This document is part of the Coordination and Support Action "Preparation and Launch of a Large-scale Action for Quality Translation Technology (QTLaunchPad)", funded by the 7th Framework Programme of the European Commission through the contract 296347.

Publishable Summary

Project Duration:	July 2012 – June 2014
Reporting Period:	July 2013 – June 2014
Coordinator:	Prof. Dr. Hans Uszkoreit

DFKI GmbH

Alt-Moabit 91c

10559 Berlin

Hans.Uszkoreit@dfki.de

Summary of project objectives

Quality translation is in higher demand today than ever before. European industry, administration, and society urgently need progress in translation technology to fill existing translation requirements, extend multilingual communication to additional languages and services, and to reduce costs and fulfill their commitment to linguistic diversity.

However, nearly all on-going machine translation (MT) research is dedicated to gradual improvements of MT for information “gisting”, i.e., getting a rough impression of the content of a text such as a web page. Despite considerable progress in this area, the quality barriers for outbound translations – i.e., translations to be published or distributed outside an organization – have not yet been tackled, much less overcome.

The QTLaunchPad project consortium is convinced that nothing less than a paradigm shift is needed to break out of the dead-end that MT research landscape is currently trapped in. This includes a new mode of cooperation between language service providers (LSPs) and MT R&D, who have enjoyed blissfully separate existences for far too long. Instead of only adjusting known algorithms to produce marginally better results, we need a novel, systematic approach to carry out MT research in Europe, an approach that addresses the goal of producing quality translations and that takes into account the needs and priorities of European MT and LSPs (see Figure 1 for an illustration of the idea).

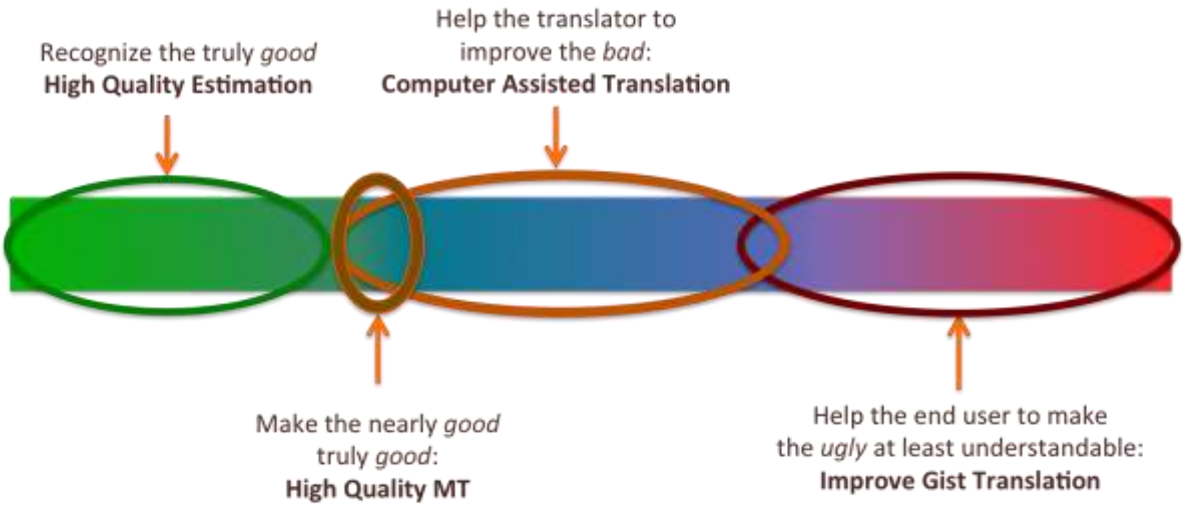
Recognize the truly *good*
High Quality Estimation
Help the translator to improve the *bad*:
Computer Assisted Translation
Make the nearly *good* truly *good*:
High Quality MT
Help the end user to make the *ugly* at least understandable:
Improve Gist Translation

Figure 1: Goals for Translation Technology Development

The QTLaunchPad support action has prepared the grounds for a new type of collaborative MT research dedicated to overcoming translation quality barriers. Central objectives have been to:

1. Assemble and provide needed **data and tools**, including specialized translation corpora, test suites, and tools for quality assessment;
2. Create shared **quality metrics for human and machine translation** and to improve automatic translation quality estimation;
3. Extend the META-SHARE **platform for resource-sharing** to the needs of quality MT research;
4. Define strategies and challenges and then **plan and launch a large-scale research and innovation action** for a breakthrough in quality translation technology.

The most important horizontal goal of the project has been to involve the different stakeholders (LSPs, translators, industries requesting translations, etc.) in the development of tools and methods and in the planning activities to assure usefulness and sustainability of its results. To this end, the project has cooperated with the Globalization and Localization Association (GALA) as subcontractor and the International Federation of Translators (FIT), who both have supported the project. In addition, numerous stakeholders such as TAUS and several LSPs and companies have provided feedback and invested effort for the project and in their own interest.

Main activities and results achieved

This section briefly mentions core results of the project. Pointers for more information on the QTLaunchPad web page are given.

