An Open Service Framework for Next Generation Localisation

David Lewis, Stephen Curran, Dominic Jones, John Moran
Centre for Next Generation Localisation
Knowledge and Data Engineering Group
School of Computer Science and Statistics
Trinity College Dublin
College Green, Dublin, Ireland
E-mail: {Dave.Lewis|Stephen.Curran|Dominic.Jones|John.Moran}@scss.tcd.ie

Abstract

The localisation industry makes strong use of language processing pipelines at the core of its bulk localisation workflows, where software text and technical manuals are translated into the languages of target markets. Natural language technologies such as machine translation and text analytics are now maturing to a stage where they are being adopted as components in these workflows. However, they also offer the opportunity to broaden the localisation business into domains where the source content is less predicable and produced and consumed more rapidly in higher volumes by a wider range of users. To exploit the business opportunities of such Next Generation Localisation, the localisation industry must adopt a more flexible, extensible and lower cost mechanism for the integration of language processing workflows across many, increasingly specialised players. This paper outlines an open services framework that is being developed by the Centre for Next Generation Localisation, that will allow industry to react rapidly to changing business models and new opportunities by exploiting service oriented architectures for service reuse and (re)composition, extensible meta-data driven interoperability and flexible service and workflow management capabilities.

1. Introduction

Localisation is the industrial process of adapting digital content to culture, locale and linguistic environment (Johnson 2007). It is a key enabling, value adding, multiplier component of global manufacturing, services, software and content distribution industry so as a business process it must be conducted at high quality, speed, volume and low cost. The localisation industry makes strong use of language processing pipelines at the core of its bulk localisation workflows, where software text and technical manuals are translated into the languages of target markets. These language processing workflows have been well tuned to this domain by the various players in the value chain, such as the multinationals that are high volume generators of content requiring localisation and the Language Service Providers (LSPs) that provide outsourced localisation services, including the management of the translation of textual content. The business drivers in this industry produce workflows that are driven by the cost reduction needs of bulk publishers, resulting in little innovation into new business areas or applications.

Natural language technologies such as machine translation and text analytics are now maturing to a stage where they are being adopted as components in these workflows. However, they also offer the opportunity to broaden the localisation business into domains where the source content is less predicable and produced and consumed more rapidly in higher volumes by a wider range of users. The potential for innovation for the localisation industry exists in several directions;

- Outwards: addressing language as the next big barrier to be overcome in the use of the Internet for global communication and value generation
- Inward to focus on the need of the individual consumer through personalisation, i.e. the tailoring of the delivery of content, not only to the users locale but also to their personal content consumption preferences and their current physical, social and task context.
- Sideways into other corporate activities of existing knowledge- and service-intensive localisation clients, e.g. customer care and customer relations management or leveraging Web 2.0 technologies to engage with crowd-sourcing or open innovation value networks.

We identify such a shift and broadening in the localisation industry as *Next Generation Localisation*. This will involve making the workflows for linguistic processing and translation much more customer driven, rather than product driven as currently. It will require dealing with a much wider range of content sources, including user generated content and highly transient content that provides much of the value found in Web 2.0. It will also involve leveraging a wider range of linguistic human skills and value exchange models, beyond the scope of today's professional translators.

This presents a major challenge in systematically integrating fine-grained, on-demand quality into web content and web application localisation. This requires integrating mechanisms to determine and deliver quality, reliability and speed that match immediate user requirements into such web offerings. Though linguistic technologies allow us to automate some tasks, such as machine translation or entity recognition, the bounds in the confidence of the quality of outcomes needs to be understood and carefully managed. Key to this is empowering the user to assess that quality and demand more if required and indicate the level of quality they are willing to pay for in a given context. Content owners must then be able to adaptively tailor allocation of localisation resources (whether human or automated) to a wider and more dynamic range of quality targets.

To exploit the business opportunities of such Next Generation Localisation, industry must adopt a more flexible, extensible and lower cost mechanism for the integration of language processing workflows across many, increasingly specialised players. Service Oriented Architectures (SOA) offer a viable route to addressing this challenge. This paper outlines an Open Service Framework that is being developed by the Centre for Next Generation Localisation (www.cngl.ie) that harnesses the power of SOA to enable industry to react rapidly to changing business models and opportunities through service reuse and composition, extensible meta-data interoperability and flexible service and workflow management capabilities.

