**Globally Speaking**

**Episode 055**

**Advancing the TAUS Dynamic Quality Framework (DQF)**

M I’m Michael Stevens.

R I’m Renato Beninatto

M And today on Globally Speaking, Renato you have invited someone to talk about quality, and I know you hate talking about quality. So why would you do that?

R Well as I say, quality doesn’t matter. It’s something that you don’t talk about. It’s like you don’t negotiate with terrorists; you don’t talk about quality. That’s my position.

 However, our guest today is a guy that has a fantastic history in our industry. He is one of the fathers of localization as a service, and has been involved in all the efforts in the history of the industry related to measuring and controlling quality. He’s a friend of mine. I’ve worked with him in the past, and I’ve learned a lot with him.

 But the thing that I like the most about our guest is that this is a guy who understands our industry so well, that the world that we live in today in the translation and localization space is the world that he envisioned 20 years ago.

 Twenty years ago he was talking about these concepts of automated translation, translation-as-a-service, plug and play, automatic project management. He was too early for his time, but I think that the initiative that we are going to talk about today, DQF, is something that is more than ripe.

M Yeah. And so we’re going to get to hear from him a little bit about DQF. So let’s let him introduce himself.

J Hello I'm Jaap van der Meer, director of TAUS. I think most people know me in the industry—that’s at least what they tell me when I accidentally introduce myself at receptions and conferences in this industry.

M Let’s start with what does TAUS stand for, and what does TAUS do?

J Yes, as I said, I've been around for a few decades, started sort of accidentally at a translation company around 1980, last century, straight out of university. And that became one of the first localizations, software localization companies in Europe. And in my business I've always been extremely interested in the technology, how it could help us to the maximum. After running translation companies for many years, my first company was called Ink and got swallowed up by Lionbridge.

M Is that with a C or a K?

J A K. It was the raw material. Ink island way. At that time we delivered a lot of translations still printed. You can't imagine that now.

M Translation by the pound, is what I’ve heard that called.

J Yeah. Exactly.

R But the kilo in Europe.

M And the kilo.

R You need to localize it, Mike.

J It was a stack of paper with a big floppy disc on top of that. As big as an old long-play record. Another thing you don't see anymore. Well you do see them again, of course. But not the big floppy disc. Anyway, after running translation companies nationally, European, globally, for many years, I started TAUS because of my passion for the technology. That was sort of a coincidence as well, I was running a round-table meeting at one of the first LocWorld conferences in San Francisco for anyone who is particularly interested in machine translation. That was in November 2004. Nobody was really using machine translation at the time but the big IT companies, Microsoft, IBM, Oracle, et cetera, Sun Microsystems; they all had one or two people looking at the technology to see what it was capable of. And these people came together for a full day. It was an exciting day, learning from each other. And at the end of the day they asked me “Can you organize another day like this?” And that was kind of the beginning of TAUS.

R But what does TAUS stand for?

J I knew you were going ask that because I don't want to say that anymore because it confuses a lot of people. But now I have to. It's Translation Automation User Society. And why I don't like to use the full name anymore is that people then think that the only thing we do is machine translation. But we do so much more these days. It's about technology in general, innovation, quality, what have you. But anyway, that's a short introduction of TAUS.

R Well and I would add I mean, you were also involved, you mentioned your decades of experience, what you didn't say is that you were also one of the founders, one of the initiators of LISA which was the first industry association, global industry association we had, the Localization Industry Standards Association. I think you have a knack for choosing names for associations that don't do what they say in their names, because LISA didn't do much about standards, and TAUS doesn't work only on translation automation, but that's beside the point. But one of the things that we have avoided like the plague here in this podcast, is talking about quality.

M Only when we're forced to.

R Yes. It's one of those topics that I don't like to talk about because they never end, but if there is somebody in the world that can talk about that, it's you. So we want to start with a grand opening on this topic covering the idea. Because you have always been involved with quality initiatives, with metrics, with measuring effort, and you were part of the group that created the LISA standard, not standard. The LISA QA Model that people still use today, and this is something that was done back in the 90s. But the reality of the world changed completely from what it was at that time when an Excel spreadsheet was a big advancement. What is the involvement of TAUS with quality?

