



Project no. 215231

TrebleCLEF

TrebleCLEF: Evaluation, Best Practices and Collaboration for Multilingual Information Access
IST: ICT-1-4-1, Digital libraries and technology-enhanced learning

Deliverable 6.1

Dissemination Plan

Start Date of Project: 01 January 2008

Duration: 24 Months

Organisation Name of Lead Contractor for this Deliverable: CELCT

Version 1, March 2008

Project co-funded by the European Commission within the Seventh Framework Programme

Document Information

Deliverable number: 6.1
Deliverable title: Dissemination Plan
Due date of deliverable: 03|08
Actual date of deliverable: 03|08
Author(s): Amedeo Cappelli, CELCT
Participant(s): All
Workpackage: 6
Workpackage title: Dissemination
Workpackage leader: CELCT
Dissemination Level: PU
Version: 1.0
Keywords: dissemination, publications, events, press, media, web]

History of Versions

Version	Date	Status	Author (Partner)	Description/Approval Level
01	03.03.2008	Draft	CELCT	Circulated to all partners

Abstract

This document describes the initial dissemination plan of the TrebleCLEF coordination action which aims at widening the general understanding and visibility of the project and ensuring an awareness of the results in the interested research and application communities; this first release will be maintained and updated throughout the entire project period in order to keep track of the ongoing project activities.

Table of Contents

Document Information	1
Abstract	1
Executive Summary	3
1. Introduction	5
2. Dissemination Strategy.....	6
2.1 Target Communities.....	6
3. Dissemination Activities	7
3.1 TrebleCLEF Portal.....	7
3.2 Production of Promotional Material	9
3.3 Organisation of Events.....	9
3.4 Training Activities	10
3.5 Participation in Standardization Activities	11
3.6 Contacts and Collaborations	12
3.7 Technology Transfer and Exploitation	12

Executive Summary

This document describes the dissemination plan of TrebleCLEF.

First of all, the objectives of TrebleCLEF are presented, which clearly motivate the setting up of a detailed dissemination plan for the project. The primary goal of TrebleCLEF is the dissemination of knowledge and spreading of excellence in the domain of Multilingual Information Access (MLIA), by creating a community of researchers and developers and facilitating future collaborative initiatives between them. Another important goal is to assist and stimulate the development of European cross-language retrieval systems in order to guarantee their competitiveness in the global market.

In this sense, particular attention will be devoted to the organisation of all initiatives aimed at increasing awareness of the importance of MLIA in the information society, in both academic and industrial subjects at an international level. As a matter of fact, TrebleCLEF intends to reinforce the importance of MLIA in research programs and to propose mature technologies to the industrial community to be used in real applications. For this purpose, the TrebleCLEF dissemination activities will make available tools, data, and all information concerning MLIA best practices to the scientific and application community, which will be identified and involved in the project activities by the means of specific modalities –i.e., setting up a mailing list, creating a User Advisory Board, establishing relationships with projects and initiatives focused on similar subject domains, etc.

This document describes all the initiatives which will be taken in order to promote the general understanding and visibility of the project and ensure the dissemination of the results in the community.

The activities on which the project intends to focus, are the following:

- TrebleCLEF Portal
- Production of Promotional Material
- Organisation of Events
- Training Activities
- Participation in Standardization Activities
- Establishment of Contacts and Collaborations with groups, projects, institutions working in related areas
- Technology Transfer and Exploitation

The portal will be a primary instrument for project dissemination. It has been created at the beginning of the project and will be maintained during the entire project lifetime. The structure and the functions of the portal are described in details in section 3.1 of this document.

Other dissemination actions are planned, which mainly consist in the production of material to be distributed in various events and to different users, such as flyers, posters, videos, slides of workshops and tutorials etc. Such materials are illustrated in section 3.2 .

Particular attention will be devoted to the organisation of events, as presented in section 3.3. Specifically, the organisation of the annual CLEF workshop has to be considered as the central dissemination activity of TrebleCLEF, since one of its main goal is to build on and extend the relevant results already achieved by CLEF. The workshop will be held in conjunction with the European Conference on Digital Libraries in order to maximize the possibility of dissemination of the results in the DL community.

