# TAUS Quality Dashboard

An Industry-Shared Platform for Quality Evaluation and Business Intelligence

September, 2015


TAUS
Enabling Better Translation

This document describes how the TAUS Dynamic Quality Framework (DQF) generates a Quality Dashboard serving the interests of all stakeholders in the global translation industry. Through an easy-to-use plug-in translators and managers share data and reports that give them valuable statistics, benchmarking and analytics.

The Quality Dashboard is developed for and with the TAUS members and users.

# Table of Content

# 1. Background

## Old school: 'one size fits all'

Translation quality evaluation is problematic. In 2011 TAUS conducted a survey among its members. We found that despite very detailed and strict error-based evaluation models the satisfaction levels with both translation quality and the evaluation process itself are very low. QE models are static, that is, there is a 'one size fits all' approach. Little consideration is given to multiple variables such as content type, communications function, end user requirements, context, perishability, or mode of translation generation (whether the translation is created by a qualified human translator, unqualified volunteer, a machine translation system or a combination of these.)

## Dynamic Quality Framework

In November 2011 TAUS published the foundational report for the Dynamic Quality Framework (DQF). (Download this report.) The solution proposed in this report was that a dynamic evaluation model would be developed that takes into account the changing landscape of diversification in content types and the adoption of automated translation technologies.

## Content Profiling

In collaboration with members and researchers from Dublin City University TAUS designed a model for profiling content types based on three vectors: utility, timeliness and sentiment ("UTS"). The UTS scoring started to give users guidance on methods for translation quality evaluation.

## DQF Knowledge Base

In 2012 TAUS started developing a knowledge base for everyone working in the global translation industry documenting different approaches to translation quality evaluation. The models described in the knowledge base are: adequacy (accuracy), fluency (readability), usability, community feedback, error typology, productivity measurement. The knowledge base contains use cases, templates, metrics and specifications. In consultation with users and members TAUS developed best practices for post-editing, sampling and different quality evaluation techniques that are all part of the knowledge base.

## DQF Error Typology

A vast majority of the providers and buyers of translation services manage their quality program with an error typology template. The LISA QA model which was developed in the nineteen-eighties forms the basis of the error categories being applied in most cases. The SAE J2450 is another metric that is well-known in automotive translations. TAUS worked with Dr. Sharon O'Brien to develop a more up-to-date version of these error typologies and made it available under DQF. In 2014 DFKI published the MQM (Multidimensional Quality Metrics). Under the European funded project QT21 TAUS and DFKI have harmonized the DQF and MQM error typologies.

## DQF Tools

On request of the members TAUS developed various tools for quality evaluation and made them available through the web site. Since early 2014 users can go to the TAUS web site to do the following types of quality evaluation: adequacy, fluency, error review, productivity measurement, MT ranking and comparison.

## DQF Users and User Group

An active user group has been formed around the DQF tools, allowing the TAUS team to constantly improve and fine-tune the tools. The DQF tools were designed for incidental quality evaluations, for instance for tracking and benchmarking of translators or translation vendors from time to time, or for the comparison of the quality of MT engines. Among the users the wish came up to use DFQ more frequently and to

make it part of day-to-day translation production control. Some companies used the DQF error typology, templates and metrics and implemented these in their processes.

## DQF as industry reference

Quality evaluation is one of the most pressing themes in the translation industry. Operators realize that, unless they start relying on common metrics and using the same methods and tools, they run the risk of comparing apples with oranges and taking uninformed decisions. DQF is more and more referenced as an industry standard and best practice for translation quality evaluation.

## DQF training program

In 2013 the Localization Institute started offering a quality management training program featuring DQF. In cooperation with TAUS Willem Stoeller has now developed a much improved and more focused DQF training program which is offered online through the TAUS web site. Through this course industry professionals will be trained in the DQF methods and best practices ensuring a consistent usage of the evaluation tools.

## Free access to DQF for students

Since December 2014 TAUS offers translation students at universities around the world access to all TAUS Data and resources, including the DQF knowledge base and tools through the Free Academic Membership Program.

## Further reading and videos

Check out the TAUS web site for various reports about quality evaluation and DQF use cases presented at TAUS QE Summits, hosted at Microsoft, Adobe and Localization World Conference in 2013 and 2014. See also the DQF API specifications and the list of translation tools that have already integrated DQF.

