
- ▶ Communication of research findings / activities (through a regular newsletter) / promotion / organization of special issues /tracks
- ▶ Information about calls / EU call
- ▶ Integration with other social media
 - ▶ LinkedIn group
- ▶ MAIN GOAL: keeping people in touch!
 - ▶ Making researchers aware that a community exists working in these fields
 - ▶ Making researchers aware that we have some «tools» for improving coordination and networking

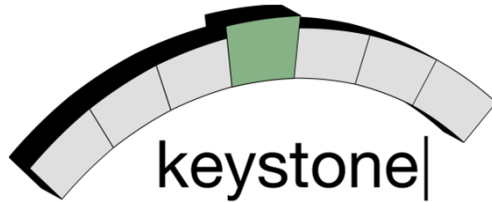


Problem

- ▶ How to make people, other than the MC members, participant in the Action activities?
- ▶ How to make KEYSTONE a reference for the research in the area?
- ▶ How to promote the population of contents in the website by the Action participants?



- ▶ Standard STSMs need to take place according to the following rules:
 - ▶ minimum duration of 5 days; maximum duration of 90 days;
 - ▶ carried out within a single grant period and within the Action's lifetime
- ▶ Specific provisions had been adopted by the CSO to foster the participation of ESR according to the following rules:
 - ▶ minimum duration of 91 days; maximum duration of 180 days;
 - ▶ carried out within a single grant period and within the Action's lifetime
- ▶ Which are the criteria for approving / selecting among STSM applications?
- ▶ Calls? Open Submissions?
- ▶ How to promote the submission of applications (by ESR in particular)?

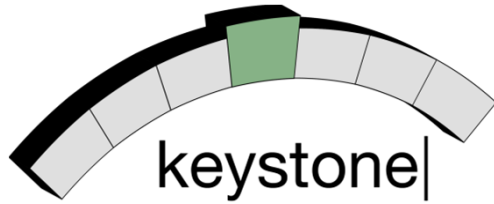


Meetings

- ▶ Goal
 - ▶ Share / discuss issues, ideas and results
 - ▶ Promote new research activities
 - ▶ Promote the participation in EU proposals

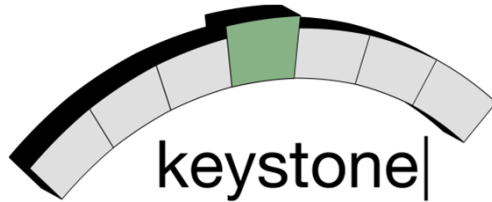
- ▶ Organization
 - ▶ Keynote Speakers
 - ▶ Panels, Round tables

- ▶ Participation of external experts
 - ▶ e.g. Experts / proponents / officers of calls H2020



Discussion

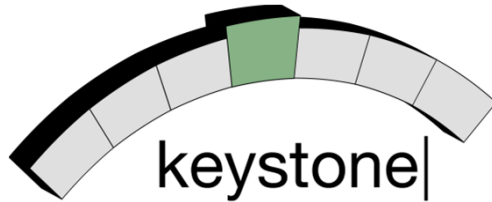
1. WEBSITE (and other dissemination channels)
2. STSM
3. MEETINGS



Four working groups

- ▶ **WG1: Representation of structured data sources**
 - ▶ How to extract a set of metadata describing data sources and to index them for their retrieval
- ▶ **WG2: Keyword search**
 - ▶ How to select the relevant sources for a query and match keywords with the data structures
- ▶ **WG3: User interaction and keyword query interpretation**
 - ▶ How to express constraints, to disambiguate the meaning of the query to exploit feedback
- ▶ **WG4: Research integration, showcases, benchmarks, and evaluations**
 - ▶ How to integrate the constituent functionalities into a search engine and how to evaluate it

- ▶ Four tasks addressed by all the WGs
 1. Critical review of the existing emerging techniques
 2. Definition of open/closed issues in the field and definition of a roadmap for achieving the KEYSTONE outcomes
 3. Coordination of research activities
 4. Dissemination of the outcomes



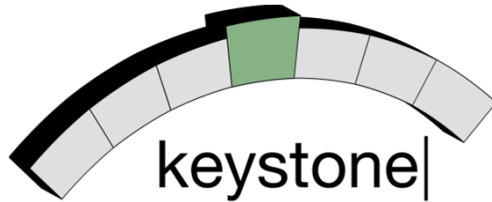
Preliminary Scientific Programme

▶ **WG1 Representation of Structured Data Sources**

- ▶ Study possible metadata formats for describing data sources
- ▶ Study and extend techniques for publishing structured data sources on the Web in a semantically-rich form.
- ▶ Study of standard methodology in order to perform the publication of LOD
- ▶ Study metadata for the description of Social data and the types of metadata available in each platform
- ▶ Discovery and Indexing of Structured and Semi-structured Data Sources
- ▶ Indexing of data sources based on summarizations of data contents.
- ▶ Implementation of territory dashboards, based on a geographic granularity and a theme, using open data available
- ▶ Generic data visualization tool for open data sets.

