



10TH ANNIVERSARY

Newsletter

THE LOCALISATION INDUSTRY STANDARDS ASSOCIATION

LISA © 2001 VOLUME X № 3 • ISSN № 1420 – 3693 • REPRINT

August 2001

GLOBALIZATION TECHNOLOGY

Globalization Technology Will Prevail!



THIS ISSUE IS PRIMARILY ABOUT technology, technology and technology. The intent is to provoke the imagination with some “out-of-the-box” thinking. Too many technology discussions these days are about everything except technology: financing, management, missed expectations, etc. In tough times it is tempting to focus on why things can’t get done as opposed to what can be done. It almost feels as if technology has failed and stopped advancing but, let’s face it, technology won’t stop tomorrow or the day after.

Technology moves ahead; not all projects, not all companies, but technology as a whole moves ahead. Forget the temporary setbacks of implementation errors, failed experiments, management problems, cash flow problems, etc. These are merely punctual events; technology still moves ahead.

Globalization technology is moving ahead. UNICODE is certainly moving ahead (more so than appears at first sight, see François Yergeau’s article). Various companies have been, or are still, experimenting with automated workflow, translation memories and various other approaches to improve the localization process. These include, just to name a few: Bowne, Compaq, GlobalSight, Idiom, LionBridge, Microsoft, Sentius, Trados, Tridion, Uniscape and WizArt. There are many others I have not named including university labs and home garages. It doesn’t matter who does it, technology happens because people have a vision they believe in.

The question is: what is the vision? Where is globalization technology going? Where will we be in 5 years? In 10 years? Will the future be bleak or will there be a whole new world of challenging opportunities?

Guardians of Humanity

Let’s consider what our industry does. Our role is benign: we don’t pollute, we don’t make weapons, we don’t even

clone human beings! What we do, to a large extent, is keep a measure of humanity in man-machine interactions by respecting the language and customs of the user. As these interactions get more complex, so does our role (localizing voice applications, for example).

Take a step back and look at the big picture. On a social and historical level the role that we are playing in the information age is more than simply worthwhile, it is fundamental: we are responsible for keeping human culture present in technology. In the cyber-revolution, we are guardians of humanity.

I submit that globalization technology will persist and progress and all for the betterment of mankind!

We Are Not Alone

Yet, as pointed out in Don Plumley’s article, the results so far have been somewhat disappointing. **Why?**

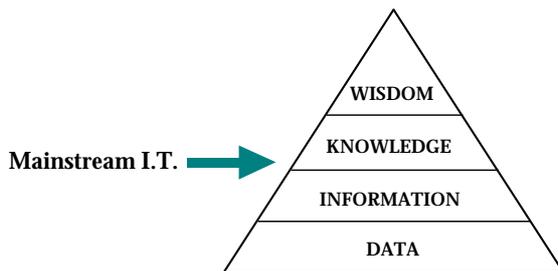
Let me first point out that we are not alone in this respect. There is a multi-trillion dollar industry that is even more “technologically impaired” than ours: the health care industry. Amazing amounts of money are spent on various drugs and treatments, yet it is extremely difficult to get any data about their real efficiency because patient records are still kept on paper (of which more than 25% are unreadable even by the doctors who wrote them!).

I spent three years working as architect for a medical knowledge base. Working with a terminology composed of 300,000+ medical noun phrases in several languages proved to be very similar to localization. In both cases, simple tasks (translating a word, coding a disease) have to be performed thousands of times by a knowledge worker supported by dictionaries and other online resources. In both cases, there is a production process involving knowledge workers managing terminology.

Still in both cases, the reason why technology has not yet delivered the desired results is because the problem is much more complex than anticipated. IBM entered medical terminology management and eventually pulled out; one

U.S. government official was worried that it was simply too big a problem for any single company to handle. I've heard the same thing about globalization technology.

Consider the well-known pyramid below. Mainstream information technology, with spreadsheets and databases, has handled the information level. A lot of what could be done with "information" has been done. Knowledge management is the next big hurdle for mainstream I.T., and it's not that far away, especially since there are a few trillion dollars (in health care) that need to see it solved. We shouldn't wait that long, since globalization knowledge is much less complex than medical knowledge.



The enabling technologies are the same for localization and for computerized patient records: innovative task-oriented ergonomic design for production work, knowledge management for knowledge workers and workflow to connect the pieces together in an integrated production process. Speech recognition is an added bonus to accelerate data input, be it in the emergency ward or on the translator's desk.

