META≡NET

A Network of Excellence forging the
Multilingual Europe Technology Alliance

Vision Document

Vision Group Translation and Localisation

Results of first two meetings

Editors:	Aljoscha Burchardt, Georg Rehm
Dissemination Level:	Public
Date:	3 December 2010

Grant agreement no.	249119
Project acronym	T4ME Net (META-NET)
Project full title	Technologies for the Multilingual European Information Society
Funding scheme	Network of Excellence
Coordinator	Prof. Hans Uszkoreit (DFKI)
Start date, duration	1 February 2010, 36 months
Distribution	Public
Contractual date of delivery	n.a.
Actual date of delivery	n.a.
Deliverable number	n.a.
Deliverable title	Vision Document- Vision Group Translation and Localisation: Results of first two meetings
Type	Report
Status and version	Draft
Number of pages	19
Contributing partners	DCU
WP leader	DFKI
Task leader	ILSP
Authors	Aljoscha Burchardt, Georg Rehm using Input from all members of the Vision Group "Translation and Localisation" listed in Section 3.2.
EC project officer	Hanna Klimek
The partners in META-NET are:	Deutsches Forschungszentrum für Künstliche Intelligenz (DFKI), Germany
	Barcelona Media (BM), Spain
	Consiglio Nazionale Ricerche – Istituto di Linguistica Computazionale "Antonio Zampolli" (CNR), Italy
	Institute for Language and Speech Processing, R.C. "Athena" (ILSP), Greece
	Charles University in Prague (CUP), Czech Republic
	Centre National de la Recherche Scientifique – Laboratoire d'Informatique pour la Mécanique et les Sciences de l'Ingénieur (CNRS), France
	Universiteit Utrecht (UU), Netherlands
	Aalto University (AALTO), Finland
	Fondazione Bruno Kessler (FBK), Italy
	Dublin City University (DCU), Ireland
	Rheinisch-Westfälische Technische Hochschule Aachen (RWTH), Germany
	Jozef Stefan Institute (JSI), Slovenia
	Evaluations and Language Resources Distribution Agency (ELDA), France

For copies of reports, updates on project activities and other META-NET-related information, contact:

Table of Contents

1 Executive Summary

Language industries will contribute a considerable percentage of Europe's future economic growth. META-NET focuses on understanding the particular goals and constraints of research, user industries, provider industries, administrations and government, surveying the current state and identifying the gaps with the goal to create a shared vision. To this end, it operates three Vision Groups.

The industry and research sectors discussed in the first two meetings of the META-NET Vision Group "Translation and Localisation" are facing fundamental changes if not paradigm changes in the next few years. External societal and technological factors such as new mobile applications or globalised workflows and cross-border online communication put new demand on innovative solutions. At the same time, free online services increase the price pressure while the possibilities for affordable quality improvement have reached a deadlock.

This document is the outcome of two whole-day meetings among top European industry stakeholders and researchers that discussed needs and visions, opportunities, and threats. This document is organised as follows. Section 2 introduces the topic and section 3 describes the Vision Group constitution and meetings. Section 4 presents the central needs and visions that have been discussed, organised into domain-specific content and general needs and visions that also apply to other areas of Language Technology. To give a first idea of the content, the headlines of the sections on needs and visions are listed below:

Needs

- Need #1: Humans (content providers, translators/post-editors) need to be better prepared for and included in translation technologies and their development.
- Need #2: A technological breakthrough in Machine Translation.
- Need #3: Translation has to have greater respect for context and translation adequacy must be judged in context.
- Need #4: Future-proof and transparent business models and translation workflows.
- Need #5: Security must be guaranteed throughout the whole development chain.
- Need #6: Interoperability regarding tools and data as well as configuration and personalisation.
- Need #7: An ecosystem that concerts R&D activities as well as economic cooperation.