# 2. TAUS Industry Program

TAUS is a think tank and platform for industry-shared services and resources for the global translation sector. The TAUS industry program is focused in three main areas:




The TAUS Academy is a knowledge center where users can find training courses for post-editing and quality management with DQF, a library of reports and articles on translation automation and quality evaluation, industry best practices, use cases, technology briefings and webinars. The TAUS Academy also contains directories of translation technologies, certified post-editors and profiles of member organizations.

The Data Cloud is the largest industry-shared repository of translation data, containing more than 60 Billion words in 2,200 language pairs. Users can download and pool data for the training and customization of translation technologies. Members can choose to share data publicly or limit the sharing in a Private Vault. TAUS provides an API allowing the integration of the Search and Data Pooling functions in other translation work environments.

The Quality Dashboard is an industry platform for statistics on translation, benchmarking translation activity and quality and analysis of translation performance and production. TAUS provides API specifications allowing translation technology providers and users of translation services to plug TAUS DQF into their work environment.

This document is about the TAUS Quality Dashboard. Translators, project and vendor managers as well as buyers of translation who are interested in using the TAUS Quality Dashboard find all the relevant information in this document. Developers, interested in using the plug-in to integrate their technology with the TAUS quality Dashboard can find the API specifications on the TAUS web site.

# 3. DQF Vision

## Industry-shared metrics

Translation quality is a pressing theme in the translation industry. The diversification in content types and rapid adoption of translation technologies (including machine translation) drives the need for more dynamic and reliable methods and measurements of translation quality evaluation. The vision behind DQF is to standardize the methods and tools of quality evaluation and to aggregate the scores and measurements and make these available through industry-shared metrics.

## QE as business intelligence

Quality evaluation is costly and troublesome in many translation environments today. Measurements are often subjective and anecdotal. Industry-shared metrics will lead to more reliable measurements that give translation operators and producers (translators) useful benchmarks and insights that help them to adjust and improve processes. Today, for most companies quality evaluation is a necessity without remedy. The industry-shared metrics will turn quality evaluation into business intelligence steering and supporting management decisions.

## Industry benefits

DQF is an industry collaborative program that elevates the discussion about quality to an industry level. The resulting benefits on an industry scale are:

1. Increased credibility of the quality assurance function and process throughout the industry. This will improve the image of the industry in general and attract more customers, allowing the translation sector to grow faster.
2. Lower cost in general for all translation operators as a result of the shared investments in resources and tools.
3. Increased satisfaction with both vendors and clients by referring to standard metrics.

## Business benefits

Users of DQF can benefit from this industry collaborative program in multiple ways:

1. Sharing knowledge and standard references. Stepping away from the rigid outdated error typology review as the only way to measure translation quality, each individual translation operator does not have to research new methods on his own. Referring to common methods and best practices avoids confusion and disputes and ensures consistency.
2. Reduced costs. Sharing resources, tools and knowledge allows each individual stakeholder to reduce the costs of overhead in quality research and tool development.
3. Increased efficiency. Agreed best practices and shared tools and metrics reduces friction in each client-vendor relationship and makes it easier to collaborate in a multi-vendor environment.
4. Increased flexibility. Industry consensus and benchmarking on quality evaluation methods and scores allow for a more flexible use of often more efficient and less costly methods of quality evaluation.

# 4. What is the Quality Dashboard?

The Quality Dashboard delivers on the DQF vision described here above. It is an industry collaborative platform for the global translation services sector, helping all stakeholders – *translation buyers and providers, technology developers and translators* – to get deeper insights in the processes and the technologies. At the basis the Quality Dashboard will provide productivity and efficiency metrics across content types, industries, processes used, technologies applied and by language pairs.

Productivity is the throughput or speed expressed in the number of words per hour. Productivity is currently the most common way of measuring translation performance. Productivity is an absolute score. Efficiency is a new score introduced by TAUS. Efficiency is calculated based on productivity and the number of edits a translator makes per hour. The TAUS efficiency score is dynamically normalized against the highest and lowest numbers of words and the highest and lowest number of edits per hour, aggregated in the database.

While the productivity score is a good first performance indicator, we believe that the TAUS efficiency score gives both translators and managers a more reliable measurement, especially when used in combination with the filters for technology, process and content. See chapter 7 for a full description of the TAUS efficiency score.