Furthermore, three brain-storming workshops will be organised with the aim of bringing together systems developers and MLIA user communities to study and discuss system-oriented aspects, user requirements and user-oriented evaluation for multilingual information access systems.

Finally, a MLIA Technology Day aiming at bringing together researchers, software developers and

users to facilitate communication among them will be organised. This event will reinforce the awareness of the importance of MLIA technologies in the industrial and users community. The Technology Day is also to be intended as a preparatory action for the transfer and exploitation of the evaluation activities of the project. In fact, the evaluation methodologies and tools developed during the duration of TrebleCLEF will be exploited by offering software companies and general users services and collaborations to improve their products.

The project intends also to organise training activities to disseminate the results of Best Practice studies resulting from the work carried out in WP 3 (Best Practices for system-oriented and user-oriented MLIA), WP 4 (Best Practices for Test Collection Creation) and WP 5 (Best Practices for MLIA Language Resources). These activities will also comprise the preparation of tutorial to be proposed in the most relevant conferences in the area covered by MLIA.

Furthermore, a summer school on MLIA theory and practice will be organised. The school will be structured in a series of lectures and training sessions. The target of the school will be both students and professionals interested in multilingual information systems.

Finally, since TrebleCLEF will open or increase ongoing collaborations with many international institutions and initiatives, it will examine the feasibility of stipulating conventions for widely distributed resources and methodologies, in order to develop standards for tools and evaluation metrics, and guidelines for human evaluation useful for sharing and reusing MLIA technologies (see section 3.5).

1. Introduction

The objective of **TrebleCLEF** is to build on and extend the results already achieved by CLEF. The dissemination of knowledge and the spreading of excellence in the domain of MLIA, has been the primary goal of CLEF since its beginning. Through the organisation of system evaluation campaigns, the main aim has been to create a community of researchers and developers who study the same problems and to facilitate future collaborative initiatives between groups with similar interests. CLEF has also established strong links, exchanging ideas and sharing results, with similar cross-language evaluation initiatives in the US and Asia, which are working on other sets of languages. The goal has been to assist and stimulate the development of European cross-language retrieval systems in order to guarantee their competitiveness in the global marketplace.

The main objective of **TrebleCLEF** will be to reinforce this role by proposing a program of activities for knowledge dissemination in research and industry communities; by making available mature technologies to be used in real applications in MLIA; and by creating a network among all players in MLIA domain (researcher, developers and end users). For this reason, workpackage 6 (Dissemination) will be a key activity of the project. In particular, this workpackage will consider the results achieved by the coordination activities of WPs 2,3,4 & 5 and will exploit them in carefully planned dissemination actions.

The tools and data collected and produced, and all the information about MLIA best practices will be made available to the scientific and application communities through a precise program of dissemination. Targeted workshops and tutorials will be organised to present and discuss the main issues of MLIA. Workshop proceedings will be published both in paper and electronic format. Particular attention will be devoted to the publication of guidelines and best practice case studies, in order to disseminate reports about successful activities to be taken into consideration for future applications.

Thus, **TrebleCLEF** is promoting a two-fold Dissemination Action aimed at both scientific and application communities in order to:

1. continue to encourage community-building and collaboration by providing a forum for the discussion of results and making the scientific data, experiments and results produced during evaluation campaigns publicly available.
2. disseminate know-how, tools, resources and best practice guidelines for applications (and in particular digital libraries) and system developers through the organisation of workshops, tutorials and training sessions.

This Dissemination Plan provides a specific description of the various dissemination activities which will be carried out within the above framework. It will be updated during the project lifetime in order to identify additional initiatives. Details will be given in the 12 and 24 month project progress reports.

The planning of the dissemination activities will identify promotional material and dissemination processes that will optimize the knowledge of the target communities about **Treble-CLEF** results. Particular attention will be paid to technology transfer. Detailed information on the dissemination strategies is provided in the following sections of this document.