Visions

- Visions 1: New Workflows
- Visions 2: New Architectures

- Visions 3: New Applications
- Visions 4: New Modes of Research and Training
- Visions 5: New Computing Models, Business models, and Standards for Sharing Knowledge Resources

2 Introduction

2.1 The META-NET Vision Building Process

A central objective of META-NET is the preparation of a major concerted effort geared towards the creation of the needed technological foundation for the European multilingual information society. An essential instrument to this end is the forging of a strategic alliance involving, in addition to the top level R&D centres, the active participation of European LT and ICT industry and many private and public stakeholders, including the language communities themselves.

The Vision Groups are a central instrument within META-NET. Each of the three groups brings together researchers, developers, integrators and (actual or potential, corporate or professional) users of LT-based products, services and applications. The goal of the groups is to generate domain-specific visions and roadmaps in the form of technology forecasts. This includes ideas for innovative applications of language technology and scenarios for the future knowledge society which can be supported by advanced technology. The visions produced will be gathered by the Technology Council which will consolidate them into a Strategic Research Agenda. The Agenda will contain high-level recommendations and suggestions for joint actions to be presented to the EC and national as well as regional bodies. The three Vision Groups are:

1. Translation and Localization
2. Media and Information Services
3. Interactive Systems

The Vision Groups are scheduled to meet twice a year. In 2010, two rounds of meetings have been successfully completed. Their output will be discussed at the 1st Technology Council Meeting (16/11/2010), which, in turn, will prepare the draft SRA. Preliminary findings of the Vision Groups will be presented at META-FORUM 2010 (17/11/2010). Additional meetings will take place in 2011, the goal being to provide input to the SRA draft and to further elaborate the visions.

This document is intended to provide a distillation of ideas, opinions and visions expressed by members of the Vision Group Media and Information Services during its first two meetings. It aims at serving as a basis for discussion on the challenging and innovative LT based

scenarios that need to be addressed by 2020 and providing seed ideas for the drafting of the SRA to the Technology Council.

2.2 Machine Translation and Localisation

Human-quality machine translation (HQMT) opens new doors by enabling innovative application scenarios and use cases in areas like document processing, information services, translation and localisation industry, language services, mobile applications, education and e-commerce. Unfortunately, however, the automatic, fully software-based translation of arbitrary texts on a level that matches that of a human expert translator does not exist yet.

More than 60 languages are spoken in Europe. Overcoming linguistic borders and language barriers by means of translation and localisation are, therefore, central challenges for the continent's goal of a digital market in which every European citizen can participate in his or her own mother tongue. At the core of the demand side is instant translation[1] of any type of content (written texts and documents, spoken language, etc.) into at least all European languages in a very high quality and also at a reasonable price.

The situation on the supply side is the exact opposite. Machine Translation (MT) looks back on roughly 40 years of research concerning the rule-based paradigm (RBMT) and about 20 years in the statistical paradigm (SMT). Quality-wise, a plateau seems to have been reached. Huge efforts are needed to offer slightly higher quality for a particular language, domain, or customer. At the same time the business models of translation companies are seriously challenged by freely available online services such as Google Translate. Although Google is not a Language Technology provider, their access to data, sheer computing power, and human resources make it possible to offer a free service with a considerable performance in general domains.

An issue that is specific to translation is the relation between human translators and MT technology. Traditionally, many translators have viewed translation (rightly so) as a highly creative process, almost as an art, rather than as a service provision that makes sure that the information, entertainment or cultural value that lies in the original product is transformed appropriately into the target language and cultural context. Consequently, any technology that comes even close to taking over part of the creative process is considered a serious threat. This is changing, as more and more translators are beginning to accept and use technology support including MT in translation work flows.

In Section 4, we will present the Vision Group's analyses and visions in the aforementioned areas (society, technology, economics) in much more detail.

---

[1] For brevity, in this document, we will use "translation" in a wider sense covering also localisation, different modalities, etc.