Here is a mock-up of one of the reports on the Quality Dashboard:


The user can select one of the metrics – Productivity, Efficiency, Adequacy, Fluency – and select attributes from the filters – Languages, Time, Technology, Process, Content, Industry, Project, Translator/vendor, Customer. Translators and managers can also track statistics of their work and projects and the distribution of the segments (from TM+Human, TM+MT+Human+Post-edit, MT+Post-edit, TM only, MT only).

The Quality Dashboard is developed for and with the members and subscribers. QE Summits are organized twice a year, hosted by TAUS member organizations, to bring users together and agree on best practices and new features and support for the Quality Dashboard. TAUS also organizes bi-monthly user group meetings and webinars for new users.

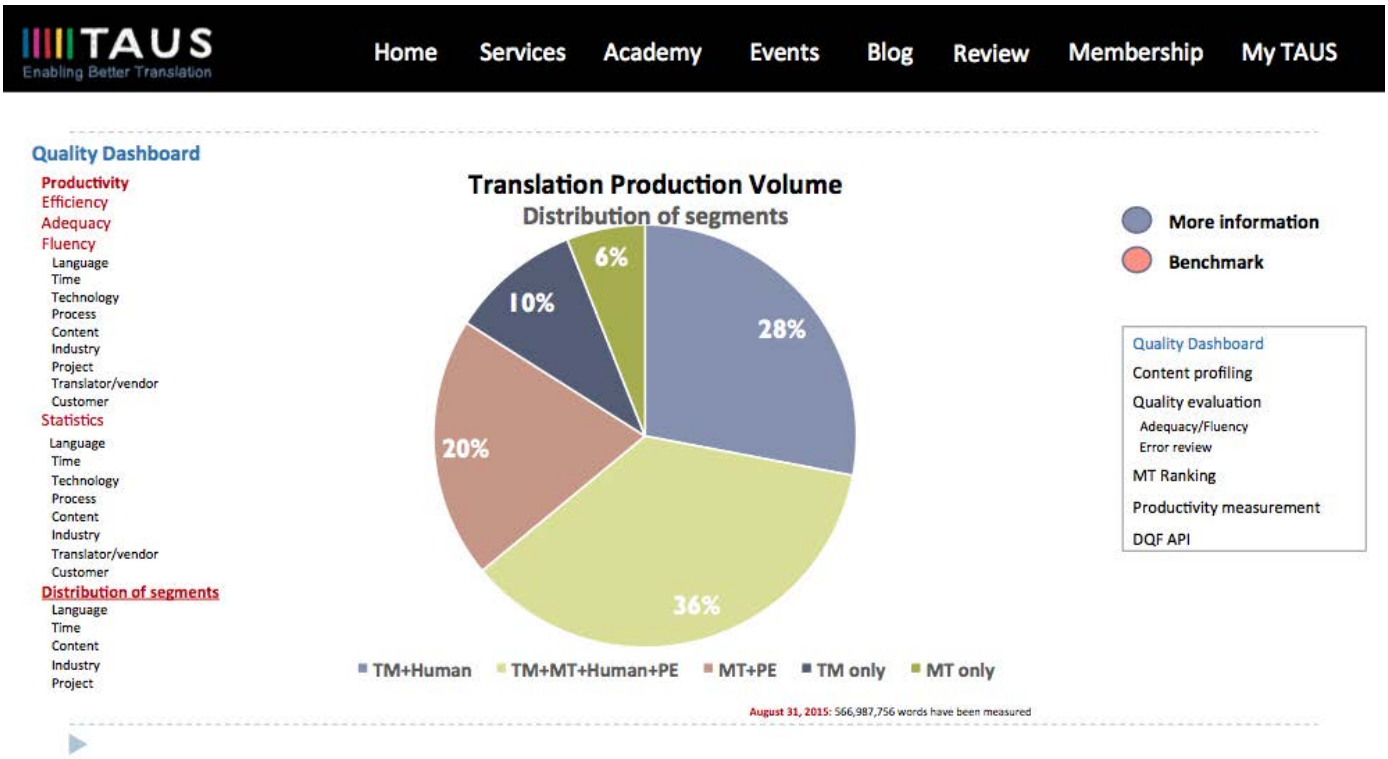
The Quality Dashboard will follow the needs of the industry and continuously add new tools and features. On the roadmap for 2015 is the release of a full Quality Control environment (based on the harmonized DQF and MQM error categorization) with the possibility for companies to do error annotation and correction. There will also be an API for the Quality Control tool to allow members and subscribers to plug this service into their common translation management environment.