2. Dissemination Strategy

The main goal of the dissemination activities is to promote the general understanding and visibility of the project and ensure an awareness of the results in the interested research and application communities.

The project activities and results will be disseminated using the following mechanisms:

- Treble-CLEF Portal
- Production of Promotional Material
- Organisation of Events
- Training Activities
- Participation in Standardization Activities
- Establishment of Contacts and Collaborations with groups, projects, institutions working in related areas
- Technology Transfer and Exploitation

Within this framework, the following list of tasks has been identified as relevant towards satisfying these goals:

- organising the annual campaigns of the Cross-Language Evaluation Forum (CLEF) and the related workshop;
- producing and distributing promotional material for the attendees of the CLEF conference;
- publishing proceedings of the conference with well-known publishers;
- organizing workshops around the key topics of interest to TrebleCLEF activities
- attending other related conferences and events and distributing prepared dissemination material;
- stimulating the adoption of solutions and guidelines in research and industrial projects through tutorials and training;
- submitting articles to international conferences and journals and presenting and promoting the project via networks and other actions;
- preparing and maintaining of dissemination material, flyers, brochures, posters, CDs, DVDs, etc.;
- contacting the major cultural and political organisations, associations, content producers, publishers, to increase awareness about Treble-CLEF activities.

2.1 Target Communities

An important activity of Workpackage 6 will be to identify the target communities that will be interested in the results of Treble-CLEF. This involves the creation of sets of mailing lists for different kinds of communities and the setting up of a User Advisory Board. We will focus on both research and application sectors.

Multilingual information access is a multidisciplinary research area and includes the following fields: information retrieval, system evaluation, image and speech processing, human computer interaction, natural language processing, geographical information systems, etc. Members of the Treble-CLEF consortium are active in most of these areas and no major difficulties are expected with respect to dissemination in these research communities via the organisation and participation in events and the publication of results at conferences and in scientific journals.

The application communities provide more of a challenge. However, as a starting point, TrebleCLEF will take advantage of the network created by the DELOS Network of Excellence. Moreover, we have already begun discussion with some of the major European DL activities (e.g. EDLNet, TEL and MICHAEL); TrebleCLEF is a member of EDLNet and has already provided consulting services with respect to the implementation of multilingual access functionality in Europeana; TrebleCLEF is

represented in Workpackage 2 of this Network by Carol Peters. Similarly an action of collaboration has been begun with TEL; the CLEF 2008 evaluation campaign offers a task on TEL data. The aim is to identify the most effective retrieval technologies for searching the sparse data provided by library catalog cards.

One of the tasks of Workpackage 6 will be to identify and establish contacts with other application communities potentially interested in the activities of TrebleCLEF.

3. Dissemination Activities

In this section we provide details on the dissemination activities planned.

3.1 TrebleCLEF Portal

The TrebleCLEF portal will be a primary instrument for project dissemination. In order to present the project and to stimulate awareness of its existence, scope and possibilities a project website was created at the beginning of the project and will be maintained during the entire project lifetime.