3 Vision Group Translation and Localisation

The Vision Group Translation and Localisation is the result of merging the two originally foreseen groups "Text Translation" and "Software and Content Localisation". Early on in the process of assembling lists of members this merger was decided in the META-NET Standing Committee due to a significant amount of overlap with regard to the suggested members. The Vision Group covers a comparably large area within Language Technology, both in terms of technologies and applications. The focal point is Machine Translation. According to the original plan, the group covers the following:

- **Fields**: Technical documentation, consumer information, official bulletins, user interface localisation, games, translation/localisation services, information gathering, etc.
- **Stakeholders**: Software companies, game companies, large users of translation/localisation, translation companies, localisation industry, language service providers, etc.

3.1 Recruitment Process

The recruitment process started by collecting suggestions for Vision Group members among members of the META-NET consortium and the external members of the META-NET Standing Committee. Out of these suggestions the Standing Committee generated a shortlist taking into account aspects such as coverage of topic areas, regional coverage, having a good balance of internal and external partners as well as industry vs. academic participants. A target number for participation in the first set of meetings was set to 10-15 external members. Invitations were sent out both electronically and by traditional mail. Feedback was collected by email and by calling the invitees after the deadline for feedback had passed. Below we list the key figures concerning the recruitment process of the Translation and Localisation group (listing external members that are not in the META-NET consortium only):

- Initial member suggestions: 49
- Shortlist (invitations sent out):   27
- Confirmations: 21
- Refusals: 1
- Meeting attendants (cumulative):   14

Concerning META-NET's ambitious goal of establishing a strong community within only a few months time, the organisers of the group are pleased with the positive feedback and high number of external participants that the conveners have been able to mobilise and who joined the group at rather short notice.

3.2 Meetings

The Vision Group Translation and Localisation met twice in 2010:

- 1st Meeting: July 23rd, 2010 at DFKI, Berlin, Germany.
- 2nd Meeting: September 29th, 2010 at Hotel Le Plaza, Brussels, Belgium.

The participants of the meetings are listed in the table below.

	Name	Organisation	Country	Sector/Role	1st meeting attendance	2nd meeting attendance
1	Andrew Bredenkamp	acrolinx	Germany	Authoring support	☒	☐
2	Aljoscha Burchardt	DFKI	Germany	META-NET[2], minute taker	☒	☒
3	David Filip	Moravia Worldwide	Czech Republic	Translation and Localisation	☒	☐
4	Stefan Geissler	Temis	Germany	Semantic Text Technologies	☒	☐
5	Josef van Genabith	DCU/Centre for Next Generation Localisation	Ireland	META-NET, convener	☐	☒
6	Daniel Grasmick	Lucy Software	Germany	Machine Translation	☒	☐
7	Jan Hajic	Charles University, Prague	Czech Republic	META-NET	☒	☐
8	John Judge	DCU	Ireland	META-NET, minute taker	☐	☒
9	Martin Kay	Stanford University/ U. des Saarlandes	Germany/ USA	Machine Translation Research	☒	☐
10	Stefan Kreckwitz	Across Systems	Germany	Translation Workflow Support	☐	☒
11	Gudrun Magnusdottir	ESTeam AB	Sweden	Machine Translation	☒	☒
12	Elisabeth Maier	CLS Communication AG	Switzerland	Language Services	☒	☐
13	Jörg Porsiel	VW	Germany	Translation, rapporteur	☒	☒
14	Artur Raczynski	European Patent Office	Germany		☐	☒
15	Georg Rehm	DFKI	Germany	META-NET	☒	☒
16	Johann Roturier	Localisation Department, Symantec	Ireland	Localisation	☐	☒
17	Svetlana Sokolova	ProMT	Russia	Machine Translation	☒	☐
18	Volker Steinbiss	RWTH Aachen	Germany	Machine Translation Research, rapporteur	☒	☒
19	Lori Thicke	Lexcelera; Translators Without Borders	France	Translation	☐	☒
20	Gregor Thurmair	LinguaTec	Germany	Translation, Localisation, Globalisation	☒	☐
21	Hans Uszkoreit	DFKI	Germany	META-NET, convener	☒	☒

Table 1. List of Vision Group Translation and Localisation meeting participants

The first meeting was intended to achieve mutual understanding concerning the goals of both META-NET and especially this Vision Group, to define the sector, collecting needs and topics to be discussed in depth later on, and discussing the appropriateness of the group's

---

[2] META-NET members are from Language Technology research if not indicated otherwise.

