Project annual public report

Project acronym: MORMED

Project title: Multilingual Organic Information Management in the Medical Domain

Grant Agreement number: 250534

Project website address: http://www.mormed.eu

Coordinator's Contact Details

Project coordinator: Dr. Iraklis Paraskakis, Senior Research Officer, South East European Research Centre

Tel: +302310253478

Fax: +302310234205

E-mail: iparaskakis@seerc.org

Project Overview

The World Wide Web is used nowadays as the main and preferred repository where people search, access and publish information on any topic. Nevertheless, information for restricted and highly focused communities (e.g., communities of rare diseases) is not easily available and not always easy to find. Such diseases are not widespread, although they appear all over the world, and information resources are dispersed and difficult to find. Relevant information exists but is available only to specific groups of people or to specific countries. Language barriers further impede the dissemination of interesting information and the exchange of experiences between the involved parties. Thus communication channels between the various stakeholders, e.g. researchers and general practitioners (GPs), GPs and patients, are cumbersome or even non-existent due to this two-dimensional language barrier of national languages and specific terminology.

People concerned about, suffering from or working on such rare diseases, should benefit from recent technological developments in order to have access to information of interest, to publish their experiences and socialise. The intuitiveness and usability of Web 2.0 applications, further enhanced with semantic capabilities, and their combination with just-in-time automated translation services reassure that anyone could participate in a thematically focused online community regardless of the educational and cultural background and the national language.

Towards this direction a platform for multilingual thematic communities is proposed, which will be offered by a specific service provider to a targeted audience involved with the Lupus disease and Hughes syndrome. This platform will support efficient online collaboration for content publishing and social networking in a multilingual environment. Semantically enriched Web 2.0 tools are combined with comprehensive automated translation and text summarisation tools to form the MORMED platform. MORMED will not only facilitate the workflow of publishing content by end users (e.g. researchers, medical doctors, GPs, patients), but also the workflow of managing the platform itself by administrators, and handling the translation requests by translation experts.

Goals and Objectives

The main goal of the MORMED project is to establish and offer a multilingual thematic community platform for rare diseases as a service to groups of interested users.

The main objectives of MORMED are to:

- Integrate and deploy a multilingual thematic community platform for rare diseases which will promote information exchange and social networking.

- Make existing content accessible via this platform and stimulate new, user generated content.

- Use existing linguistic resources to improve the automated translation quality.

- Test and assess the platform by evaluating the experiences of users involved in the rare diseases domain, using various languages (e.g. English, Spanish, German and Hungarian).

The applicability of this service is reinforced by several supplementary objectives in order to:

- Align relevant corpora collected from users, from EMEA[1] and possibly ELRA[2], in order to achieve high translation quality and to provide a broad linguistic basis for future deployment in the medical domain.

---

[1] European Medicines Agency

[2] European Language Resources Association

- Prepare the commercial service after the end of the project by deciding on the optimal business model.

- Distribute the information about the service.

- Motivate an extended user community to take part in the service test during the project.

- Collaborate with relevant similar initiatives, in particular with the Thematic Network that will be funded under objective 5.2 of the ICT/PSP work programme.

The project will integrate results of the on-going OrganiK[3] project, with backend translation tools provided by the SME partner Language Technology Centre (LTC) and with auxiliary tools (such as summarization tools). Most of the technology is available as open source software. Technically, the project will combine a semantic Web 2.0 knowledge management platform with advanced automatic translation tools in order to support information exchange and to bridge language barriers. System architecture and integration aim at providing an open platform, based on open standards that is extensible and scalable, and applicable to different market areas.

The system to be developed will consist of the following main components:

1. The OrganiK platform

2. The multilingual production system operating in the background: LTC Communicator, further extended with text summarization tools (open source)

The OrganiK platform is a next generation knowledge management system that manages and promotes social structures. OrganiK combines social software applications and semantic technology. This includes social software applications (i.e. Wiki, Blog and Microblogging work spaces, collaborative bookmarking application, search component), enhanced with semantic technologies (i.e. semantic search capability, recommender system, content and user behaviour analyser, collaborative filtering engine, etc). The translation tools are based on LTC Communicator, a workflow system incorporating translation tools and machine translation.

