Language Intelligence
for the European Defence Agenda

09 April 2019

LT-Innovate is the Language Technology Industry Association. Language Technology makes the world intelligent, multilingual and interactive. LT-Innovate groups over 200 companies and organisations active in the field of multilingual ambient intelligence: speech interaction, deep meaning processing and multilingual communication & cognition, a set of technologies that are at the very core of what is commonly referred to as Artificial Intelligence (AI).

LT-Innovate is represented in the European Defence Agency (EDA) through its members Bertin-IT, Coreon, Expert System, Sail Labs, SDL and Vocavio.

The global political megatrends of the 21Ist century make it mandatory for European states to join their forces to ensure their capability for independent self-defence. As in many other areas of European integration, it is very hard to make national systems interoperable cross-border. The goal of this paper is to outline the contribution that the Language Technology (LT) / Language Intelligence Industry (LI) could make to creating, strengthening and maintaining a European defence capability commensurate with the challenges our continent will have to meet in the years ahead.

As a preliminary, LT-Innovate would recommend that EDA undertake an audit of existing military capacities to determine where/how they could be significantly strengthened by LT/LI technologies. Below, we are addressing some of the main use cases our member companies have come across so far.

1. Knowledge Interoperability for a European Army

THE CHALLENGE: Data supporting troops

Today, data has become a more important resource than oil or steel, also in defence. Half of data is textual, in Europe multilingual and, thus, does not cross language borders. Most people intuitively see cross-border as a translation problem only. But this is only the surface. The real challenge is to make knowledge interoperable, to empower armies to collaborate smoothly even if their knowledge is structured differently and stored in different languages. This kind of interoperability is critical for decision making processes, particularly in a tactical environment where time to operation is crucial.

THE SOLUTION: Knowledge for defence

Multilingual Knowledge-based Systems, a symbiosis of knowledge graphs and terminology, enable cross-border interoperability. They are the key to frictionless knowledge exchange between organizations. They enable the key trends in efficient use of multilingual data: Search, Machine Learning, AI, and the Internet of Things (IoT), “weapons” every modern army needs to have in its stock. A key factor to this approach is the definition and implementation of a multilingual, shared model or schema in the form of a taxonomy to map the terminologies and semantic representations coming from different sources. Storing and deploying an army’s knowledge in a Multilingual Knowledge System can be labour intensive. Luckily, in most countries a wide range of raw linguistic resources is already available: translation memories, glossaries, or simply content.

PROPOSED ACTION: Deploying knowledge[1]

In the short term, the following actions should be prioritised:

[1] Relevant reference schemas and terminologies include:
- CISE data model
- EIXM
- EUROPOL SIENA
- S1000D

- Identify areas of cross-border interoperability (for example, a European armament project, central command, or international manoeuvre).
- Research and identify existing linguistic assets and convert them semi-automatically into a Multilingual Knowledge System.
- Deploy the Knowledge System together with a knowledge consumer to obtain a proof-of-concept of increased European defence preparedness.
- Standardize the procurement and technical publications for Air, Land and Maritime equipment by using technologies that enforce the S1000D standard.

2. Relying on Verified Intelligence

THE CHALLENGE: Acquiring intelligence from multiple sources and modalities in support of multiple analytical tasks

Military intelligence uses information collection and analysis approaches to provide guidance and direction to assist commanders in their decisions. A basic need of the intelligence analysts is to understand in advance intent coming from targets and to profile them using different semantic attributes (EX-ANTE analysis). At the same time, a similar approach can be applied to past events in order to understand the related facts, entities and relations involved (EX-POST analysis). This type of behavioural analysis is a new challenge in the Defence and Intelligence sector, including aspects such as multimodal subject recognition based on content (text), emotion, and voice print analysis. To acquire intelligence from multiple sources and modalities in support of multiple tasks, it is necessary to provide automated means to identify and extract knowledge in the different languages and modalities, e.g. text, video, audio, of interest from a variety of sources in a single, unified representation including the normalization of multilingual content through machine translation that enables analysis through machine inference and reasoning. Artificial intelligence approaches can then leverage such knowledge to assist strategic intelligence activities.