J Yeah, thanks for that question because—and by the way the LISA QA model and then the SAE J2450 standard or metric—I was heavily involved with that as well so indeed, I don't know what it is that attracted me to this topic. But let me try. Why, let me say at least why it came up again in the TAUS think tank. Because originally TAUS was just a think-tank, you know bringing people to think deeply about some of the basic matters that concern us all. Like how can we help this world communicate better? How can we reach more people? I mean we're only a couple of hundred thousand professionals working all this content and the mountain of content keeps growing. So we need to be able to do more. So, as a think-tank we came up with the recipes for using machine translation initially. But then the problem that came up in our conferences, in our round-table meetings, was also how do we know we delivered a quality that the users really need?

 And that problem became bigger and bigger as people started to use more technology, or let's say, innovative ways of producing the translation because crowdsourcing was another big trend of course in the last 15 years or so. And at the same time you saw a diversification of content. In the old days we were just translating user instructions, manuals, and what have you, but now people need to translate social media, knowledge bases, support articles and what have you. And in this environment the old issue of how do you know you're delivering the quality became more pressing, more urgent.

 So as a think-tank, we were drawn in to this like what are the models that people are using? Indeed, LISA QA, SAE J2450, ISO standards, what have you, but none of them really help us because we need to have different levels of quality, different levels of output. So we started again with our think-caps on drafting a model that was more dynamic. And that's how we came up with a *Dynamic Quality Framework.*

 And by the way this is not something that we did just on our own. TAUS is always engaging with our members and user groups, and we also for this particular project we started back in 2010, I believe, already worked very close together with Dublin City University with Sharon O'Brien, so we got the academic input. Ultimately, we came out, I think in 2010, with the white paper called Dynamic Quality Framework. That was the blueprint for a model for assessing quality based on content profiles. And the content profile, I mean it was maybe a bit pretentious, I don't know. But we set up this what we called UTS scoring—Utility Timeliness and Sentiment. So three factors, by which you could profile the content. Is this content that requires speed? Like a fires alert?

 You don't worry that much about fluency, but the accuracy is important. Or is this content that really, you really want to touch people's hearts? It's really about the brand, an advertisement, or a marketing leaflet. You need something more fluent. You need something that really speaks very well, is personable, and what have you.

M One question I have in just your description so far, it seems like, as the industry commercialized more, that people with business backgrounds, solely, who didn't have the language background entered into, along with the evolution of the content that was being created, that that outpaced the LISA standard, and that's where DQF came in. What are other factors do you think were happening at the time that made you say okay, now is why we need to create this?

J Good question, Michael. If I place it somehow in our own description of the evolution of the industry, I would say this is typical for the convergence era that we live in and how we label this decade that we work and live in now in our industry. Translation is not an isolated activity anymore. It links with marketing, with support, with anyone in the social media, in the search, in the IT department, so the translation localization function is very connected to the overall business. Which clearly, of course, changes the demands, the requirements, and we need to think more business-like, as you say.

M It still has that foundation of, with the iron triangle of quality, cost, and time.

J Oh yeah.

M But the DQF is able to bend the triangle.

R Yeah. I think that the key word in DQF is dynamic, right?

M Mm-hmm (affirmative).

R Because it changes according to the type of content. One of the problems that existed before in every quality methodology that is usually driven by the translators is that there is a huge focus on errors and mistakes. The LISA QA model, you would fail a translation because there was one error, and sometimes, depending on the type of document that you’re translating, that error might be irrelevant. Or it was… You fail a document that is ten words long yeah it's turning into a problem, right? So the element of having a dynamic framework that adapts to the type of content and it's just a matter of anything else that you want to measure is agreeing up front what you're going to measure, right? But tell us more about the DQF and what are benefits of a program like this, and eventually we want to know how it works and how our listeners can be drawn in and participate.