Language is... What People Speak

Speech is the fundamental expression of language; it's what we learn first, it's our preferred mode of communication. "Do you *speak* X?" is how people enquire if you know the X language. Speech technology is progressing; it also won't stop. Inevitably, over the next few years, speech technology and globalization technology will become increasingly inter-dependent.

Speech synthesis has evolved considerably and some synthesized voices are now of near-human quality. If you imagine synthesized speech à la Stephen Hawking, I suggest you listen to the demos at <http://www.speechworks.com/demos/speechify.cfm#>.

Speech recognition has also become mainstream and multilingual. The SpeechWorks article in this issue, by Jose Elizondo et al., describes localization issues specific to voice applications. Interestingly, human beings will always imagine a person when hearing a voice, so voice

localization issues will include the accent, personality and gender of the virtual agent.

Speech technology will not only provide opportunities for localization work, it will also be an enabling technology for improving translation tools and reducing cost of translation. In the short-to-medium term, we will see a voice-enabled translator's workbench, which will allow translators to enter data verbally, several times faster than typing. Interestingly, this can be accomplished by combining voice recognition with machine translation (MT): MT can provide a list of candidate translations for each given sentence or word to translate (this can be even done offline, in advance); the list of candidate translations dramatically constrains the vocabulary space for the voice recognizer. The translator never sees the MT candidates; he simply gets the benefit of a quick voice recognizer with a quasi-perfect recognition rate.

Again, with a new emphasis, language is what *people* speak. Governments can recognize and support languages, but it is people that define and use language. I don't know if there will ever be a Cajun Windows, but we need to keep an open mind on what languages are, how they evolve and how we decide to support them. The article by Marilyn Mason and Jeffrey Allen shows us that French Creole, like French, has many varieties that need to be taken into account.

Ergonomics and Production

Ergonomics is another fascinating field of growing importance. It is similar to localization in that it is "human-centric"; its objective is also to ensure a quality interaction between man and machine. However, whereas localization is concerned primarily with cultural issues, ergonomics is mainly task-oriented, i.e. ensuring that the application interface allows users to perform their tasks naturally and efficiently.

There is yet another unfortunate parallel between ergonomics and translation: they both look easier than they are because they are close to our daily lives. Many people believe they can handle ergonomic design or translation without any special training. In fact, studies have shown that the most efficient interfaces developed using ergonomic techniques are often not those the users or the designers expected, i.e. how to produce efficient intuitive interfaces is often counter-intuitive! Ergonomics is a domain in its own right; you only need remember how frustrating software or web site interfaces can be to understand it is both important and non trivial.

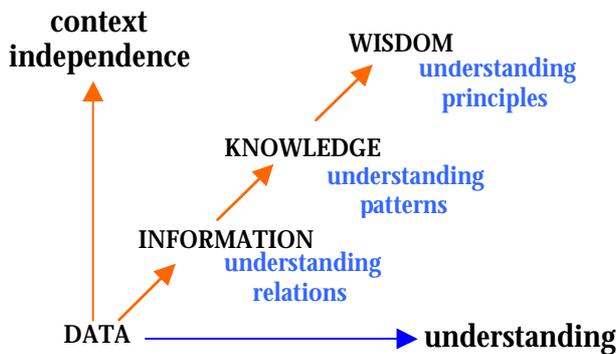
A perfect example of the value of good ergonomic design is the TRADOS translator’s workbench that has the lion’s share of the market. This success is not due to the presence of “translation memories” (which are present in most translation tools) but rather to the way the translation memories (and glossaries, etc.) are presented to the translator. What is important is that the translator can work efficiently, hour after hour, translating text with this tool.

Ergonomic design techniques are particularly useful in production environments where the same tasks are performed thousands of times. Localization is precisely such a production process involving project managers, internationalization and localization engineers, DTP specialists, translators and testers.

Most of the people involved could use better, more scalable, more task-oriented tools. Ergonomic design techniques provide proven methods for achieving this. Smarter tools supported by knowledge management will be even more task-efficient, while workflow technology will connect them together to optimize the process itself.

Knowledge Management

Let’s start by clarifying that knowledge management is not artificial intelligence; it can in fact be very simple. The basic idea is just to store knowledge and re-use it later on. The chart below expands on the pyramid seen previously:



It identifies context independence as the main quality distinguishing various levels of understanding. Context independence facilitates re-use of knowledge.