Technologies for the Multilingual European Information Society
Vision Group Translation and Localisation: Results of first two meetings

META=NET

composition. No input was requested by the participants in advance to ensure the greatest possible open-mindedness. The meeting was structured into the following three sessions:

▪ Session 1: Mutual understanding

- • Introduction of participants

- • Gaps, needs and dreams

▪ Session 2: Irritation and Provocation

- • Time horizon and scope of the visions

- • Your visions

▪ Session 3: Anticipation

- • Drafting of visions

- • Additional names end expertise

- • Preparation of next meeting

For the second meeting, written input by the Vision Group members was requested in advance, in the form of responses to the following three questions:

A. What are attractive, plausible, powerful, innovative language technology scenarios (applications or a combination of applications, including use cases) that in your opinion could be realized through massive concerted research, development, and innovation actions?

B. Can you think of novel research paradigms (approaches and modes) that could support real breakthroughs of type A?

C. What are expected social, technological or economic developments that have to be considered as prohibitive or enabling factors in the planning of A and B?

The second meeting was structured into four sessions:

▪ Session 1: Introduction and wrap up of the first meeting

▪ Session 2: Participants' input (questions A-C above)

▪ Session 3: Synthesis and priorisation of seed ideas for public discussion

▪ Session4: Preparation of report

At the meeting, nine short contributions were presented by external members, five of which were sent to the organizers in written form after the meeting.

3.3 Coverage

The composition of the group roughly follows the breakdown into stakeholder groups in the vision group planning phase and is considered to be balanced by the group members themselves. The only deviation from the original plan is the absence of representatives from the games industry. In fact, a number of personal discussions revealed that the localisation and

Technologies for the Multilingual European Information Society
Vision Group Translation and Localisation: Results of first two meetings

META-NET

translation needs in computer games are handled mostly by human translators as the number of sentences to be translated is usually not too large and, more importantly, for pieces of discourse that are quite short, it would be extremely difficult to transport the exact message, mood and style using MT.

One area that should probably be represented in future meetings is that of e-learning. This need was formulated in a discussion at the second meeting, where the idea was brought up that e-learning may turn out to be a good use case for the language and translation technologies and applications under consideration. Also, human translators should be included in an extended feedback process, e.g., via the online discussion forum at http://www.meta-net.eu.

4 Visions on challenging and innovative LT-based scenarios

4.1 Domain-specific needs

1. **Translation is an emotional topic that needs to be handled with care**. In contrast to other LT applications, translation involves at least two stakeholder groups that have strong views with regard to the material to be translated: the content providers and the human translators/post editors. In the case of content providers, the main interests lie in the proper function of the translation (is it accessible by all that need it; does it adequately transport the meaning, entertaining and cultural value, etc.?).[3] The more the translation process is automated, the more control can and should be given to the content providers, who are often the experts on the subject matter.

   As regards translators, personal discussions and also public discussions at several META-NET events, such as META-NET panel discussions, have shown that many translators consider themselves artists closely engaged in a creative process rather than as providers of a linguistic service. Moreover, to date most translators have not been trained for using Language Technology in translation workflows, let alone for doing post-editing. These aspects have serious consequences for the future improvement of technology, workflows, and training.

   Need #1: Humans (content providers, translators/post-editors) need to be better prepared for and included in translation technologies and their development.

2. **Despite technological progress, the goal of high quality MT has not been reached**. Quality is a central issue. As has been said in the introduction, current MT technology is already quite mature, but a dramatic quality improvement cannot be

---

[3] Depending on the confidentiality, data security (trust) becomes another major concern (see Section 4.2).

expected. A problem seems to be the Zipfian distribution of issues in MT: some improvements can be realised with moderate effort, but unreasonably many resources are needed for small improvements in the long tail. This is a problem in particular for SMEs given their limited resources.[4]

Need #2: A technological breakthrough in Machine Translation.