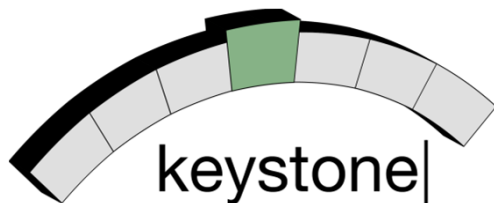
▶ **WG2 Keyword Search**

- ▶ Critical analysis and develop techniques for summarising and indexing structured data sources
- ▶ Develop and extend approaches with advanced techniques exploiting statistics and semantics, knowledge extracted from the metadata and in query logs
- ▶ Develop techniques for distinguishing between schema-related and instance-related keywords
- ▶ Develop techniques for generating structured queries, taking into account different join paths and query execution plans
- ▶ Develop techniques for querying in multi-source scenarios
- ▶ Study, analyse and develop techniques that are made to scale.



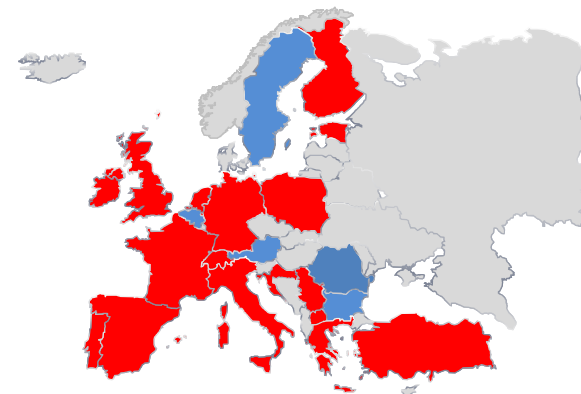
Preliminary Scientific Programme

- ▶ **WG3 User Interaction and keyword query interpretation**
 - ▶ Develop of user interfaces
 - ▶ Exploring how semantics can be used to enhance search precision and recall in specific industry domains
 - ▶ Exploring how metadata and semantics can be used to semantically enhance the description of business services and service networks
 - ▶ Investigating how the precision of search algorithms is enhanced by exploring metadata, folksonomies and ontological concepts
 - ▶ Develop of functions to enrich user keywords
 - ▶ Study of paradigms for formulating keyword queries, including a "common query language" for accessing different storage engines
 - ▶ Investigating the application of NLP for (1) the disambiguation of keywords, (2) the recognition of named entities (3) the identification of multi-word keywords or in the segmentation of aggregated keywords
- ▶ **WG4 Research Integration, Showcases, Benchmarks And Evaluations**
 - ▶ Identify existing solutions and prototypes, software libraries, LOD, datasets and query sets.
 - ▶ Develop a set of showcases/scenarios with a special emphasis on business and cloud services and social Web
 - ▶ Study criteria for evaluating the current techniques and prototypes
 - ▶ Benchmark the overall solution developed
 - ▶ Development of datasets to serve as baseline for future benchmarking initiatives
 - ▶ Performance issues on big Data



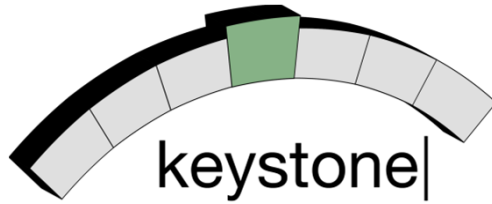
The Network

	Country	WG1	WG2	WG3	WG4
MARTINCIC IPSIC	HR		■	■	
LIIV	EE		■		
TERZIYAN	FI			■	■
BOUCELMA - MANOLESCU	FR			■	
DELEV - MADJAROV	MK			■	
DEMIDOVA – NURNBERGER	DE	■	■	■	
SALAMPASIS – STAVRAKAS	GR				
BRESLIN - AVRAM	IE	■		■	
HACOHEN-KERNER	IL		■		
GUERRA – VELEGRAKIS	IT	■	■		■
ABELA	MT	■	■		
SCHALK - HOUBEN	NL	■			■
SZYMANSKI - NGUYEN	PL	■			■
CARDOSO – MATOS	PT			■	■
STANKOVIC	RS			■	
TRILLO LADO - FARINA MARTINEZ	ES	■	■		
FALQUET – CRESTANI	CH			■	■
ACAR	TK		■	■	
MISSIER – CALI'	UK	■	■		