# 5. How to use the Quality Dashboard

## Productivity scores

Productivity measurement is just one way to evaluate translation quality. It can be an indicator for the quality of the translation technology (like machine translation engines) and/or the performance of the translators and translation suppliers. It may also be an indicator for the complexity of certain content types, languages or industry domains.

On the public TAUS web site pages the statistics can be checked by everyone who is interested in tracking and comparing translation production. One statistic report for instance gives an overview of the distribution of translation production.




Or a report of the average productivity by language pairs.

Productivity measurement can best be done through the DQF API that is offered to all service and technology providers as an easy plug-in. The translator and manager are requested to create an account. Productivity measurement can be controlled by the manager and the translator. Every time a translator completes a job s/he receives a link to a productivity report.


Upon completion of a project, the manager will also receive a link to a benchmarking report. This way, all stakeholders in the translation process – clients, managers, vendors, translators – can track and compare productivity scores. The statistics and basic benchmarking reports will be part of a free offering and can be viewed on the TAUS site.

## Efficiency scores

With the launch of the Quality Dashboard TAUS introduces a new metrics: the Efficiency Score. This score will enhance traditional productivity measurement. Productivity is widely used for measuring the throughput of translators or quantifying the quality of MT engines. While the productivity score is a good first performance indicator, it ignores a number of factors that should be taken into account when assessing productivity and quality (e.g. final quality of the translation, edit distance, difficulty of source).

The TAUS Efficiency Score can be applied to every translation project: translation from scratch, translation with translation memory, PEMT or a mix of these three. The Efficiency score is flexible in that the number of variables used to calculate it and the ways the different measurements are taken into account vary based on user requirements and the available data.

The Efficiency score can be an absolute score calculated based on one given project or a relative score that is calculated using all the relevant data in the DQF database. It can be calculated using the two obligatory variables (core variables) or by adding some optional variables to the calculation to increase precision and credibility. It can be calculated to measure translator efficiency but the focus can also be on CAT/TMS efficiency or MT engine efficiency.

## Variables

More and more translation jobs have a mixed nature: one can post-edit MT suggestions, insert TM matches or translate segments from scratch in the very same translation job. There is no hard divide anymore

between MT, TM and human translation. This should be reflected in a new metric. In the TAUS Efficiency Score, time is measured for producing (and, if needed, updating) each segment regardless the segment origin (MT, PE, Human or all).

While edit distance and the edits per hour are calculated in many translation tools, this measurement tends to only be applied to evaluate MT engines and less so for evaluating post-editing productivity. Simply because no one has come up with a method that would combine a productivity score with edit distance information and normalize the score in a dynamic way. This is exactly what the TAUS Efficiency Score does when it is based on the core variables.

Finally, in order to unify two measurements (processed words per hour and edits per hour that is based on edit distance) one needs to convert relative scores into absolute scores. We create the Efficiency Score on an ongoing basis using data from the DQF database. The more data and the more homogenous data is used to calculate the score the more precise and meaningful that score will become.

To sum up, for the Efficiency Score based on the core variables, we measure time for processing segments while tracking the segment origin. Next, we measure the edit distance and calculate the edit distance per segment (minimum number of edits needed to get from A to B) and produce the number of edits per hour. Finally we normalize and unify the two measurements. For more precision and credibility, we can base our calculation of the score on additional (optional) features.

Below is a mock-up of a combined productivity and efficiency report benchmarking 'My MT engine' with Microsoft's and Google's MT engines.


The advantage of both the Productivity and Efficiency Scores is that it is part of the translation production process and it does not require additional steps, resources or time to collect the scores. All productivity, efficiency, statistics, distribution of segments reports are generated on the Quality Dashboard through the use of the DQF plug-in during translation production.