3. **Translation is not only a linguistic business**. Context and common sense knowledge play a decisive role in providing a situationally adequate and personalised translation. At the same time, the quality/value of a translation must not only be assessed solely in terms of lingustic considerations. For example, a translation might be slightly ungrammatical, but can still be very helpful for quickly gathering information, e.g., about a company profile. The main value is often in the information, entertainment, etc., and not necessarily in an exact translation.

Need #3: Translation has to have greater respect for context and translation adequacy must be judged in context.

4. **Translation workflows and business models are not up-to-date.** Translation is, as mentioned above, a complex process. The prevailing pipeline architecture (content generation – translation – deployment) is not flexible enough to adjust to today's and even more so to future needs. On the one hand, too much is translated without a real need already today (e.g., technical documentation, language-variants of detailed manuals). A lot of money is thus wasted non-productively. At the same time, the need for translation is rising, e.g., in context of cross-border e-business. Whole areas such as Web 2.0 content (discussion fora, user-generated videos, etc.) are currently not translated at all.

Different demands require customised solutions and have special requirements, e.g.,:

- **Outbound high quality translation** (corporate, manuals, PR material, etc.): not time critical, quality matters

- **Inbound cross-lingual information gathering**: trust and data security are important

- **Instant translation for private users** (online content, communication): must be quick and cheap, should be personalised, quality is not the main issue

It is by no means clear who will pay what for instant or on-demand translations with or without human correction loops. Business models are lacking and so is proper expectation management.

---

[4] The connected problem of too little interoperability will be discussed in 4.2.

Need #4: Future-proof and transparent business models and translation workflows.

4.2 Domain-independent needs

We expect a certain amount of overlap in the domain-independent needs formulated in the three Vision Groups. This is why we list domain-independent needs only briefly here, relating them to specific challenges in translation and localisation.

5. **New functional requirements**. A general impression is that many solutions are driven by the supply side (what is technically/technologically possible today) rather than by a proper analysis of what is needed by end-users and business users. Important new requirements that emerge from analysing user needs include:

▪ Trust management (clear terms of conditions, pricing models, confidentiality, etc.)

▪ Ensuring data security (Where does the computation take place? How are corpora, terminologies, and translation models transported and stored?)

Need #5: Security must be guaranteed throughout the whole development chain.

6. **Interoperability**. As in other areas of Language Technology, data and software migration and reuse is difficult in MT. What is needed:

▪ Standards for creating data (language models, terminologies, background information such as domain ontologies, corpora) with the goal of sharing, including:

    • Metadata

    • Resolved copyright and legal issues (typically cross-border in translation), e.g., who owns the machine translation?

    • Technical solutions for integrating different MT paradigms (using Translation Memories, statistical, rule-based, template-based, etc.), which is even more important in the light of the recent trend of hybrid translation.

▪ A platform for standardisation efforts (a successful example from the Information Extraction community is UIMA, but much more is needed here)

Another interoperability problem in translation is that solutions that work for one text type often do not carry over to other cases.

Need #6: Interoperability regarding tools and data as well as configuration and personalisation.

7. **R&D and business cooperation**. One question that was discussed in all Vision Groups is how the division of labour between SMEs, larger companies, and providers of

free software and services such as universities and, e.g., Google should develop in the future. It seems that the number of Language Service Providers (LSPs) is increasing while the effort spent for adaptation of core language technology is reduced. A more recent challenge is adaptability to changing data (e.g., user-generated content, patents).

Need #7: An ecosystem that concerts R&D activities as well as economic co-operation.

8. **Language awareness.** Raising language awareness among pupils and students, i. e., pointing out to the younger generation the importance of (one's native) language in general as well as the benefits of (e-)learning a foreign language in particular is an important educational goal. By being able to speak another language than one's mother tongue, people will have easier access to more information and will also be able to communicate with people from around the world and share knowledge and views.

Need #8: (Educational) measures to raise language awareness among pupils and students.

4.3 Domain-specific visions

The domain-specific visions are subsumed under three headlines: New Workflows, Architectures, and Modes of Research and Training.