MORMED is envisaged to be used by various types of users:

1. Researchers collaborating in research projects. They could reach and communicate with GPs and patients; researchers could post recruitment information for clinical trials and thus make this available to GPs for dealing with patients.

2. Medical doctors and General Practitioners (GPs) requiring information on rare diseases (Lupus disease and Hughes Syndrome) and accessing international information resources.

3. Patients, also represented by patient support groups, like family and friends, requiring and exchanging information on international level. Patients can

---

[3] Seventh Framework Programme, Research for the Benefit of SMEs, Grant agreement no.: 2222225

contribute their experiences which can be used as case studies to further fuel research activities and improve common practices.

4. Administrators of the MORMED platform, responsible for general maintenance tasks

5. Translation experts, such as translators, post-editors, terminologists and language technology experts, responsible for accessing, updating and populating the automated translation service component.

The envisaged architecture of MORMED, which will result from the integration of the OrganiK platform and LTC Communicator, is illustrated in the figure below.
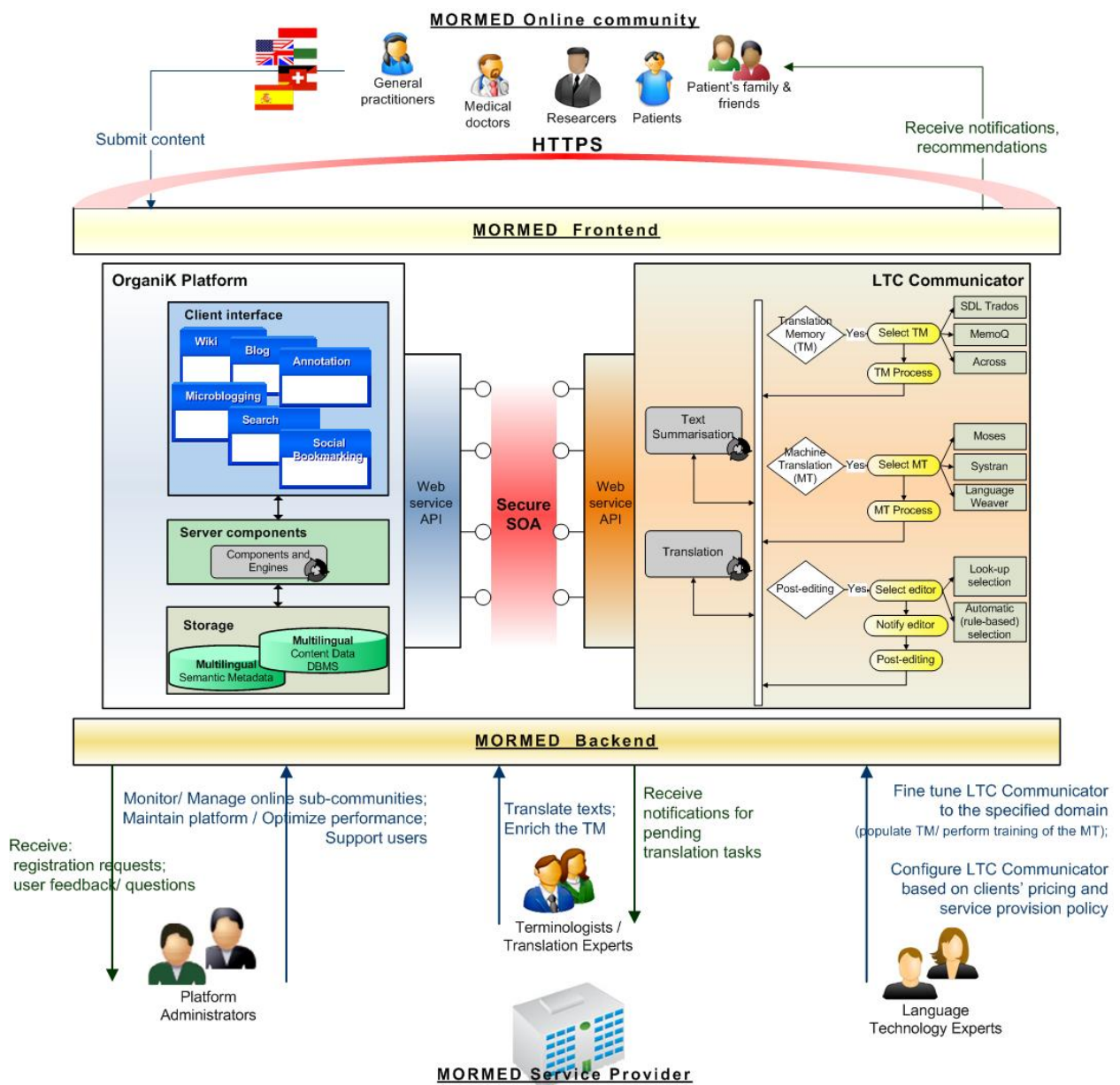
Figure: MORMED Architecture