A central result is the second version of the **Multidimensional Quality Metrics (MQM)**. In collaboration between QTLaunchPad and the MultilingualWeb-LT project, the first version was adopted as the basis for the *Localization Quality Issue* data category in ITS 2.0, which has been standardized as a World Wide Web Consortium (W3C) recommendation. MQM 2.0 builds on this version and maintains compatibility with it (as a superset with a defined mapping to ITS 2.0), and was developed in close collaboration with language service providers (LSPs), including close consultation with Welocalize, text&form, Enlaso, VistaTEC, Logrus, Multiling, and Linguatech, plus discussions at several workshops and webinars organised by QTLaunchPad together with GALA. MQM also received extensive feedback from the American Translators Association (ATA) and the standards committee of the International Federation of Translators (FIT). MQM 2.0 features include:

- 102 issue types in a hierarchy.
- A “core” of 21 common issue types.
- A flexible scoring mechanism that emulates the scoring system of the LISA QA Model 3.1 by default, but which can be easily adapted to any needed values
- A set of twelve “parameters” based on the ISO/TS-11669 specification, which help guide users in selecting task-relevant issues.

More information can be found on the QTLaunchPad web page.[1]

[1] http://www.qt21.eu/launchpad/content/multidimensional-quality-metrics

Supported by the project, MQM has been implemented in translate5[2], an open source client-server proofreading, editing, and data procurement environment. Early adopters of MQM include Mozilla, who state in a support letter:

> "We anticipate scaling the MQM to our general community (including all 90 languages), targeting translation quality assessment at our Firefox OS mobile product, which is currently being sold on smartphones in several European countries."

Using a subset of MQM, the project consortium has prepared several **error corpora**[3] that document the **barriers for high quality MT** found in near-miss translations together with **test suites** that can be used to quickly assess the performance of new MT engines. These corpora include filtering and searching options that enable users to explore patterns in the data and make selections of data relevant to specific questions. This work has been performed in very close cooperation with GALA and several Language Service Providers (LSPs) and has served as a very important experience showcasing the new, integrated mode of cooperation between MT research and LSPs.

Work done so far for **quality estimation baseline and extensions** included benchmarking existing systems for quality estimation, along with the development of a baseline software involving a large number of basic features used in previous work in the field, as well as new language and MT system independent features introduced as part of the project. An open source Java software for feature extraction was developed at USFD. The code is stored in a GitHub repository. USFD integrated functionalities for machine learning using the toolkit SK-learn[4] and an in-house implementation of a powerful learning algorithm for the task (Gaussian Processes). Altogether, these form the full pipeline necessary for quality estimation: from data pre-processing and feature extraction to feature selection, building and testing of models. The software is available for download at http://www.quest.dcs.shef.ac.uk/. It was used as the official baseline system for the WMT13 and WMT14 shared tasks on quality estimation (second and third edition)[5], organized by Lucia Specia and colleagues In addition, a number of papers have been published that describe the software toolkit and report experiments with it.

Working under our shared **communication and dissemination** plan, awareness of QTLaunchPad has been raised, gaining particularly good exposure at a number of events. This consists of: setting up social media channels and expanding their memberships (Twitter: 113; Facebook: 86; LinkedIn: 154; Mailing List: 371); preparing and disseminating printed materials relating to the workshops and other events/presentations; preparing content for dissemination via GALA showcases, blogs, and the GALAxy newsletter; and also for regular and targeted stakeholder publications (ATA Newsbriefs, Multilingual Magazine, CNGL outlets, etc.). These activities have strongly added to the momentum of the project and created new relationships with stakeholders while strengthening existing links with project partners. In its first year, QTLaunchPad, in addition to several talks given by consortium members at different conferences, has organized workshops and tutorials at the following events:

- MultilingualWeb[6], Rome, March 12-13th, 2013
- GALA Annual Conference[7] Miami, March 17-20th, 2013

[2] http://www.translate5.net/

[3] http://www.qt21.eu/deliverables/annotations/

[4] http://scikit-learn.org/

[5] http://www.statmt.org/wmt13/quality-estimation-task.html, http://www.statmt.org/wmt14/quality-estimation-task.html

[6] http://www.multilingualweb.eu/rome

- Localization World, London, June 11-13th, 2013
- GALA 2014, Istanbul, March 23-26th, 2014
- CNGL Scientific Expo, Dublin, May 13-14th, 2014
- LREC 2014, Reykjavik, May 26-31st, 2014
- EAMT Conference, Dubrovnik, June 15-18th, 2014

In order to **gather requirements from the QTLaunchPad stakeholder groups** and survey their needs and expectations in relation to quality translation, a questionnaire was developed with external input and distributed via GALA to survey their members' needs and interests. Data from 467 industry-based participants were collected from the questionnaire, which focused on translation quality assessment and technologies to tie in with the aims of QTLaunchPad, including MQM technologies, and the RIAS for QT21 (see below). Participants were also given contact details for the project and many indicated their interest in becoming involved with the project. This process was coordinated by DCU in collaboration with GALA and FIT, and was key to establish additional support for the project and for longer term cooperation. A paper entitled "A Survey of Machine Translation Competencies: Insights for Translation Technology Educators and Practitioners" used the data collected as part of this survey. This publication is of particular interest to freelance translators and translator trainers, especially to inform the design of translator training programmes with a special focus on translation technology. At the project end, this paper has been recommended for publication in the international peer-reviewed journal *Perspectives: Studies in Translatology* (Routledge, Taylor & Francis), pending minor changes that the authors are in the process of implementing. This forthcoming publication based on the survey testifies the importance of the data collected by the project for the wider community.