Some examples of key tasks include e.g. anomaly identification, looking not only at the usual sequence of events on a related topic or geographic area, but especially at unusual things that may require a timely reaction. Also, fact or claim checking, automatic scoring of content based on its trustworthiness, allowing analysts to rapidly identify misinformation campaigns and malicious information actors. Another key

scenario is related to the interception of enemy transmissions, the extraction of information from the speech through noise-resilient automatic speech recognition (ASR) methods and its interlinking with other pieces of information extracted from text, though LT/LI, video, or structured knowledge bases, e.g. Wikidata. This scenario highlights the relevance of cross-modal (and cross-lingual) approaches in military intelligence. The interplay between the different modalities not only brings additional intelligence to the picture but also contributes to fill in the gaps when the signal is poor, e.g. with noisy audio or incomplete database records.

THE SOLUTION: Integrated knowledge representations

Many of the AI algorithms used today to address such challenges are computationally and data hungry. Furthermore, manually tagging and curating the necessary datasets required for training supervised approaches is often impractical in terms of costs and reaction time. Thus, it is necessary to come up with hybrid approaches that allow leveraging pre-trained models, fine-tuning them on specific tasks, hence reducing the amount of data and resources required for training. Also, the use of free sources of supervision in cross-modal data management, like the correspondence between figures and their captions or between audio and video is key to produce visual, audio and linguistic features that outperform those resulting from training on individual modalities in tasks like e.g. multi-modal classification or machine comprehension. Another key element is the reuse of existing structured resources, i.e. knowledge graphs, which can be interoperable with neural architectures and at the same time can inject such semantic knowledge, previously curated by knowledge engineers. Approaches like Vecsigrafo go in this direction[2].

PROPOSED ACTION

Active research in the following areas and their application to the military and intelligence domain in a cross-lingual, cross-modal scenario are key in order to address the above mentioned challenges:

[2]R Denaux and JM Gomez-Perez, 2019,.Vecsigrafo: Corpus-based word-concept embeddings-bridging the statistic-symbolic representational gap in natural language processing. To appear in Semantic Web Journal (http://www.semantic-web-journal.net/system/files/swj2148.pdf).

- Hybrid approaches to address cross-modal and cross-lingual scenarios, like video, speech and text analytics, enabling learning from such modalities plus structured knowledge graphs. The goal here is to leverage the information contained in mutually supportive modalities[3] to augment the information that is otherwise captured through uni-modal and mono-lingual approaches. Potential application scenarios include embedding semantics in speech to text and machine translation algorithms. This will also alleviate the need for manually annotated data since correspondence is a source of supervision itself.
- Automated language identification and translation of digital content, obtained from both open and closed sources, with high accuracy and actionability, in a secure manner, in order to facilitate data fusion and target -language analysis in the analyst's native language. Translation of conversational chat languages, such as Arabizi, as found on social media channels.
- Neologisms and slang analysis, acronyms, spelling mistakes, jargon, neologisms and abbreviations, enabling a high precision analysis of SMS, blogs and social media sites and fine-tuning on short utterance.
- Automating linguistic support for emerging dialects: Related to the previous line of research, here we propose assisted methods to derive NLP support for dialects based on existing resources for the main language they branch from, e.g. Arabic -> Araby[4].
- Encoded messages[5] and anomaly detection under different forms including unknown concepts or variations thereof.
- Computational stylometry to profile the style of writing/speaking of an individual, supporting e.g. the identification of connections of such individual with criminal groups using a marked terminology, mapping several user ids on social media.
- Bot/no bot analysis, trust-modelling: Identifying, particularly in social media, when a campaign is being promoted by humans or bots can be useful to detect disinformation campaigns or spurious interest around a topic, preventing virality.

[3]Examples of mutually supportive data modalities include, to name but a few: the audio and images in a video, a figure and its associated caption in a scientific publication and an entity, e.g. the European Union, its Wikipedia page and its representation in DBpedia.

[4]Araby is an alphabet used to communicate in Arabic over the Internet or for sending messages via cellular phones. It is a character encoding of Arabic to the Latin script and the Western Arabic numerals.

[5]Article published in El País: La Inteligencia Artificial que detecta amenazas terroristas: https://elpais.com/tecnologia/2019/01/02/actualidad/1546466761_871053.html

- Reasoning and Inference is one of the most important tasks in the intelligence process. It is useful in order to help the analyst to create a narrative about real facts and hypothesis linked to a specific target.
- Emotion mining, not just using a standard positive/negative/neutral evaluation approach (sentiment) but also providing a finer granularity, with different kinds of feeling, e.g. stress, fear, trust, anger, etc. and assessing the emotions related to a specific topic, entity or target of interest.
- Event discovery: Understanding events is fundamental for intelligence and defence analysts. Events can be deducted from the extraction of categories and specific facts, entities, locations, and temporal references.