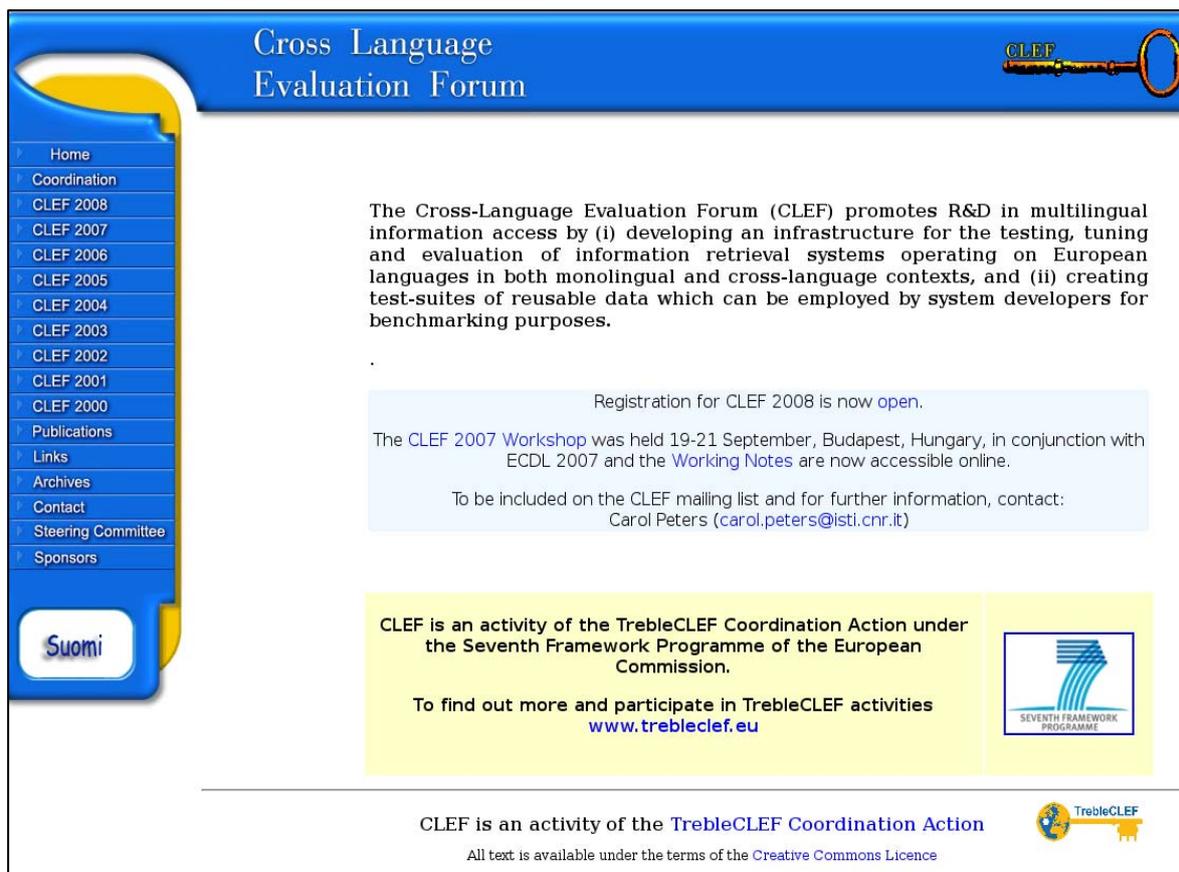


Figure 1: Treble-CLEF website home page

It can be viewed at <http://www.trebleclef.eu/>. The website is divided into public and password protected areas. The public area provides information on TrebleCLEF research directions, events and publications

together with News on activities of interest to the TrebleCLEF community. The password protected area of the website contains information of interest to members of the consortium: deliverables, information on management issues, meetings, phone conferences, minutes, etc.

A major issue that had to be considered when creating the website was the relationship between TrebleCLEF and CLEF. CLEF is a very well-known activity and the website is visited very frequently. However, CLEF is now an activity of TrebleCLEF and we felt that it was important to make this fact very evident on the CLEF website. The TrebleCLEF website has thus been designed to be compatible with the original colours (blue and yellow) and symbols (a key) adopted by the CLEF site – with a clear link to the CLEF site (www.clef-campaign.org). At the same time we have redesigned the CLEF site in order to encourage visitors to CLEF to visit the TrebleCLEF site and learn about our activities. However, it is our intention to extend the current website in the near future with additional functionality so that it becomes a real portal offering access to a number of resources and services.



The screenshot shows the homepage of the Cross Language Evaluation Forum (CLEF). The header is blue with the text "Cross Language Evaluation Forum" and a key icon. A navigation menu on the left lists various categories like Home, Coordination, CLEF 2008-2000, Publications, Links, Archives, Contact, Steering Committee, and Sponsors. The main content area features a paragraph about the forum's mission, a registration notice for CLEF 2008, and information about the CLEF 2007 Workshop. A yellow box highlights that CLEF is an activity of the TrebleCLEF Coordination Action under the Seventh Framework Programme of the European Commission, with a link to www.trebleclef.eu. The footer includes the TrebleCLEF logo and a Creative Commons license notice.