2. Background

The localization industry has already undertaken a number of separate document focussed standardization activities to support interoperability between different localisation applications. The Localisation Industry Standards Association (LISA – www.lisa.org) has developed various localisation standards:

- Translation Memory Exchange (TMX) for exchanging Translation Memory (TM) database content (TMX 2005). Such content is key in eliminating the re-translation of content segments that have previously been translated. TMs also support fuzzy matches, where translations of similar source segments can be considered by translators. Many TM tool providers have implemented support for TMX in their products.
- Term Base eXchange (TBX): XML Terminology Exchange Standard, to allow terminology to be exchanged between content author and translator tools (TBX 2008). An XML linking standard for terms, called Term Link, is also being investigated.
- Segmentation Rules eXchange (SRX), for exchanging the rule by which content is originally segmented. There has been very little support to date for SRX because segmentation is the main component that distinguished TM tools. Segmentation has direct consequences for the level of reuse of a TM. A TM's value is significantly reduced without the segmentation rules that were used to build it.
- Global information management Metrics eXchange (GMX): A partially populated family of standards of globalization and localization-related metrics

The Organization for the Advancement of Structured Information Standards (OASIS – www.oasis-open.org), which produces e-business standards has had a number of initiatives, the most notable being XML Localisation Interchange File Format (XLIFF 2008). XLIFF is the most common open standard for the exchange of localisable content and localisation process information between tools in a workflow. Many tool providers have implemented support for XLIFF in their products.

The W3C, which develops many core web standards, has an Internationalisation Activity (www.w3.org/International) working on enabling the use Web technologies with different languages, scripts, and cultures. Specific standardisation includes the Internationalisation Tag Set to support internationalisation of XML Schema/DTDs (ITS 2007).

To date, therefore, though file interoperability is supported in places, standard localisation processes and workflows and associated open interfaces addressing common interoperability issues have not yet been widely adopted. Outside of proprietary scenarios, digital publishers and service providers cannot easily integrate their processes and technologies and monitoring end to end process performance is extremely difficult. This implies lost business opportunities for many and missed opportunities for significant performance improvement for most of the stakeholders.

SOA coupled with workflow technologies are therefore well placed to address this lack of interoperability and end-to-end process management. Anecdotal evidence suggests that elements of the localisation industry were quick to consider the use of Web Service technology. In 2003 Bowne Global Services presented a case study (Reynolds) showing how they connected Interwoven's TeamSite Content Management System (CMS) to their in-house workflow engine (then named Elcano) using Web Services. IBM also presented a white paper discussing how web services and workflow management feature such as supported by their WebSphere product range could stream line the localisation process (Flinter 2003). However, these were focussed on integration within the enterprise, and end to end web service solutions have been slow to emerge, though several tools now make internal APIs available via Web Services for enterprise integration and support of custom client applications, e.g. for accessing TM content. Two examples of interfaces provided to human translation services are those provided by Translated.net and by Lionbridge to their Freeway system. Web Service interfaces to Machine Translation systems are more straightforward due to less branching logic and, as such, more common. Examples include the WebSphere Translation Server and the Google Translate API.

In 2007 the OASIS Translation Web Services (TWS) 1.0.3 draft specification (Reynolds 2007) was released with the aim of standardising the communication between translation providers and their clients (Reynolds 2003)(Bargary 2006). TWS remains the only real attempt to define web-services to support the end to end localization process. However, TWS has a limited scope. Rather than aiming to support the dynamic composition of language services into flexible localization workflows, it concentrates on supporting the negotiation of "jobs" between service providers. It is primarily intended to support the efficient out-sourcing of localization and translation jobs and it does not address the composition of language-services to form automated workflows.It is not clear to what extent this draft specification has found traction in industry to date.

3. Open Service Framework

4. Therefore, in order to deploy web-services to

support such composition, there is little standardisation to rely on. Thus, a first step in addressing the problem is to design a set of web-services and their interfaces suitable for the task. In designing these services, it is worthwhile to recall the general goals of service-oriented architectures; the services should be designed to be as flexible and general as possible and they should neither be tightly coupled to one another, nor to the overall system which they are part of. Furthermore, in keeping with the general trends in service designs (Foster 2008), variability in service behavior should generally be supported through the passed data-structures rather than through different function signatures.

Our ultimate aim is to establish a Unified Localisation Factory (ULF) that will enable future web content and service providers of all sizes instantiate localisation processes tailored to their needs and those of their customers. The ULF will allow future localisation-focussed applications that leverage advanced language and digital content management technologies to be rapidly integrated at low cost. This requires an Open Service Framework for presenting and assessing individual technologies, applications, evaluation techniques, design patterns, interoperability standards and workflows is a SOA.

This framework will consist of the following;

- Core Principles and Concepts: The core shared domain knowledge that characterise the vision of Next Generation Localisation.
- Process Map: A Business Level Reference Framework expressed using business process modelling concepts.
- Methods and Techniques: The procedural guidance needed to apply the framework, to evaluate that application and to contribute to the refinement of the framework in an open manner.
- System Services Architecture: The software system architecture needed to ground the application of the framework in operational software systems.
- Reusable Elements: Specifications, models, service definitions, APIs, software components and various forms of design patterns (e.g. for workflow, software integration, SOA etc) that can be used in a specific applications of the Open Service Framework

5. Next Generation Localisation Process Map

As the Open Services Framework aims to support interoperability across next generation localisation workflows consisting of multiple parties and their various services and applications, then ultimately it must support this through the definition of common meta-data. The benefits from the identification of common meta-data models in a particular domain are to provide the foundation for interoperability standards. Such standards thereby:

- Reduce cost of system integration
- Support multi-vendor system architectures, increasing the benefits of vendor competition by reducing lock-in for the different process actors
- Maximise the reuse of data and processes and the software services that underpin them.