1. **New Workflows (re. Needs #1-4).**

   In the Vision Group discussions the experts agreed that many (if not most) problems with MT arise from poor quality of source texts to be translated including spelling and grammar errors, ambiguity, missing context, incoherent use of terminology, etc. One example of a novel workflow to optimise the translation process is **going upstream**, i.e., linking the translation process tighter to content production. The concrete vision is intelligent authoring support not so much in the sense of controlled language, which has been tested with limited success in the past, but rather in the sense of anticipating the problems that can arise in the translation process. A straightforward example is the resolution of an anaphora such as "it":

   Usage example: While authoring a document (to be translated), the envisioned authoring system would notice that the anaphor "it" in *after opening the compartment holding the battery, remove it* is ambiguous and would, e.g., highlight both "battery" and "it" and prompt the author to confirm that "it" here refers to the battery. This information would then be stored in an external record of context metadata to be used later on for disambiguation in the translation process. The same interface could be used for many kinds of semantic markup such as, e.g., preparing online content for the

semantic web (Linked Data, etc.). In cases where the use of controlled language is acceptable or even an added value (e.g., corporate terminology, style, etc.), semi-automatic re-writing support can be integrated.

An improvement more in the **downstream direction** is a better usage of translations and corrections provided by human translators for improving MT and authoring systems (**crowd-sourcing**). This scenario links to the one above as in some (many?) cases, it will be necessary to comment on the correction for the system to be able to correctly categorize the error and prevent it in the future. This seems to be the only way to approach the long tail of translation issues mentioned above.

Usage scenarios: Translators and post-editors will be working anytime and anywhere. Their corrections and modifications are recorded and fed back into MT systems and the rest of the translations to be done/corrected should immediately gain from the additional knowledge acquired by the system. For example, a post-editor changes *ladies and gentlemen* into *folks*. The system will offer a ranked selection of explanations such as "style", "length", or "corporate", etc. such that the update mechanism of the system can contextualise the correction and apply it in related contexts.

2. New Architectures (re. Needs #3,4,6,7)

As has been argued above, many current MT architectures (in-house, cloud- or mainframe-based, using proprietary formats on the one hand, free web-based services on the other hand) suffer from severe drawbacks concerning, e.g., interoperability, sustainability, and security. Visions in this field take up current computing trends such as software as a service, cloud computing, etc. For MT, a combined vision is **translation brokering at a large scale** and **translation on demand**. This vision is meant to streamline the translation process such that it (a) becomes simpler to use and more transparent for the user (a one-stop shop), and at the same time (b) respects important factors such as subject domain, language, style, genre, corporate requirements, user preferences etc. (divide and conquer). Technically, what is required is maximum interoperability of all components (corpora, preprocessing, terminology, knowledge, maybe even pre-trained translation models) and a cloud or server/service farm (see section 4.4) of specialized LT services for different needs (text and media types, domains, etc.) offered by SMEs or large companies.

Usage scenario: A customer sends a text to be translated to the brokering engine. The text is classified and analysed according to the aspects mentioned above. The customer gets a configurable offer of different ad-hoc translation pipelines that could be used (+/- preprocessing; +/- domain terminology by supplier X, Y, Z, human post-editing, etc.). Every option is explained, the estimate in quality improvement is indicated as is the

price. The customer can choose either the recommended pipeline or add and remove several options. It is straightforward to extend this vision to include linguistic processing other than MT such as automatic summarisation, terminology extraction, or automatic linking with background information such as Linked Data.