# 6. From Quality Evaluation to Business Intelligence

The main function of quality evaluation is to adjust and optimize processes. That is true for most industry sectors. In the translation industry however quality inspection is limited to the output: the translated text. Quality evaluation is a necessity without remedy. For most buyers of translation, as well as providers and translators, translation is like a black box. They miss the insights, the data and the business intelligence to take well-grounded decisions about processes, tools, resources, budgets.

The Quality Dashboard will help all stakeholders getting the indispensable statistics, benchmarking and analytics to take well-informed decisions. In this chapter we list examples of the questions that different stakeholders may have and that the Quality Dashboard will be able to answer.

## Buyer's questions

**1. How are we performing compared to industry peers, in terms of production throughput, quality (adequacy, fluency, number of errors and efficiency)?**

- The Quality Dashboard provides the benchmarks by language, content type, industry.

**2. Should we consider using MT for support content?**

- The Quality Dashboard shows the percentage use of MT on support content, and also shows the average productivity and quality scores.

**3. How is our translation technology ranking compared to industry?**

- The Quality Dashboard shows the scores for different technologies, also broken out by language, industry, content types.

**4. How good is our MT engine?**

- The Quality Dashboard shows the productivity and quality scores compared to industry.
- In the Analytics the manager can run additional evaluations and measurements to compare the sores of specifically customized engines with averages.

**5. Are we paying our translators fairly?**

- The Quality Dashboard shows productivity (number of words per hour) for translators on average and per translator.

## Project manager's questions

**1. What is the correlation between the productivity (number of words per hour) and the number of edits per hour, on total and per translator?**

- The Quality Dashboard provides the numbers and the benchmarks by translator, language, content type, industry.

**2. How did our scores change since we changed our translation technology platform?**

- The Quality Dashboard shows the productivity and quality scores by any selected time period.

**3. Should we change to the NAME technology for the translation of our web content?**

- The Quality Dashboard shows the productivity and quality scores by any selected technology.

**4. Where did our translators take their translated segment from?**

- The Quality Dashboard shows distribution of segments and presents the percentage of human translation, machine translation, translation memory, post-editing.

**5. How did project A compare with project B?**

- The Quality Dashboard shows the productivity and quality scores by any project.

## Translator's questions

**1. How is my productivity compared to industry averages?**

- The Quality Dashboard provides the numbers and the benchmarks language, content type, industry.

**2. How am I doing on this project compared to the other project?**

- The Quality Dashboard shows the productivity and quality scores by any project.

**3. Am I using the best tool for the job?**

- The Quality Dashboard shows the productivity and quality scores by technology in comparison to other technologies

# 7. How to Get Started?

There are several ways to get started using the TAUS DQF and the Quality Dashboard. Please see below to learn how to generate reports in the TAUS Quality Dashboard or integrate the TAUS DQF API in your translation technology.

## For Users

You can access the TAUS Quality Dashboard here and obtain free reports with industry average data on productivity. To see more advanced and personalized reports and benchmark your own data against industry average, you need to have a TAUS account. If you do not have an account with TAUS yet, you can subscribe here.

Visit the DQF users page to find your translation tool(s) of choice and follow the instructions on how to get started. Once you start using DQF for your projects, the TAUS Quality Dashboard will generate productivity and efficiency reports for you.

Should your tool(s) not be listed, please let us know. Our Member Services team is always available to help you and answer any questions you may have at memberservices@taus.net

## For Developers

TAUS welcomes any translation technology provider who would like to integrate DQF into their CAT tool or Translation Management System. You can obtain more detailed instructions on how to become an integrator in the "For Developers" section of the TAUS website.

You can obtain the free TAUS DQF API specifications directly from the TAUS website or from Github. Please write an email to dqf@taus.net if you have any questions on how to get started. We are happy to welcome you to the developers community.


TAUS is a resource center for the global language and translation industries. Our mission is to increase the size and significance of the translation industry to help the world communicate better. We envision translation as a standard feature, a utility, similar to the Internet, electricity and water. Translation available in all languages to all people in the world will push the evolution of human civilization to a much higher level of understanding, education and discovery. We support buyers and providers of language services and technologies with a comprehensive suite of online services, software and knowledge that help them to grow and innovate their business. We extend the reach and growth of the translation industry through our vision of the Human Language Project and our execution with sharing translation memory data and quality evaluation metrics.

For more information about TAUS, please visit: https://www.taus.net