However, localisation, in common with many application domains, possesses multiple stakeholders operating multiple systems in multiple interlinked business processes. These factors complicate efforts towards convergence and agreement on common industry-wide meta-data. Attempting a programme of meta-data modelling for interoperability standards therefore requires a common business-level reference framework in order to understand and discuss the different data and meta-data requirements at different interoperability points.

Other industries have successfully used a Process Map as a business level reference framework within which detailed business process definitions, and thereby specific interoperability models, can be worked upon within a shared set of terms and associated meanings. This requires an abstract process map that is not a reflection of any one company’s model and therefore provides neutral means for discussing shared interoperability concerns. Example of the use of such process maps in other industries are: the electronic Telecommunication Operations Map (eTOM) used by the TeleManagement Forum to support industry interoperation agreements between vendors of telecommunications management packages (Reilly 2009). Another example is the Smart Building Process map used to enable standardization of the exchange of data between CAD tools and building operations tools in the construction and facilities management industry (SmartBuilding).

Currently the Localisation Industry does not possess such an agreed process map, so as part of the Framework we propose a novel ‘Next Generation Localisation’ Process Map. The scope of this should be business processes covered by our broad vision of Next Generation Localisation, beyond conventional localisation workflows into areas of: crowd-sourcing, integrated language technologies such as machine translation, speech processing and text analytics that use statistical approaches; information retrieval; digital content management and personalization; web service development and governance. The NGL Process Map provides a top down common analytical frame within which specific business scenarios can be modelled. By overlaying specific business process flows of the process map we can start to identify where: existing standards such as XLIFF, TMX can be applied and if necessary extended and where new meta-data agreements are needed, the proposal of which is an activity within CNGL. The NGL Process Map therefore acts as a stakeholder-neutral medium to communicate requirements, seek solutions and contextualise the design and agreement of interoperability standards. Of course the structure of the Process Map itself will therefore influence the direction of such deliberation, so we remain open to proposals to refine this structure.

The structure currently proposed for the Process Map consists of an orthogonal grid of abstract Stakeholder Types ranged against shared Business Process areas. The Business Process areas represent recognisable collections of activities that span the localisation business process lifecycle and includes the processes related to process improvement. The areas can be individually populated with more specific processes, both for abstract business modelling and for capturing specific business scenarios. The Stakeholders differentiate the broader range of actor types involved in Next Generation Localisation, ranging beyond those just concerned with the generation and translation of content to include directly the consumer, online communities and software developers. These can in turn be specialised as niche stakeholders are identified or when applying the grid to a specific concrete business scenario. In this the process map can be used at both an abstract industry-wide level and for the analysis of specific business relationships and their interoperability requirements. The orthogonal structure breaks the domain down into a set of regions, the boundaries between which become the primary foci for capturing requirements. As the map is used as a frame for specific business scenarios and their associated process flows, solutions to interoperability issues at these boundaries can be collated and after review combined into broader interoperability models at the top levels from which future concrete scenarios spanning the same boundaries can select appropriate solutions.

In the current Process Map the two orthogonal axes are defined as follows:

Stakeholder Types:

- *Corporate:* This contains processes performed by organisations employing workers in a professional capacity. It encompasses any processes that are performed for monetary exchange, thereby encompassing public bodies and NGOs. Several specific sub-categories of this pool have been identified for the NGL domain: Content Generator; Language Service Provider; Translation Agency; Translation Sole Trader; Web Search Service Provider and Content Service Provider.
- *Consumer:* This contains processes conducted by the ultimate consumer of content. It is distinguished from other stakeholders in that it does not consume content for the purpose of providing it to other processes. Process for this stakeholder may annotate content to provide feedback to other processes, but only as a secondary activity to the consumption of that content.
- *Community:* This stakeholder represents processes that are subject to collective decision-making and content processing work performed by volunteers. It therefore excludes any activities performed for monetary reward directed to those performing it. The processes are distinct from those of the consumer in that they are indented to produce results of value to some other party and they are knowingly performed as part of a collaborative effort.
- *Service Developer*: This contains processes related to the development of new software services that can subsequently used by processes elsewhere in the process map. It excludes processes related to the localisation of that software, in such cases the processes should be conducted as part of the corporate stakeholder.

Business Process Areas:

- *Content Generation:* This includes authoring, internationalization and the development of terminology, domain models and content style guidelines.
- *Content Localisation:* Translating content from a source language to one or more target languages and making other locale specific changes to content.
- *Content Consumption:* The user driven consumption of content including search and personalsiation of content.
- *Content/Asset Management:* The collection, storage, refinement and general husbanding of reusable digital assets, e.g. TM, term-bases, guidelines, user models, annotations, quality assessments etc.
- *Process Management:* The process involved in monitoring, analysing and modifying business processes with the view to improving performance metrics.