Figure 2. CLEFwebsite homepage

Through this portal, the public documents and all publications produced within the project will be made accessible, along with links to specific sites and documentation about the relevant technological issues. An RSS feed will be added so that registered members can be kept up-to-date with TrebleCLEF activities and we are evaluating the possibility of producing an electronic Newsletter. CELCT will eventually be responsible for the newsletter but all partners will contribute.

CLEF already provides an area on the existing website dedicated to information on language processing tools and resources for MLIA system development; this area will be considerably expanded and reinforced in **Treble-CLEF** in order to provide anyone who wants to build a MLIA system with both information on how to proceed and also on how to identify and find the various building blocks needed to build a particular system.

As **Treble-CLEF** is primarily aimed at the evaluation of scientific approaches underlying technological developments in MLIA, a special effort will have to be made to provide transparent information on the nature of the project activities to a non-specialised audience. A concept for this is now being developed from the start of the project and will be updated as the work proceeds. A part of the project website will be dedicated to the presentation of the project work to the broad public and an information slide-show will be developed for presentation at any information events organised by the Commission or other European institutions.

3.2 Production of Promotional Material

The following actions are planned (the website will be a major means of distribution for many of them):

- Design and printing of posters and project flyers to be distributed at conferences and other international events
- Employment of the project logo on all the project related resources including the project web pages, flyers, deliverables, and many other relevant materials and publications in order to make the brand as widely known as possible
- Regular issue of press release when events are organised; dissemination also via the ELRA newsletter and ERCIM News – the bulletin of the European research Consortium for Informatics and Mathematics
- Production and publication of reports on research activities, and guidelines for best practices
- Production and publication of TrebleCLEF tutorial slides
- Production of Videos of project-related activities, eg tutorials, lessons, etc., accessible from the website and also included in promotional DVDs
- Production of slides for training and dissemination material, accessible from the website.

3.3 Organisation of Events

A number of workshops will be organised as part of the activities of TrebleCLEF. The scientific Committee for all workshops will include one member from each project partner. CELCT will be responsible for the logistic organisation of all workshops. In all we plan to hold 5 workshops:

- two annual CLEF workshops in which the participants in the campaign meet to present their experiments and discuss their results; these workshop will be held in conjunction with the European Conference for Digital Libraries in order to maximise the possibility of dissemination of the results to the DL community; the main CLEF workshop is held over 2.5 days but normally there are some additional preliminary one day meetings between smaller groups of participants; working notes are produced for dissemination at the workshop and on the TrebleCLEF portal; the post-workshop proceedings consisting of fully revised and expanded papers will be published by Springer in the LNCS series; the workshops are organised as part of Workpackage 2.
- 3 brain-storming workshops. The first two will be organised as part of Workpackage 3 – one will bring together system developers to study system-oriented aspects, while the second will bring together MLIA user communities to discuss user requirements and user-oriented evaluation of multilingual information access systems; the third will bring together researchers to share their experiences in using log analysis and will be organised as part of Workpackage 4.
- a “MLIA Technology Day” aimed at putting researchers, software developers and users together in order to facilitate a two-way communication of research advances and results researchers, as well as the reverse flow of requirements and ideas from users and developers; this will be directly organised by Workpackage 6.

The workshops listed above are specified in the Technical Annex and specific funding has been

assigned for their organisation.

However, when possible and when opportunities arise (and depending on budget) the Consortium will also organise other events. In this respect, TrebleCLEF has already organised a workshop to be held at the European Conference for Information Retrieval- ECIR '08 Glasgow, Scotland 30 March 2008: “Workshop on Novel Methodologies for Evaluation in Information Retrieval”, see <http://ecir2008.dcs.gla.ac.uk/>. A proposal for a TrebleCLEF panel on evaluation at the European Summer School in Information Retrieval (**ESSIR**) in 2009 has also been accepted

3.4 Training Activities

It is the intention of **Treble-CLEF** to disseminate the results of the Best Practice studies resulting from the activities of WP 3 (Best Practices for system-oriented and user-oriented MLIA), WP 4 (Best Practices for Test Collection Creation) and WP 5 (Best Practices for MLIA Language resources) through training activities.