J Yeah, thanks. I guess I need to describe it a little bit more before we can even talk about benefits because the dynamic as you rightfully said that's the key characteristic, if you like, that being able to adapt to the content, the audience, the context you need different quality levels, but the next question that most immediately comes to mind is “then but how do you measure?” And that's exactly the point here. You need to have a way of measuring it precisely, with the phrase that we use very often now if you can't measure you can't manage, so how have we been managing the translation business for the last couple of decades? We were not measuring. We basically just in most cases asked an in-country reviewer, a subject matter expert, or a linguist, “what do you think of this translation? Is it good or bad?” And that was in that sense the way the business was run …

R Is.

J Yeah it is.

M It is still run.

R The majority of the people are not in DQF unfortunately but the reality is that most of the work done in the world is done in this archaic way, right? Where you have initiatives like this that we want to help promote because they bring the whole concept of quality to a different level.

J Yes. Yes. And as Michael was saying, again correctly, we're now connected to many different sorts of functions in the enterprise, and we need to think more businesslike, so how do we do this? Well, let me very briefly give you an evolution of this think work that we did at TAUS because we published this whitepaper which was just a theoretical framework. Okay you need you to do content profiling; you need to do this and this and this, but then the members came to TAUS; we have members right, you know, who have a lot of influence in what we develop and what we document and so on, and put out as research and reports, so they said this is good thinking, but how can we use this?

 So, there were these tools that actually let us use this framework, so we developed the DQF tools, and they're still running on our website, and at first instance they were developed for our members to evaluate machine translation output, so you can still go—it's on our site. You upload a machine-translated text, and you can do productivity post editing, and so you get a productivity score. You can do adequacy fluency, A/B testing, anti-comparison, and that's nice. But then the members came back and said, but we like to not just do our incidental MT evaluation; we'd like to score all of our human translations, everything, so what we need from you is an API, because then we can integrate these tools into our CAD tool, into our workflow system.

 So, there we go. We started developing an API, and this is three years ago. So we delivered the API, and then of course we were dependent on the CAD and translation workflow developers to use this API and create plug-ins in their technologies.

 It's been quite a long process but when that started to happen, I tell you we were surprised. We didn't know that when we started this whole venture, but there is so much data collected in the actual translation process. When a translator's working in Trados or in Memsource or in XTM or whatever tool, there is a tracking of how many milliseconds they work on a segment. How many mouse clicks. How many edits they make. Combine that with the metadata that we ask the project manager or whoever is starting the project to provide like content type, industry sector, process, which MT tool is behind this process, which technology is used. All of these data points plus then the production data that's collected during the process of translation and the review creates a magnificent set of data that changes the whole business completely, because now we're in a data-driven business. Now we can measure everything. And that's the aha in our evolution.

 We had no clue when we started this, but that was the moment when we said “this is a game-changer for everyone.” It's for everyone. Let's not position this as the big brother's watching you, buyer can see it all. We want this to be in the hands of the translator of the LSPs. Everyone in this cascaded supply chain should have its own view of this rich data set and be able to say I fire my customer because I don't make any money off this machine translation ingenious crap. It doesn't deliver me the benefits that they told me it would give.

M That is a huge benefit, and it's no surprise that it puts companies that are smaller size into the conversation with these data-rich companies like the Amazon and the Google and the Facebook. The first time I heard of DQF was having coffee here in Seattle with an Amazon employee. And he said have you heard about this, and I was still talking about the LISA standards at the time and I had to go home and do some homework on it, but because of the rich-data level that it gives you, it's really appealing, and it makes for a better business conversation.

R So Jaap. I'm a software publisher. I am an LSP. I want to participate in TAUS. How does that work? Do I have to buy something? Do I have to donate something? What is the way that you participate in this effort because as far as I understand, DQF is not something that you do alone, right? You're part of a community, and you're sharing some of your metrics, so that you can leverage them.