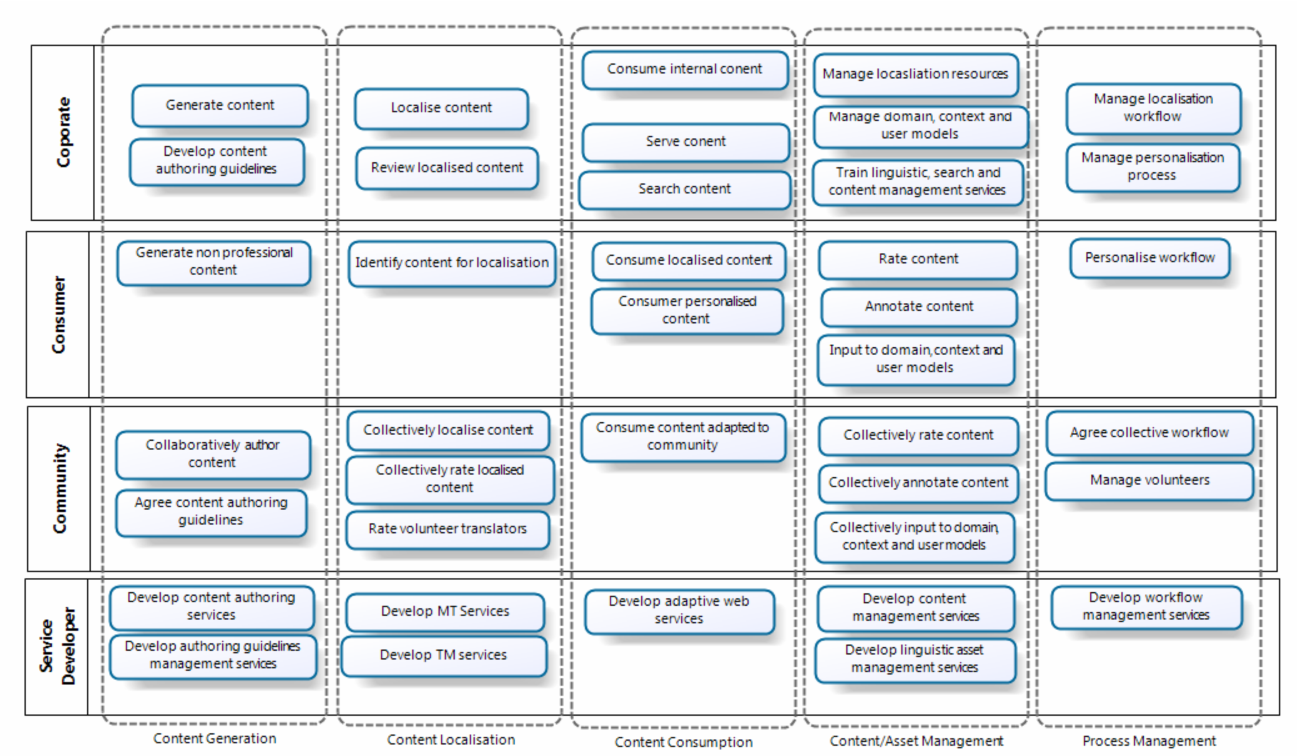

Figure 1: Next Generation Localisation Process Map high level processes in each region

Figure 1 outlines a set of processes that we have identified as populating the process map. The significance of this model is the emphasis given to activities beyond the traditional corporate work-flow, highlighting the important of managing the dynamic relationship with the consumer of content; of leveraging the collective intelligence of online communities and integrating the software service developer into SOA-based process improvement.

6. Language Processing Services Architecture

Adoption of an SOA advocates software integration through well defined functional interfaces that can be invoked remotely, typically using the Web's HTTP protocol with input and output parameters encoded in XML. The W3C have standardized an XML format, The Web Service Description Language (WSDL), for describing and exchanging such service definitions. Web services can be composed into more complicated applications using explicit control and data flow models that can be directly executed by workflow engines. This allows new workflow applications to be defined declaratively and immediately executed, thus greatly reducing the integration costs of developing new workflows and increasing the flexibility to modify existing ones. Such web-service based service composition is known as Web Service Orchestration. OASIS has a standardized web service orchestration language called the Business Process Execution Language (BPEL), which has resulted in the development of several commercial execution platform and BPEL workflow definition tools, which support workflow definition through drag-and drop interfaces. This approach has already been by the LanguageGrid project (Inaba 2007) for the rapid development of linguistic applications by defining BPEL orchestration of web services. These services offer access to language resources provided in a mutual manner by different academic and research organisations. Resources include parallel text, cross-lingual dictionaries, machine translators and morphological analysers. In our prior work in this area (Lewis 2008) we used BPEL composition of a machine translation service and a language identification service to integrate more flexible content handling into WorldServer, a well established localisation workflow management product from SDL Inc (www.idiominc.com). This work highlighted how linguistic processing services for localisation workflow can be readily abstracted into services that take a source language segment and either adds target language segment, sort target segments or annotate segment pairs. From these a wide range of specialised linguistic services can be derived and composed to address linguistic processing needs for localisation. However, the simplicity of this linguistic processing service taxonomy does not reflect the need to configure and train the systems underlying these services, this being the processes where core value is derived. Further, as modern linguistic processing is increasingly statistical, the monitoring of statistical performance analysis over the various source language content flow being localised becomes a vital part of the processes. For this reason we add to the core linguistic processing service interface taxonomy, two parallel abstract interface types:

- A Service Configuration interface: via which the component can be configured to operate in the desired manner, e.g. by providing a domain trained statistical model to a Machine Translation service component
- A Service Monitoring and Logging interface via which operational data about the performance of the component can be remotely monitored or locally logged. This interface has generic operations for the configuration of the behaviour of monitoring and logging, e.g. producing event on a threshold being reached.

These parallel interfaces are seen a essential to developing web services for Next Generation Localisation due to the key role played by ongoing process monitoring in the design and deployment of new, improved processes, tailored to a wider range of business, social and individual consumer needs. The definition of these interfaces may be implemented as simply as dedicated operations on a web service, though ongoing work of the Service Component Architecture group at OASIS (www.oasis-opencsa.org) and the web service activities at W3C (www.w3.org/standards/webofservices/) promise a more standardised mechanism for assembling web service components with multiple interfaces.

Figure 2 gives an indication of the different combinations in which systems can be assembled in accordance with the service architecture as a range of service components deployed and accessed from a variety of client applications. These could range from service invocations made by existing Globalisation Management Systems (GMS) and Computer Assisted Translation (CAT) software. Platforms such as WorldServer and Trados from SDL and Catalyst from Alchemy already have extension APIs that allow invocation of third party services. A communication bus between the service components and the client applications based on WSDL/SOAP providing the best operational support for workflow-based clients, such as BPEL web service orchestration engines, which can then explicitly define fault and compensation handling workflow branches. However it is recognized that many applications may be better suited to RESTful service invocation models typical of web mash-ups, e.g. in JavaScript web browser clients or PHP web server modules. This mode of operation can also fit naturally with invocations from web application platforms such FaceBook and Twitter or for invocation from third party mobile software clients such as iPhone Apps, however in these cases WSDL client access is also commonly available.

The service components currently being developed in the CNGL divide broadly into those that provide linguistic processing and those providing on-demand personalised information access information to multilingual content. The lingistic processing services such as Machine Translation, Speech Synthesis and Recognition and Text Analytics provide value that depends on statistical training over large volumes of data. The personalisation services adapt content to particular user preferences and context, based on rich meta-data assembled about the user, the context and the setting of the interaction. In both cases, the value of the service depends on appropriate configuration, which must therefore become a major element of any progressively improving workflow or application. At the same time, it must be acknowledged that statistical language processing and meta-data driven adaptation will never deliver complete accuracy in all cases, so integrated operational monitoring is needed to support ongoing reconfiguration of individual components through statistical retraining or improved meta-data modelling (which itself may rely on statistical techniques for semantic annotation of content or social network analysis of user activity). Therefore to support operational monitoring and process improvement for end-to-end workflows that span our NGL stakeholder types, shared data-model for monitoring and configuring the operation of the various server components, via the parallel interfaces identified above, will be essential.