In blue, countries that will be shortly included

(Sorry! Malta and Israel are not represented in the map)



Organisation



- Chair / Co – Chair
- Scientific Coordinator
- Dissemination Coordinator
- Training Coordinator
- STSM Coordinator
- WG Leaders

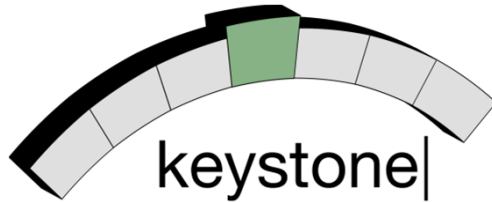
**Executive
Scientific Board**

Networking

- ▶ Scientific Coordination (**Executive Scientific Board**)
 - ▶ 4 Working Groups
 - ▶ Intra WG workshops
 - ▶ Open annotated bibliography
- ▶ STSMs (**STSM Coordinator**)
 - ▶ Exchanges (Early Stage Researchers), visiting (Senior Researchers)

Dissemination

- ▶ Publications (**Editorial Board**)
 - ▶ Website (with tools enabling discussion)
 - ▶ Newsletter
 - ▶ Scientific publications
- ▶ Events
 - ▶ Open WG Meetings / Industrial days
 - ▶ Training schools (**Training Coordinator**)
 - ▶ Workshops and tutorials



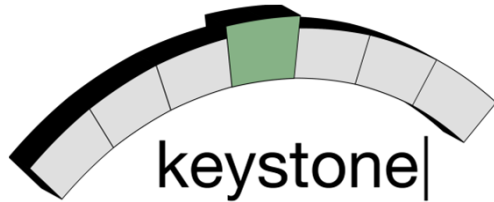
Responsibility Position

- ▶ Chair:
- ▶ Co – Chair:
- ▶ Scientific Coordinator:
- ▶ Dissemination Coordinator:

- ▶ WG1 Leader:
- ▶ WG2 Leader:
- ▶ WG3 Leader:
- ▶ WG4 Leader:

- ▶ Financial Rapporteur 1:
- ▶ Financial Rapporteur 2:

- ▶ Training Coordinator:
- ▶ STSM Coordinator:



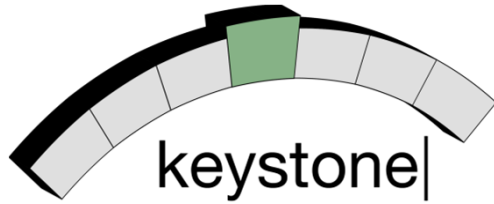
Responsibility Position: proposal

- ▶ Chair: **Francesco Guerra (IT)**
- ▶ Co – Chair: **Jorge Cardoso (PT)**
- ▶ Scientific Coordinator: **Yannis Velegarakis (IT)**
- ▶ Dissemination Coordinator: **John Breslin (IE)**
(supported by Catarina Ferreira (FR) and Maciej Dabrowski (IE))

- ▶ WG1 Leader: **Raquel Trillo (ES), Stefan Dietze (DE)**
- ▶ WG2 Leader: **Elena Demidova (DE), Julian Szimanski (PL)**
- ▶ WG3 Leader: **Omar Boucelma (FR), Ioana Manolescu (or Bodgan Cautis) (FR)**
- ▶ WG4 Leader: **Paulo Rupino (PT), Ngoc Thanh Nguyen (PL)**

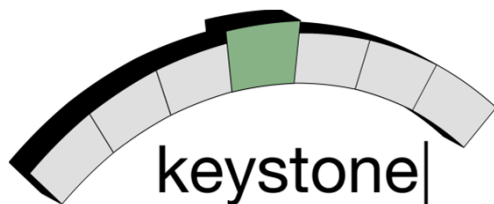
- ▶ Financial Rapporteur 1: **Antonio Farina (ES)**
- ▶ Financial Rapporteur 2: to be defined