Innovation

Current translation services are merely document-focused, failing to engage teams of end-users, thus creating virtual communities and helping them to surpass the language barriers. Traditional translation tools do not integrate community networking initiatives. Additionally, current translation tools and associated workflows typically do not engage the 'end-user' into the process and are solely focused on the translation back-end views. Finally, the marketing costs of promoting such an innovative translation offering can be prohibitive for small translation service providers.

MORMED will address the above limitations by offering the following innovations:

User experience: MORMED will provide the opportunity to individuals who belong to specific communities of practice (in our pilot case, the Lupus community (Researchers, Doctors, Patients)) to benefit from translated informal information and ad-hoc knowledge exchange. Instead of focusing on translating formal documents, MORMED platform will act as a central translation hub for communities formed by various individuals connected by a common interest (in our pilot case the Lupus rare disease).

Business model: Instead of running a translation service based on billing each document that needs to be translated, the business model of MORMED will follow the paradigm of various Web 2.0 social networking sites and rely on micro-payments of each individual user, based on a Freemium business model. This will work by offering basic services for free, while charging a premium for advanced or special translation features to each individual user.

Back-end technology: The back-end technology of MORMED will support semantically enhanced corpora based on domain-specific taxonomies and ontologies (in our pilot the Lupus disease) and streamlined translation workflows (based on LTC Communicator functionalities).

Benefits

The major benefit is the accessibility to information in the specific domain of rare diseases, such as Lupus and Antiphospholipid Syndrome (Hughes Syndrome), regardless of the national language and education level. Although researchers have the possibility to access their colleagues' research work, there are two main barriers that will be overcome by MORMED: the different language barriers and the lack of social networks in ones own country as well as across Europe. MORMED platform will thus establish communication flows and promote active discussions and exchange of ideas between them, thus fuelling further research initiatives. Medical experts and Doctors will be able to be informed for promising research results and best practices. Patients and patient support groups can access the information provided, or seek information in the community, and, moreover, benefit from experiences shared by other patients either in their country or in countries across Europe.

Furthermore, the major European policy relevant to MORMED is i2010 - A European Information Society for growth and employment, which is the EU policy framework for the information society and media. It promotes the positive contribution that information and communication technologies (ICT) can make to the economy, society and personal quality of life. MORMED aspires to have an impact on the personal quality of life of people suffering from rare diseases, and hence also on the society as a whole. If successful, the service could be applied to or extended to other areas to increase this impact. Moreover, the first objective of i2010 is to establish a Single European Information Space offering affordable and secure high-bandwidth communications, rich and diverse content and digital services. MORMED is oriented towards multilingual content and digital services, combining information in a specific sector and making it available to all users, overcoming language barriers that currently prevent interactive information exchange. This creates a clear relation to this i2010 objective as the project results in a true single information space in the area targeted.