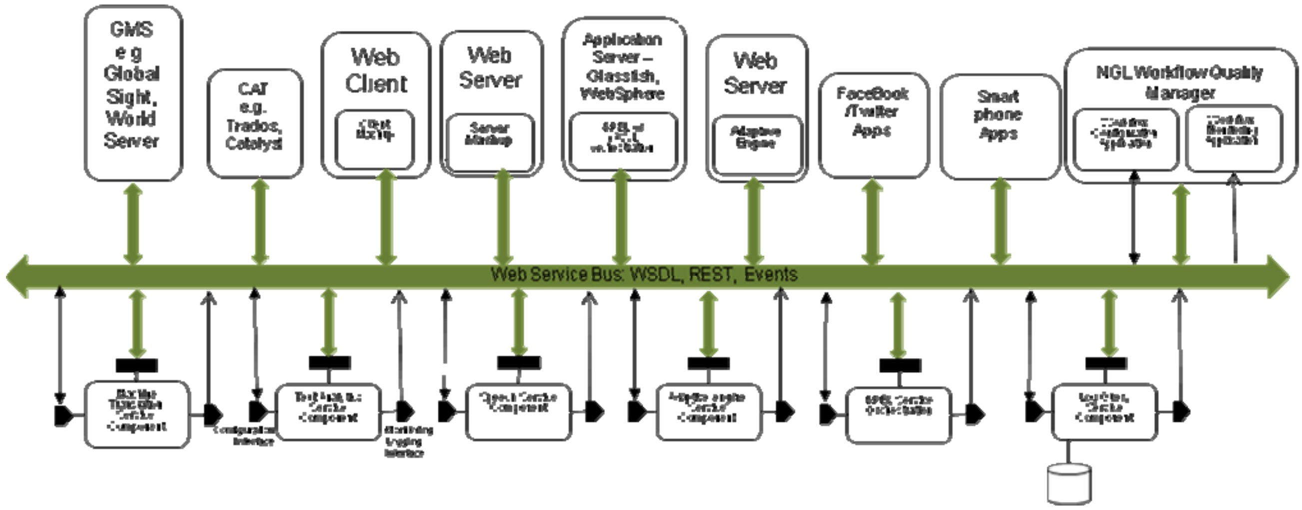

Figure 2: Potential configurations of the NGL Open Services Architecture

7. Example Next Generation Localisation Processes

In the CNGL project, we are applying this Framework as we develop a series of demonstrator systems that bring together advanced technologies from the Centre's academic partners and the current real world needs of its industrial partners or of broader real-world applications.

In relation to advancing the current localisation workflow with statistical machine translation and text analytics service, figure 3 shows an idealised reference workflow in the Business Process Modelling Notation that we are using to explore this area. It highlights how any instance of this workflow can exploit multiple instances of component sub-workflows or activities. The importance of intelligently selecting between these instances at various points in the workflow are then highlighted (the diamond, star-filled BPNM icon for a complex gateway decision). So for example a content publisher may select different instances of service providers who can perform the Translation Job sub-workflow. Within this, the service provider may divide the job and select different Translate Section sub-workflows, which employ different Translation Memories and then themselves make intelligent decisions about selecting between different Machine instances and post-editors on a segment by segment basis. After reassembly, the content provider may again select between different providers to perform the Review Job sub-workflow. Key to the aim of developing mature, optimising and transparent processes, as well as publishing the result, the quality checked translations and other process monitoring data are fed back passed to a process. One feedback path already in common practice assembles Translation Memories for future use. Another uses them to train future MT instances, potentially by grouping content by domain or style to get more accurate results with more computationally efficient MT engines. Also, feedback may be provided on problems encountered with terminology and content consistency. By supporting these steps with web services and using web service orchestration, configuration and monitoring, such decision making can become highly dynamic, itself being driven by statistical analysis of the content against domain categorisations. This accompanying reduction in process management overhead means this process can be conducted in a wider range of scenarios, beyond transitional bulk translation scenarios. For instance, cheap, or zero cost MT coupled with suitably motivated crowd-sourcing for post-editing and reviewing tasks may move the value of the process away from translation and towards the intelligent husbanding of TMs and training of MTs, allowing niche operators who combine linguistic and natural language technology skills to emerge.