The project delivered of a collection of **training materials** to the public.[8] These include the following:

- Slides and audio from two webinars organized together with GALA covering a general description of the project, a description of the multidimensional quality metric and a brief description of the quality estimation metrics,
- Slides from talks on the project, the Multidimensional Quality Metrics, and machine translation evaluation and quality estimation given at the Localization World Conference,
- Slides and a document with a more comprehensive description of the multidimensional quality metric and listing of the issue types in the metric, with illustrative examples of translation quality scorecards,
- Findings on the Translation Technologies and Quality Assessment survey recently conducted with the support of GALA focusing on language service providers,
- Annotation guidelines for an error corpus composed of human and machine translations.

To cater for the resources dimension and provide all support mechanisms needed for **documentation, sharing, search and retrieval of** all **MT-related resources**, QTLaunchPad builds upon and extends the META-SHARE infrastructure.[9] A QTLaunchPad-dedicated META-SHARE node/repository[10] has been set up and populated with MT-related language resources and/or their metadata-based descriptions of resource. For the acquisition of mono-

[7] http://www.gala-global.org/conference/

[8] These materials can be read and/or download from http://www.qt21.eu/launchpad/content/training.

[9] http://www.meta-share.eu, http://www.meta-share.org

[10] http://qt21.metashare.ilsp.gr

lingual corpora, the project's first main activity concerned the addition of new modules and/or modifications of existing ones with the purpose of providing larger and qualitatively better, in-domain corpora and of enriching the metadata we store along these documents. In more detail, a revised Focused Crawler (FC) searches in each relevant HTML document for out-links pointing to Creative Commons (CC) licenses. Also, a confidence score on the domain-relatedness is estimated by comparing the relevance score of a web page with its length, while taking into account the restrictions provided by the user through the crawler's configuration file. Finally, due to the fact that large amounts of domain-specific data are often available in PDF format, we developed a module for text extraction from PDF documents based on the open-source PDFBox java library. In the first project year, we have also performed a set of initial experiments in acquiring domain-specific parallel data from multilingual web sites.

One central objective of QTLaunchPad was the **preparation of a big research and innovation action (***QT21 - Quality Translation Technology for the 21st Century - A European Initiative*) that comprises:

- a new research paradigm based on the analytical investigation and elimination of translation quality barriers,
- an unprecedented collaboration of the best MT research actors in Europe,
- a broad and sophisticated infrastructure of resources (data and tools),
- a close integration of human translation professionals into the research process,
- a close cooperation with large translation customers,
- the wealth of results of computational linguistics research.

Planning of the big European quality translation initiative QT21 was performed by the QT21 Planning Panel, a board consisting of stakeholders who would qualify as potential participants. The planning process, which was driven and coordinated by the QTLaunchPad Extended Steering Board, consisted of several phases. The face-to-face meetings of the Planning Panel took place:

- September 13/14, 2012 at DFKI in Berlin,
- October 25, 2012 at Hotel Crowne Plaza in Wiesbaden (co-located with tekom/TCworld fairs),
- March 13, 2013 at FAO Headquarters in Rome (co-located with Multilingual Web Workshop),
- February 6, 2014 at Schloss Dagstuhl (co-located with a Seminar on Translation into Morphologically Rich Languages).

At the meetings, large space was given to the presentation and discussion of use cases and the respective research strategies needed. Several smaller sub-group meetings have worked out a master plan including a leading-edge research plan and the roadmap for the field of MT. The roadmap for MT developed within QTLaunchPad is a continuation and concretization of the roadmap for the whole field of language technologies as defined in the META-NET Strategic Research Agenda for LT in Europe presented publicly in January 2013.[11] It will be taken over by the QTLeap project after the end of QTLaunchPad. Several project proposals have been submitted to the EC's ICT-17-2014 call by members of the QTLaunchPad Planning Panel.

[11] http://www.meta-net.eu/sra-en

At the end of the project, a Quality Translation Shared Task (QTST)[12] was organized and run as part of the 9th Workshop on Statistical Machine Translation (WMT) on 26th-27th June 2014. To maximise its international visibility in the community (not only within Europe), the WMT 2014 workshop was strategically co-located with the 52nd Annual Meeting of the Association for Computational Linguistics (ACL 2014), one of the most prestigious and well-established forums for research in computational linguistics, held in Baltimore. The objective of the QTST was to road test the QTLaunchPad MT Infrastructure, resources, and approach to HQ translation, as well as to achieve close engagement of the stakeholder community on a technical level.

[12] http://www.statmt.org/wmt14/quality-estimation-task.html