J Well indeed, yeah. For us we drive this with a vision, and sometimes that's hard, because obviously we're not alone in this, and you can with the more sophisticated tools do data education as well within that sort of environment, but then you're within a silo, right? You're only within that tool environment and you can benchmark maybe your own projects historically or compare within your own organization how you do against another department if they use the same tool set and so on. But what we have in mind here is to position this where it's there, and it's available for everyone to accept it in this form, but to get one metric for the global translation industry. One platform where you can get your business intelligence, and if we succeed, then that's really the way for everyone in this industry to carry on. To do the real work. To really start adding the value. To stop arguing about issues that don't matter that much because we all make them very big.

 But I need to explain you a little bit more than what's inside under the hood, so to say of DQF because I told you we have the DQF tools. We have the content profiling, but the first thing we did was develop the best practices for fluency evaluation, for adequacy or accuracy evaluation. We have also worked with industry with all the interest groups to document the best practices for post editing which are broadly globally being referred to the TAUS post editing and that's all inside the DQF, the Dynamic Quality Framework.

 Now we're lucky something else happened because there was of course the LISA QA model that most people accepted that that was no longer supported because LISA doesn't exist. You hope that what you develop is being accepted as the standard. That's sort of in the standard's status, so to say. While we were working on this in Europe, there was this other initiative from the German Research Institute, the DFKI, the MQM metric-multi-dimensional quality metric. And we worked also with the European commission that was funding this MQM project, and there was a follow-up project, and we both said DFKI and TAUS, this doesn't make sense to have two metrics being developed at the same time. Industry will be confused again. Everybody will be asking what are we going to use? DQF? MQM?

 So, we were funded to harmonize the two models. And that is what you hear people talk about now. DQF, MQM, some people say MQM, DQF, depends on where you come from. But it's the same. And it's a harmonized model, and it's becoming a standard now. And it's even officially, it's under review by the ASTM, the American Standards Organization. It's being reviewed by ISO, I think, so hopefully it will get this rubber stamp, and under the hood, that's the core of DQF, too.

 So then there's the reporting, the DQF dashboard, allows you to see real-time reporting of your projects, and two different things are being measured, productivity and the quality. The productivity is measured totally automatic, you don't need to switch on or off anything. So it counts the number of words output per hour and the number of edits made per segment per hour. So the edit distance. These kind of entities. They're all tracked fully automatically.

 And some users are only doing this at the moment. And so they can compare machine translation, translation memory, human translation, by segment. And so the whole promise of DQF is to give you razor sharp measurements. You can get to ROI on a segment by segment basis. But what's more than that is that you can benchmark against industry averages. Within your peer groups.

M When does the TAUS top 20 quality LSPs list come out?

J Well there you go! You got it, you got it! I don't think we'll sort of publish this, by the way, I don't know.

M I know some buyers who are asking for it now.

J Well, I guess, my prediction is that we'll learn that it’s not so much about this quality. It is much more refined. It's a whole sort of, much bigger statistics. Probably, as we develop this data and people have access to all this data, I can get the data, because the one part I didn't mention that's part of the DQF, is the data connector. That allows you to download all of the data out of the DQF database, your own project data, and normalized industry data.

 So, everyone can populate their own dashboards because people won't. Everybody has dashboards, and they don't want to come to like Tableau software is using DQF, and you can imagine they don't want to go to our primitive dashboard; they have the best data visualization in the world. So they were the ones to say, "We need the data connector." I always say it's kind of a reverse API. You can download all of the data from the DQF and then populate your own dashboard.

 But as we integrate all this data, I think we'll say, "It's not about the best quality of this particular translation; there's a lot more behind it".

M It can lead to some very good business conversations, I think about when I entered the industry, the fear of being benchmarked by a client that I was serving, and they said, "Okay, we have four vendors, and this is where you're performing in related to these four vendors," and that ended up leading to improvements with our company. Better vendor-client relationship. There was great partnership that came out of that. Imagine having that across the industry. It seems like it could, like you said, help us focus on the things that are important.