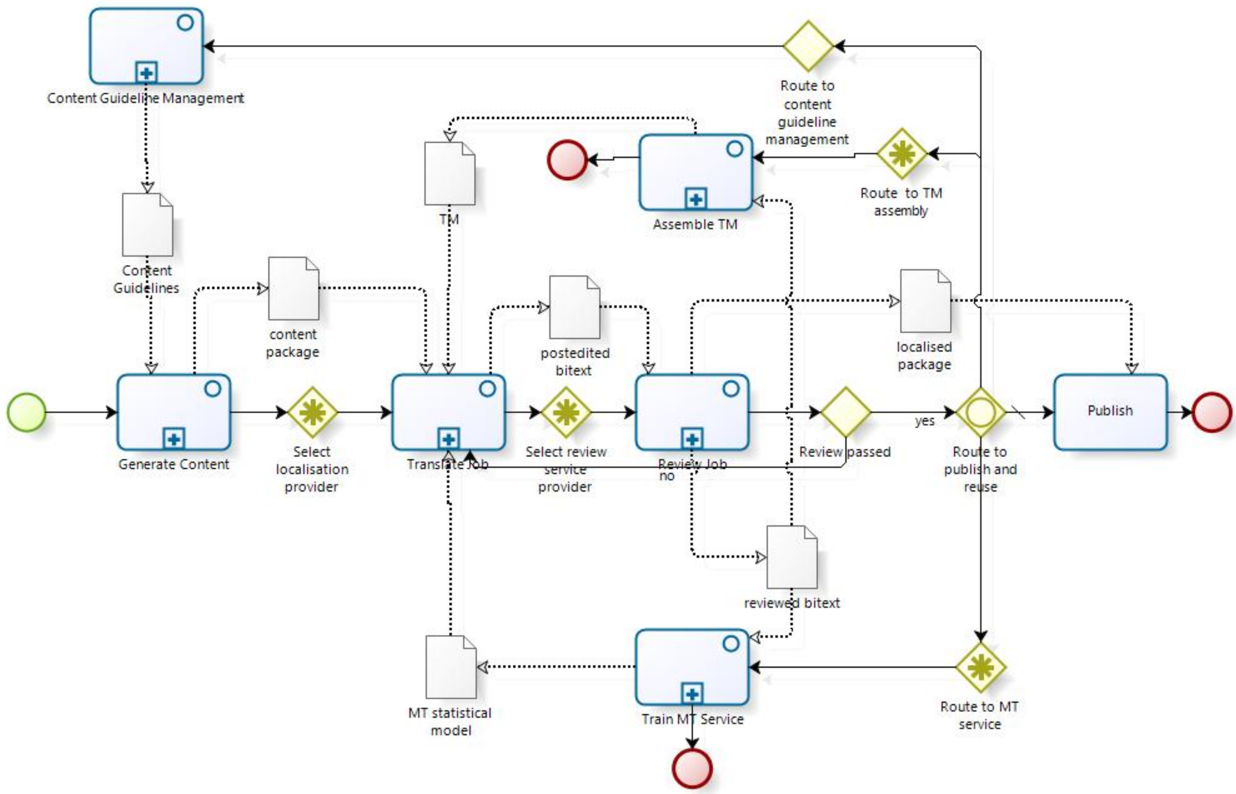

Figure 3: Generic Bulk Localisation Process Flow showing process feedback loops

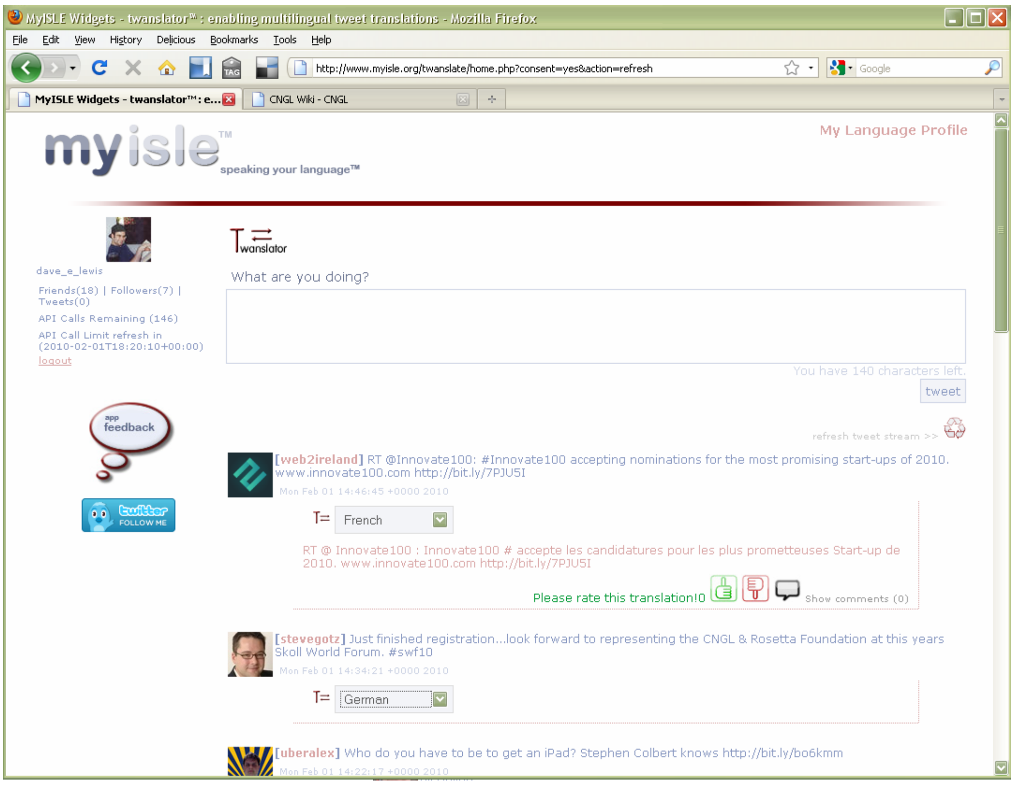

Figure 4: Screen shot of MyIsle Twanslator App at http://www.myisle.org/twanslate

By way of example consider a new application developed in CNGL to demonstrate the integration of MT and crowd-sourcing in support of multilingual Twitter consumption. This takes Twitter feeds followed by a user and passes them through, first a language identification service, which then routes this to a language specific MT engine. This reuses a BPEL sequence that combines a text analytics web service for language identification and a web service wrapper for the Matrex MT service, trained in specific language pairs. Though this orchestration was originally developed to demonstration their support for conventional localisation workflows in the WorldServer

GMS (Lewis 2009), this application allows the same technology to be integrated into a crowdsourced post-editing setting. Here users are encouraged to rate and if willing to post-edit the machine translated Tweets, providing input to further training of the MT (see screenshot on figure 4). A simple initial gauging of the user's language competence allows this rating and post-editing dialogue to be personalised to the user and also to their willingness to participate in these steps. Though simply assembled in a few days as a mash up of the Twitter API and invocation of the BPEL web service orchestration, this application can now be easily scaled to a fully managed service that can be revised and tuned over time at low cost. Such Twitter applications therefore quickly enable the study of a number of issues key to NGL, i.e. collective content annotation, rating and translation (crowds-ourcing) across a social network; machine translation of perishable content and short text form content; social network informed personalisation of content querying and content translation and automated semantic annotation based on domain personalisation and text analytics.