New computing models are needed to be able to react to the more flexible needs (see, e.g., Need #4 above). Key dimensions are illustrated in Figure 1 (on the left) and extended below:

- **Access**: Through which channels and modalities is the translation offered?
  - Desktop tools / classical GUI-based interfaces
  - Web services
  - Mobile apps, multimodal interfaces
- **Volume:** How much is translated and what is the time requirement?
  - Classical product content (e.g. manuals)
  - Asynchronous user-generated content (e.g. forum discussions)
  - Online web-based discourse (e.g. Twitter)
- **Personalisation:** How generic/individual is the translation?
  - One size fits all
  - Personalised to user
  - Personalised to context

The applications in Figure 1 (on the right) exemplify cases, in which each of the three dimensions is maximally relevant. The overall goal is to maximize all three dimensions for as many applications as possible.
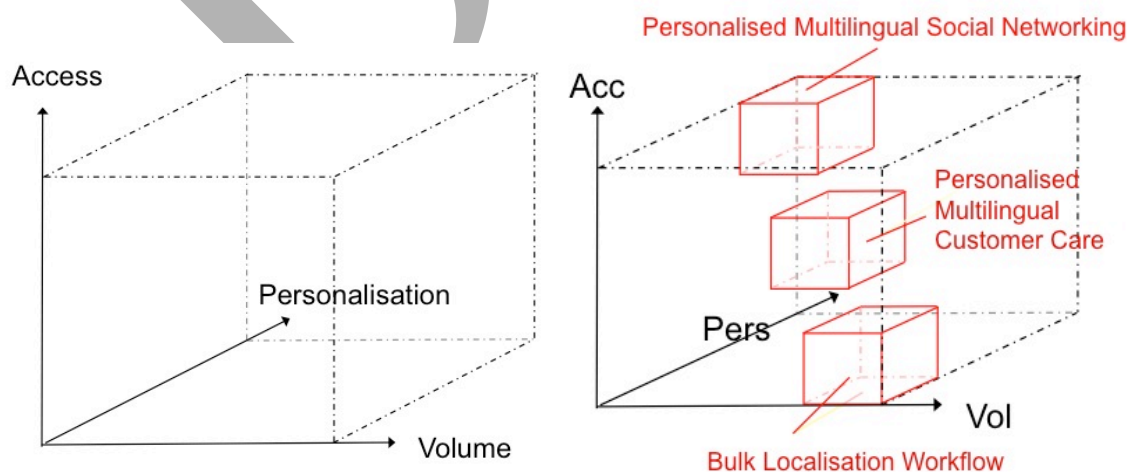

Figure 1: The Translation/Localisation Cube

Note that personalization might include not only customisation to cultural background, but also any kind of smart interaction, e.g., some sort of business intelligence.

3. New Applications

In the field of translation, the core application will be the provision of translation services, wrapped in new workflows and computing models, offering better performance and satisfying new functional requirements. New applications will either combine MT with other IT applications such as business intelligence or with technologies discussed in the other Vision Groups Interactive Systems (e.g., multilingual speech-based interfaces) or Media and Information Services (e.g., multilingual information extraction). One type of system that has been discussed in the Vision Group meeting is a variant of Douglas Adams' Babelfish: a system that understands the user's situation (in terms of ongoing communication and content under discussion, but also ambient information such as the name of a shop on the other side of the street in Beijing). The system can be accessed via speech and its output is presented via online sub-titling in the user's glasses, smart phone display, or headphones. The system understands the context and situation. It provides missing information (if needed) in an appropriate way like a smart personal assistant ("this is not a restaurant but a food shop", "on the 19th you already have an appointment at the dentist", etc.). In the same line, the system can also be used for studying new languages. This vision may be called **personal ambient translation assistant (PATA)**.

Another application that has not yet deserved much attention is **translation within a language.** This includes translations **across modalities**, e.g., subtitling AV for people with hearing impairment or re-speaking for people with visual impairment. More visionary is a translation **across style/genre**, e.g., for re-phrasing legal texts or doctor's prescriptions into everyday language or even children's language.