Work done and achievements

Within the first nine months (as the November 15th is the cut-off date for this report) of the MORMED project the first milestone M1 has been completed, which refers to the creation of user requirements analysis in work package 2 (User Requirements Analysis). User requirements analysis took place by surveying on the profile of various stakeholders and on their information exchange and sharing needs. Results have been collected and analysed, usage scenarios have been crystallised, and solution objectives have been documented in Deliverable D2.1.

During WP2 the consortium identified and defined the user requirements for the MORMED platform. That was a crucial step since it will guide and shape the choices for designing the features, functionalities and overall architecture of the platform. The procedure towards defining the user requirements involved an extensive data collection period, through semi-structured and unstructured group discussions and provided feedback that involved all consortium partners. The results of the analysis led to the definition of seven use scenarios of the MORMED platform, the formalisation of user requirements and the design of the components comprising the conceptual architecture of the platform.

Furthermore, work has been commenced on work package 4 (Preparation of Multilingual Resources and Translation Technology). Towards the realization of the second milestone the consortium has been collecting, validating, and where necessary manually enriching the multilingual content that is essential for the training and contextualisation of the underlying machine translation technology to the domain of the MORMED project.

The purpose of this work package is threefold. Firstly, to collect, validate, and where necessary manually enrich the multilingual content that is essential for the training and contextualisation of the underlying machine translation technology to the domain of the MORMED project (i.e. the domain of medicine and rare diseases). Secondly, to

evaluate the performance of different machine translation technologies against the collected and pre-processed content (rule-based and statistics-based MT systems). Thirdly, to carry out the actual preparation and population of translation memories for each language combination supported by MORMED, finalise the contextualisation of the MT systems which were selected following the evaluation with the collected content, augment the translation tooling with additional custom dictionaries and relevant bilingual material as needed, and continue repeating these tasks as appropriate throughout the evaluation of the MORMED platform by the end-users.

The first aim of the work package is at a completion stage. The second and third aims of the work package are on their way towards realisation.

Concerning work package 1 (Project Management), the project coordinator (SEERC) is enacting its role by focusing more on the administrative and financial management of the project, while the technical coordinator LTC is guiding the technological development activities.

Finally, the all of the partners of the MORMED project have been active on commencing exploitation and dissemination activities within work package 6 (Exploitation and Dissemination). Towards this end project partners have created the project's public web site (http://www.mormed.eu/) which contains information about the objectives of the project, partners profiles, generic approach, technologies used, contact details, news items, etc.

An indicative and non-exhaustive list of dissemination activities is the following:

- MEDICON 2010 12th Mediterranean Conference on Medical and Biological Engineering and Computing (27 - 30 May 2010, Chalkidiki, Greece), where SEERC has presented a paper.

- World of Health IT Conference & Exhibition (WoHIT) (15-18 March 2010, Barcelona, Spain) and World Health Care Congress (19-20 May 2010, Brussels, Belgium), both attended by LTC.

- Reunión Nacional sobre Esclerodermia, 9-10 April 2010, Palma de Mallorca (Spain), attended by HCPB.

- XIII Congreso de la Sociedad Valenciana de Reumatología, 23-24 April 2010, Peñíscola (Spain), attended by HCPB.

- 7th International Congress on Autoimmunity 5-9 May 2010, Ljubljana (Slovenia), attended by KCL

- 6th Annual World Health Care Congress Europe 2010 , 19-20 May 2010, Brussels (Belgium), attended by LTC.

- EULAR Conference 2010, 16-19 June 2010, Rome (Italy), attended by KCL and HCPB.

A more detailed list of activities is presented in the Dissemination Plan (Deliverable D 6.1) of the MORMED project.