These will consist in the preparation of a set of tutorials to be proposed in the most relevant conferences in the areas covered by MLIA R&D :

- Annual Meetings of the Association for Computational Linguistics (ACL), including the European Chapter (EACL) and the North American Chapter (NAACL)
- Information Retrieval Conferences such as SIGIR and the European Conference on Information Retrieval (ECIR)
- Digital Library conferences such as Joint Conference on Digital Libraries (JCDL) and European Conference on Research and Advanced Technologies on Digital Libraries (ECDL).
- International Conference on Language Resources and Evaluation (LREC)
- Library Conferences such as IFLA (International Federation of Library Associations) and CENL (Conference of European librarians)

The tutorials will also be publicised on the project portal and will be made available on demand to organisations requesting them.

A summer school will be organised on MLIA theory and practice (as part of Workpackage 6). The school will be structured in a series of lectures and training sessions. The program will cover the following areas: Multilingual Text Processing (covering problems of language specific tokenization, indexing, stemming); Cross-Language Information Retrieval (approaches and technologies used for CLIR); Multilingual Information Retrieval (problems involved in querying and retrieving, merging and presenting results from a multilingual collection); Resources for MLIA (information on language processing tools and linguistic resources); Multilingual User Requirements (what the user wants and how to provide it); Multilingual System Architectures; Evaluation for Multilingual Systems and Components. Prospective students will be from both industry and academic institutions. The School fees will be subsidized by the Treble-CLEF funding and a number of grants will be made available for young researchers in need. The Steering Committee of the Summer School is composed of representatives of the Consortium and of the Advisory Board: Carol Peters (Coordinator) CNR, Italy; Maristella Agosti, University of Padua, Italy; Martin Braschler, Zurich University of Applied Sciences, Switzerland; Amedeo Cappelli, CELCT, Italy; Khalid Choukri, ELRA, France; Christian Fluhr, French Atomic Energy Agency, France; Stefan Gradmann, University of Hamburg, Germany; Donna Harman, NIST, USA; Noriko Kando, National Institute of Informatics, Tokyo, Japan; Mark Sanderson, University of Sheffield, UK; Felisa Verdejo, UNED, Spain; Clare Warwick, University College London, UK.

A problem noted is that the European Summer School in Information Retrieval, held every two years, will be held in September 2009. It is important that the dates and programme of the TrebleCLEF summer school do not overlap or clash with ESSIR. The SC is currently considering possible dates and venues. It is very probable that it will be decided to hold the Summer School in Corfu in late September, immediately following ECDL 2009. A first preliminary plan should be ready by end April.

Table 1: Preliminary Table of Planned Activities (to be updated throughout the project lifetime)

Event	Place	Date	Resp.	Status
EPOCH Final Event Presentation “From CLEF to TrebleCLEF”	Rome, Italy	25-26 February 2008	C. Peters	Done
EDLNet Work Group 2 meeting Discussion of Europeana Functional Specifications	National Library of The Netherlands, The Hague	17-18 March 2008	C. Peters	Done
TrebleCLEF Workshop on Innovative Evaluation Methodologies at ECIR 2008	Glasgow, UK	30.03.08	M.Sanderson; N.Ferro; M.Braschler; J.Gonzalo	Finalized
MLIA User Communities Workshop	Segovia, Spain (tbc)	24-25.06.08 (tbc)	J.Gonzalo	Under planning
CLEF 2008 Workshop	Aarhus, Denmark	17-19.09.08	C. Peters	Confirmed
System developer’s workshop	?	?	M.Braschler	Under planning
Query log Analysis Workshop	?	?	M.Sanderson	Under planning
TrebleCLEF Panel on Evaluation at ESSIR Summer School	Padua, Italy	June, 2009	N.Ferro	Confirmed Details to be decided
CLEF 2009 Workshop	Corfu (tbc)	16-18.09.09	C. Peters	Confirmed
TrebleCLEF Summer School	Corfu (tbc)	21-25.09.09 (tbc)	C.Peters	To be decided
MLIA Technology Day	?	Autumn 2009	A.Cappelli	To be decided

3.5 Participation in Standardization Activities

The organisation of open system evaluation campaigns and the ongoing collaboration between the CLEF, TREC and NTCIR evaluation initiatives will lead to the creation of de-facto standard interfaces and data formats, with respect to the creation of test collections and evaluation methodologies and metrics. These conventions will facilitate future research efforts and commercial applications.