One application that got special attention in the Vision Group meetings is e-learning (from K12 via University to further education or learning on the job). This could be a prime application for MT and Language Technology in general:

- The potential for crowd-sourcing is high (young, many technology-trained users)
- The content is usually relatively unproblematic concerning legal/security issues
- Computer Assisted Language Learning (CALL) of one's own language or a second language is already introduced on the market
- Scenario: NGOs.
- Educational games might be a step to address games industry

4. New Modes of Research and Training (re. Needs #1-3,8)

The final domain-specific vision is that of a new, **human-centric research paradigm of MT**. The vision is to create a much more natural and organic interaction between humans and machines (including robots) than is foreseen in the current dualism between machine-aided human translation and human-aided machine translation. Some of the roles of the humans in this vision are providers of data, insights, quality judgments, and critique. Humans are first collaborators of research, later users of the resulting systems. In this scenario, systems must be launched much earlier as **prototypes that are really used**. They need not be products, but they should be completed including ergonomics (HCI field studies) and convincing use cases.

At the same time, new **modes of training students** (and software?) **for pre- and post-editing** MT texts is needed. This is still missing in the current curricula for translators as is **training MT research students in translation.** Interdisciplinary cooperation in this area is mandatory for a real breakthrough in MT performance and usability. Finally, multilingual language technology – if integrated into everyday software – can be a key for stopping the further erosion of peoples' (foreign) language skills. This, however, requires political endorsement and enforcement.

4.4 Domain-independent visions

5. **New Computing Models, Business models, and Standards for Sharing Knowledge Resources**

   The most important domain-independent starting points for visions that have been discussed in the meetings of the Vision Group Translation and Localisation shall be mentioned only briefly in this document. The first vision is the further development of new (and not so new) **computing models** and the use for building an **MT ecosystem**:

- Software as a service (flexible, interoperable, supports translation-on-demand, but trust issues need to be adequately addressed)
- Modular, interoperable, plug and play component architectures (for secure in-house and cloud computing) based on agreed standards and meta-data conventions
- Flexible and adaptive workflows connecting (subsets of) components to address particular tasks at configurable quality and speed
- Trusted service centres (as mediators between users and providers)

In parallel, new and **transparent business models** have to be developed to support the idea of translation brokering laid out above. Likewise, new ways of **sharing resources and knowledge in R&D** are needed, among other things to address the Zipfian quality of many MT issues.

Technologies for the Multilingual European Information Society
Vision Group Translation and Localisation: Results of first two meetings

META≡NET

Finally, machines also need to learn more (and self-directed) about the human environment and also draw on artificial knowledge heavily (user needs, ambient semantics, web, etc.). A prerequisite for this is the definition of **standards for sharing knowledge resources** such as encyclopaedic lexica.

5 Conclusions

This report on the first two meetings of the META-NET Vision Group "Translation and Localisation" has presented a detailed overview on the requirements for making the affected industry and research sectors future proof. Some needs arise from the outside such as emerging trends in society and technology use, other needs are founded in the state-of-the-art in MT. In an attempt to address these needs, starting points for visions and innovative directions for research and development have been formulated. The most central starting points for visions in the domain Translation and Localisation identified are:

- Visions 1: New Workflows
  - Intelligent **authoring support** in the sense of **anticipating** problems that can arise in **the translation process**
  - Preparing MT and authoring systems for **crowd-sourcing**, e.g., having the users comment on the correction for the system to be able to correctly categorize errors and prevent them in the future
- Visions 2: New Architectures
  - **Translation brokering at a large scale**
  - **Translation on demand**
  - Flexible architectures implementing the **Translation/Localisation Cube** (Access-Volume-Personalisation)
- Visions 3: New Applications
  - **Personal ambient translation assistant (PATA)**
  - **Translation within a language, across modalities**, **across style/genre**, etc.
- Visions 4: New Modes of Research and Training
  - A **human-centric research paradigm** of MT
  - **Early prototypes** that are really used
  - **Training** students **for pre- and post-editing**
- Visions 5 (domain independent):
  - New, transparent business models
  - New computing/deployment models (e.g. service clouds, trust centers)

- New ways of sharing resources and knowledge in R&D

As has been said above already, the core of the sector is Machine Translation and the visionary potential of a single clearly defined task is likely to be limited. Still, the technologies for realising HQMT, innovative applications using MT, and last not least new workflows closely coupling humans and machines in MT offer a high potential for visions that are both convincing and realistic.