R Yeah. This is why I don't like to talk about quality. Because if I start pulling the thread, it's never going to end. But I keep on, and a comment that I would like to make to what you're saying is that, essentially, this helps you measure the linguistic quality, but that's only one element in your relationship between all the parties involved. There are other elements that are more important, how do you respond to problems in the quality? How do you adjust your processes to avoid mistakes that were done before? So the value of the data is not so much about, and this is where I have a hard time talking about quality, is that there is a huge focus with everybody in our industry to talk about errors and mistakes, but errors and mistakes are things that you need to make in order to be able to avoid them.

 So the value of a tool like DQF is not so much in what it catches, it's what it helps prevent. I've known Jaap for many years and the value, when Jaap says "I have a vision…”

M It usually comes true.

R It comes true. The world that we live today, in the language industry, is the world that Jaap described 20 years ago. I heard him describing this world of translation out of the wall, remember? Your vision of plug and play, and that's what we're having more and more and more today. So if he says that he has a vision, it's going to happen. So jump into this bandwagon.

M That’s great. Probably the first place to send people is to the TAUS website?

J Yeah, sure. Yeah, back to your question, Renato, about where do you start, as I said it's a very inclusive model. We didn't want to develop this as a tool for buyers to manage their vendors and translators with much more power or what have you. We wanted this to be a framework for everyone. So we have seven levels of memberships, so basically, yes, we ask people; there's free trials obviously, as for everything these days. But if people like to really start using it, well first I need to check, of course, whether you're CAD tool or translation workflow system has been integrated with DQF. Fortunately more and more do, and others are working on it.

 And then they can try it out, and then they take a TAUS subscription for freelance translators; it's 120 Euros per year, and it gives them enough capacity because we have counters on the volumes that go through the pipes, so to say. And for very big companies like Microsoft and eBay and Dell. These subscriptions are much higher, of course, but they also have much bigger volumes they can process. So that's essentially how it works.

R Wonderful

M So you, too, can be on the same level as Microsoft, Dell and eBay? By getting involved with TAUS?

R Exactly. And keep in mind it's all about getting better. It's not about catching mistakes. That's the thing that irritates me the most when we talk about quality in our industry; people want to talk about bad quality.

M You know, that's a mistake.

J Well you know, it's far beyond that. It's not at all about catching mistake …

R Exactly! That's what I like about DQF

J I know you were joking. Let me tell you, in sort of a good sequence of what the benefits are. You can start collecting when you start using this first of course you have for the first time, you have an objective way of measuring your quality. You can take informed decisions because you're collecting the data consistently and also in a way that it's comparable to all of your peers in the industry. So you can trust the data that you collect, and you can eventually get to an ROI by segment because you don't only have the review data on quality, but also the productivity data.

 So, another thing that's really very beneficial is the efficiency of reporting. When we worked on this, we were really astounded by, we're talking to people, big and small companies, very big companies. When asked, "How do you do your quality review and reporting?", they very often use some kind of a primitive spreadsheet on the side of their translation workflow where they log errors to some self-invented or some derivative of an old QA model.

 And then they have to manually develop reports and copy paste information, so it's a lot of work. And DQF, it’s real time and instant reporting on projects, benchmarking, trans reports and what have you. So big efficiency gains. But then where the real benefits come in is later, is on the business intelligence. You can really learn from your data, and take more informed decisions about comparing technologies, machine-translation engines, vendors, translators, processes, which then, in a more sophisticated form will lead to machine learning and developing algorithms that help you to automate steps in your management process, like predictive analysis, resource allocation, confident scoring.

 You hear talk about it by the big tech companies sometimes. I think they're making it bigger than it is because we're still early days there. But you need to start, and it's not just for the big guys. That's the whole point. I think we want to see this being shared so that it's available to everyone.

R So, we're going to have to invite you back to talk more about the concept of dynamic quality, quality at source, and all these other concepts regarding ... we'll have to make more progress about quality.

**END OF CONVERSATION**