Translators are knowledge workers. Each translation is a piece of knowledge, a decision made by the translator. A translation memory is a collection of translations, effectively a knowledge base. Within a given translation project, the translation of any given sentence is simple to store and highly *context independent*, thus easy to re-use.

This is why translators have an efficient knowledge re-use tool on their desks, while internationalization engineers, for example, do not (although OneRealm’s tool has taken steps in this direction).

The decisions taken by internationalization engineers are more complex: they are pattern-based, context dependent and there are many different types. Yet many internationalization tasks are well understood, reasonably simple and quite amenable to knowledge management. The same holds true for many of the actors in the localization process. Knowledge management can help optimize the entire process by simplifying or eliminating the more repetitive tasks. Considering the impact that translation memories have had on translation, extending knowledge management features throughout the entire localization process should prove to be quite rewarding.

Workflow

Workflow technology is a commercial reality that has generated considerable interest; see, for example, <http://www.wfmc.org/> or <http://www.e-workflow.org/>. Yet it remains notoriously difficult to implement. In particular, replacing human beings, as some have suggested, is certainly not the way to do it.

In a very stimulating talk at ACM 97 on the next 50 years of computing, Pattie Maes provided her vision of how man and machine will collaborate. She called her approach IA rather than AI, Intelligence Augmentation rather than Artificial Intelligence (see <http://research.microsoft.com/acm97/pm>). The idea is simply that humans and machines can work as a team because they have complementary capabilities. Machines have excellent memory and compute very fast, humans have a better understanding of language, are better equipped to deal with unforeseen exceptions and are certainly more capable of negotiating a real commitment from an employee or vendor!

A perfect illustration of this principle is the difference between MT, an AI approach trying to replace the translator, and a translator’s workbench with translation memories, an IA approach where man and machine collaborate. Needless to say, IA is not just politically correct, it’s also much more practical. Note also that a third approach has recently appeared between MT and human translation: it is called Automated Content Enrichment and consists of a machine enriching content written by man; it is described in the article by Marc Bookman.

I believe that capability-based man-machine collaboration is the correct mental model to develop knowledge-based tools and workflow. The tools help knowledge workers get the job done; control still remains in the workers hands (i.e. using a pull model will often be better than a push model).

Workflow automation is not only useful in the globalization field, it can and is being used just about everywhere. Consider how many companies have developed their own workflow engines: globalization vendors like GlobalSight, IDIOM, and Uniscape; Content Management System (CMS) vendors like Vignette and Interwoven; enterprise system vendors like Oracle and PeopleSoft and numerous dedicated workflow vendors (e.g., see the comparative study at <http://www.wngs.com/>). There are several dozen workflow engines out there re-inventing, it seems, the wheel many times. The reason may be that it is easy to develop a basic workflow system, but a lot of work to develop a scalable, reliable, flexible and **useable** workflow system.

Given the time, cost and expertise required, it does not seem wise for globalization companies to develop their own workflow technology. In fact, as Otto De Graaf points out in his article, the globalization workflow has to integrate with the CMS workflow, which has to integrate with the enterprise workflow in order to achieve the true promise of the global e-business site. For this reason, Tridion are focusing their efforts on a generic workflow interface rather than investing in a proprietary workflow. Should globalization companies follow the same principles, the globalization workflow technology of tomorrow may simply be a slight customization of a prominent enterprise workflow system.

How Do We Get There?

To develop successful technology, we start by taking a good look at where we stand. In this respect, Don Plumley's article provides a sobering reality check that can only help build a more solid plan for the future.

We then follow industry best practices for software development. Michel Balcaen's article provides some realistic, down-to-earth advice on how to develop technology in a hybrid service/product organization. To repeat some of his advice, here is my selection of principles that are well known yet often violated (with disastrous consequences):

- Dedicated R&D team with an independent budget
- Seasoned R&D staff that have produced commercial software before
- A mature development process involving users throughout (e.g. using ergonomic techniques)
- Don't re-invent the wheel; buy what you can and build only in your area of expertise
- Regular deliverables: keep a short term horizon and build on success
- And most of all, a solid vision with strong buy-in, that will guide and motivate

About the Author

Pierre Cadieux (pcadieux@i18n.ca) is president of i18N Inc., (www.i18n.ca) a firm specializing in internationalization training and consulting for embedded systems, shrink-wrap software and Web sites. Formerly VP Technology at Alis, he pioneered the transparent handling of Arabic and Hebrew languages and created the core bidirectional technology licensed by Microsoft. As Director of Localization Technology at Bowne Global Solutions, he published the first generic model of Web globalization technologies.