8. Conclusion

We have described an Open Services Framework that we are developing to enable a broad range of Next Generation Localisation. We have emphasised the need for common meta-data to support web service interoperability as well as for the configuration and monitoring of systems via web services interfaces. The proposed NGL Process Map provides a semantic, process-oriented frame for discussing such meta-data agreements as new forms of NGL processes encompassing the broader set of stakeholders is explored. Our further work will involve expanding the range of example processes, including application in personalised multilingual customer care and social networking. Web service interfaces and associated meta-data will be harvested and common models proposed. Web service definition structures will be aligned with emerging service component architectures, including support for access control and business rules. Where meta-data requirements are volatile, we will adopt triple-base models based on the W3C's Resources Description Framework (www.w3.org/RDF) to allow rapid refinement of data models. We will also explore the deployment of compute-heavy processes such as MT training onto cloud computing environments.

Acknowledgements

This research is supported by the Science Foundation Ireland (Grant 07/CE/I1142) as part of the Centre for Next Generation Localisation at Trinity College Dublin.

References

Bargary, K. (2006) "Ignite implements Web Services", Localisation Focus, vol. 5, Issue 1, Pages 22-23, Ireland, 2006

Lewis, D., Curran, S., Feeney, K., Etzioni, Z., Keeney, J., Way, A., and Schäler, R. (2009). *Web service integration for next generation localisation*. In Proceedings of the Workshop on Software Engineering, Testing, and Quality Assurance For Natural Language Processing (Boulder, Colorado, June 05 - 05, 2009). ACL Workshops. Association for Computational Linguistics, Morristown, NJ, 47-55.

Flinter, S. (2003) "A J2EE based Localization Services Architecture", IBM Whitepaper, Jan 2003, accessed from ftp://ftp.software.ibm.com/software/globalization/documents/j2ee_lsa.pdf on 25 Feb 2010

Foster, I., Parastatidis, S., Watson, P., and Mckeown, M. 2008. How do I model state?: Let me count the ways. Commun. ACM 51, 9 (Sep. 2008), 34-41.

SmartBuilding "Information Delivery Manual Guide to Components and Development Methods", SmartBuilding Norway, accessed from http://idm.buildingsmart.no/confluence/download/attachments/446/IDM2_Methodology.pdf?version=1 on 25 Feb 2010

Inaba, R., Murakami, Y., Nadamoto, A., Ishida T. (2007) "Multilingual Communication Support Using the Language Grid" in Intercultural Collaboration, LNCS 4568, Springer, 13 Aug 2007, pp118-132

Internationalization Tag Set (ITS) (2007) Version 1.0 W3C Recommendation 03 April 2007, accessed from http://www.w3.org/TR/its/ on 25 Feb 2010

Johnson, D. (2007) "Getting Started in Localization", Multilingual, Oct/Nov 2007, pp 3-5

Reilly, J., Kelly, M. (2009) "The eTOM - A Business Process Framework Implementer's Guide", TM Forum

Reynolds, P. (2003) "Web Services for Translation", accessed from http://www.translationdirectory.com/article395.htm on 25 Feb 2010

Reynolds, P. et al (2007) "Translation Web Services - draft committee specification" OASIS Translation Web Services TC, Proposed draft specification 16 May 2007, access from http://www.oasis-open.org/committees/download.php/24350/trans-ws-spec-1.0.3.html on 25 Feb 2010

Systems to manage terminology, knowledge, and content – TermBase eXchange (TBX), Localization Industry Standards Association, 29 Oct 2008, access from http://www.lisa.org/TBX-Specification.33.0.html on 25 Feb 2010

TMX 1.4b Specification OSCAR Recommendation, Localisation Industry Standards Association, 26 April 2005, accessed from http://www.lisa.org/fileadmin/standards/tmx1.4/tmx.htm on 25 Feb 2010

XLIFF "A white paper on version 1.2 of the XML Localisation Interchange File Format (XLIFF)", Revision: 1.0, 17 Oct 2007 accessed from http://xml.coverpages.org/XLIFF-Core-WhitePaper200710-CSv12.pdf on 25 Feb 2010

XLIFF Version 1.2, OASIS Standard, 1 February 2008, accessed from http://docs.oasis-open.org/xliff/xliff-core/xliff-core.html on 25 Feb 2010