- ▶ Training Coordinator: to be defined
- ▶ STSM Coordinator: to be defined



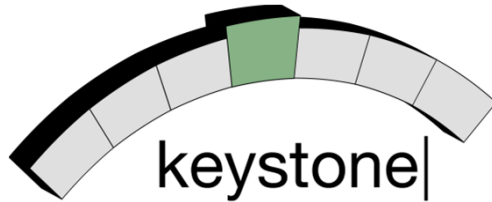
DISCUSSION

1. Scientific Program
2. Responsibility positions



Timetable

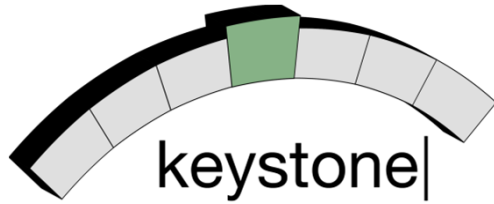
Working Groups	Year 1				Year 2				Year 3				Year 4			
	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4
WG1 – Data Representation																
WG2 – Keyword search																
WG3 – User Interaction																
WG4 – Integration & eval.																
Other activities																
Kick Off Meeting																
WG/MC Meetings																
Training School																
Open Meetings -Industrial Day																
STSMs	At least 12 involving ESR / 12 involving Senior															
Deliverables	Scientific production MC report WG reports STSM report Website operative Open bibliography Kick Off Meeting Newsletter 1 - 2				Scientific production MC report WG reports STSM report Open bibliography Open Meeting 1 Training school 1 Newsletter 3 – 4				Scientific production MC report WG reports STSM report Open bibliography Open Meeting 2 Training school 2 Newsletter 5 – 6				Scientific production MC report WG reports STSM report Final report / book Open bibliography Open Meeting 3 Training school 3 Newsletter 7 – 8			



1° Year Budget

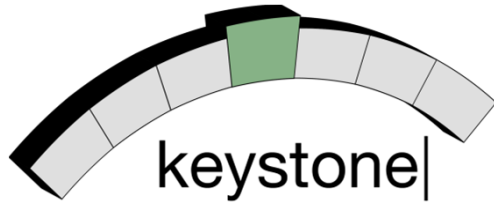
Oct. 2013	Nov. 2013	Dec. 2013	Jan. 2013	Feb. 2013	Mar. 2013	Apr. 2013	May 2013	Jun. 2013	Jul. 2013	Aug. 2013	Sep. 2013
					2nd MC Meeting						3rd MC Meeting
					WG Meetings						
					1st newsletter						2nd newsletter
		website operative									
		1st Call STSM									
						2nd Call STSM					

(1) MEETINGS	€ 90000 (~100 missions)
(2) SHORT-TERM SCIENTIFIC MISSIONS	€ 20000 (~10 STSM)
(3) TRAINING SCHOOLS	€ -
(4) PUBLICATIONS, DISSEMINATION, OUTREACH	€ 4000 (website)
(5) OERSA	€
TOTAL SCIENCE EXPENDITURE (sum of (1) to (5))	€ 114000
Financial & Scientific Administration and Coordination	€ 17100 (15%)
D. TOTAL EXPENDITURE	€ 131100



Discussion

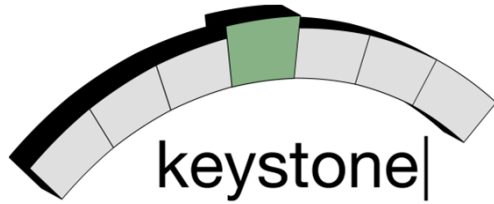
1. Budget allocation and timetable



Participation of ESR / Gender Balance

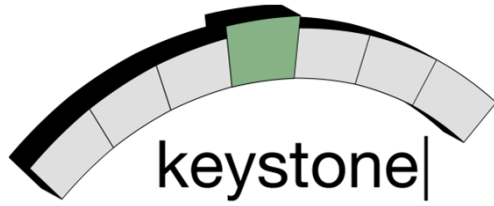
- ▶ Gender Balance
 1. Assuring equal responsibility positions in the Action
 - ▶ 3 out of 6 women in the MC have responsibility positions
 2. Promoting an appropriate balance of women and men in the Action research area
 - ▶ Participation in special events promoting the women's involvement in sectors related to engineering, technology and computer science organized by our Institutions
 - ▶ Assuring equal gender opportunity in the events promoted by the Action (meetings and training schools)

- ▶ Participation of ESR
 - ▶ Reserving positions of WG leaders to Early Stage Researchers
 - ▶ At least 3 responsibility positions are reserved to ESR
 - ▶ Several members are no ESR since few months / years
 - ▶ Focusing STSMs and Training Schools on PhD students and embryonic stage researchers



Monitoring / Evaluating Achievements

- ▶ Work done vs. the planned work
 - ▶ The WG and the Scientific coordinators shall guarantee that the work will progress as planned
- ▶ Self-assessment of the work done
 - ▶ Action reports contain information about a self-evaluation of each research group in terms of:
 - ▶ effectiveness and goal achievements, relevance, efficiency, synergies with other research units and collaborations with external partners
- ▶ Dissemination activities
 - ▶ Quality of the publications, publishing house, web site, magazine or newspaper.
- ▶ Impact on the Research Community
 - ▶ Creation and consolidation of a research community
- ▶ User-Impact
 - ▶ The KEYSTONE participants will be happy and declare it successful if / when the Action tools will become more and more popular among data publishers/searchers



Discussion

1. Final comments