We will examine the feasibility of stipulating conventions for widely distributed resources that may facilitate the development of corresponding standards will be stipulated for the following areas:

- for dissemination of tools and resources,
- for test data and evaluation metrics, including reference implementation of automatic evaluation metrics and guidelines for human evaluations.
- for organising evaluations of MLIA systems, including protocols for releasing training and test data, evaluation time, and dissemination of test results

3.6 Contacts and Collaborations

TrebleCLEF plans to establish contacts and collaborations with other projects and activities in related areas. Currently contacts have been established with the following projects/groups:

EDLNet (www.edlnet.eu): EDLNet is currently building consensus and finding solutions towards the creation of the European Digital Library. TrebleCLEF is an official member of EDLNet and advises on multilingual issues.

The European Library (www.theeuropeanlibrary.org): CLEF is currently organizing an evaluation track in collaboration with TEL and using TEL data with the goal of identifying best practices for monolingual and cross-language information retrieval on library catalog data.

Minerva (www.minervaeurope.org): MINERVA is a network of Member States Ministries to discuss, correlate and harmonise activities carried out in digitisation of cultural and scientific content for creating an agreed European common platform, recommendations and guidelines about digitisation, metadata, long-term accessibility and preservation. C. Peters has been requested to give a presentation of TrebleCLEF at a Minerva organised event in Slovenia, 5-6 June

Information Retrieval Facility (www.ir-facility.org): The Information Retrieval Facility (IRF) is a platform for experimentation, networking, and collaboration for professionals who are depending on text mining of extremely large document corpuses. TrebleCLEF and IRF are currently discussing plans for collaborative actions.

CLARIN (<http://www.mpi.nl/clarin/>): The mission of CLARIN is very compatible with that of TrebleCLEF. CLARIN is committed to establishing an integrated and interoperable research infrastructure of language resources and its technology. Representatives of CLARIN have been invited to the next management meeting of TrebleCLEF, to be held in Spain in June, in order to discuss possible collaborative activities

3.7 Technology Transfer and Exploitation

Technology transfer and exploitation of the results of the CLEF evaluation activities and of the knowhow built up by participants in CLEF with respect to the implementation of MLIA functionality is one of the main objectives of **Treble-CLEF**. In the above sections we have described how we will handle the question of technology transfer. The dissemination plan will further specify these.

A major activity will be the organisation of a “MLIA technology day” aimed at putting together researchers developers and users in order to facilitate the transfer of results from research to the market. In order to ensure that this is a success much preparatory work will be needed. For this reason, we will actively promote dialogue with prominent representatives from industry, as well as companies and large organisations with strong MLIA needs. To this end, we will create a User Advisory Board for the organisation of the “Technology Day” inviting representatives from industry and user communities. Members of the board will be involved in the organisation and in the participation of the event. The consortium expects that the cooperation with the User Advisory Board will lead to detailed evidence concerning the compatibility between various levels of MLIA quality and the requirements of real-world application scenarios.

The evaluation methodologies and tools that will be developed during the course of the **Treble-CLEF** project will be exploited in at least two ways. They will be offered either as part of a service or as part of a collaboration with several groups of interest including a) software companies which produce MLIA, b) users who require methods and technologies to improve their products and/or their internal workflow in order to increase profitability or efficiency.

An important activity of WorkPackage 6 will be the preparation of the Exploitation Plan which will assess the results of the **Treble-CLEF** activities and will propose solutions for sustainable follow-up activities.