3. Taking Informed and Rapid Decisions

This challenge is linked to challenge 2 as it focuses on how, once relevant information has been collected and structured, it can be enriched with new information (ideally, on the fly) and used to take rapid decisions. This means to be able to update the knowledge base and reason on new facts. The value of this scenario is dual. One the one hand, it can be used to better understand the situation and decide based on what happens in real time and what has happened before; on the other hand, it could be used to design training sessions for soldiers. Another important aspect, which is not addressed by LT/LI technologies, is visualisation and navigation through structured data. In order to address this challenge, one can, starting from information extracted from past events that have been stored into knowledge bases, use prediction algorithms to find past patterns and overlay them with evolving situations as to evaluate the likelihood of future events, and take decisions accordingly.

Communication Intelligence (COMINT)

THE CHALLENGE: Improve COMINT output in breadth and quality

Aside from Open Source Intelligence (OSINT), Communication Intelligence (COMINT) is a critical asset for the Forces' situation awareness and decision ability. Yet, exploiting captured communications raises very specific challenges, due to the poor audio quality and adversity of most recordings, combined with the mass of data, the difficulty to qualify relevant ones, the diversity of accents and languages, coded language,

encryption, etc. Filtering methods based on meta-data only, like caller number, callee number, device used etc. have clear limits, as people of interest often change devices and numbers, also the variety and potential switch in spoken languages make qualifying and distributing task uneasy and people/time intensive, not to mention precious time lost on a high number of all-but-empty or irrelevant recordings.

THE SOLUTION: Robust and agile adapted modules for communications processing

Speech-to-text technology brings obvious value to this challenge, but needs specific adaptation to reach sufficient accuracy on adverse COMINT data, as well as a way to ramp up on new languages and domains. It also needs insertion into a robust and modular workflow which comprises processings like voice activity detection, language identification, language switch tracking, speaker identification, keyword/key-phrase spotting, n-best processing and more.

PROPOSED ACTION

- Collect and annotate target data (with right accreditations and appropriate work environment)
- Enhance Speech-to-Text accuracy, make it robust to non-native accents and other variabilities
- Develop continuous learning schemes, fast domain and language training and adaptation
- Enhance Language and Dialects Identification, detect language and other complementary speech processing modules,
- Other specific speech intelligence tasks: detect key elements in a conversation, detect anomaly, trends, changes of topic etc. (can be combined with section 2 “intelligence”)

DIGITAL CONTENT (OSINT, WEBINT, SOCMINT, SIGINT, CYBERINT)

THE CHALLENGE: Big data processing for military purposes

Every action today leaves a digital trail. Be it the downing of flight MH17 in Ukraine, or ISIS terrorists fighting in Syria. The examples of Bellingcat and others prove the

existence of such digital trails and their importance in the analytic process. Today, relevant information is spread across a wide variety of sources, such as government controlled and private TV, radio, controlled and open social media, the open & dark web, intercepted electronic communications, connection data and many more. Such content is published with increasing speed and volume 24/7 in a variety of languages and formats (audio, video, images, textual, geo-coordinates, etc.). It has long become impossible to extract mission-critical information from such sources by mere human labour.

THE SOLUTION: Enhanced data mining solutions

- Collect and store all available such sources in an automated manner.
- Apply machine-learned models to enrich, index and annotate original sources with derived language-independent/agnostic meta-data. This includes, but is not limited to, speech-to-text, named-entity recognition, sentiment analysis, video and image analytics, optical character recognition (OCR), etc.).
- Process the collected and enriched information to detect clusters, cross-correlations, anomalies, outliers, trusted and untrusted documents.
- Visualize large data sets as to make them understandable to the human analyst.
- Automate reporting capabilities as to minimize work-load on human analysts.

PROPOSED ACTION:

Additional R&D is recommended in the following areas:

- The big ‘Why’: today’s solutions in the field suffer from a lack of transparency on why certain information is selected, displayed, or interpreted in a specific manner. Neural networks suffer from failing to deliver a satisfactory answer to the “why” question. Research is required to approach this topic.
- Cross-correlation and anomalies: find means by which to allow cross-correlation and anomaly detection of different types of source data across a variety of languages.
- Rapid operational readiness: find means by which systems can be extended to additional operational language requirements within 60 days, in order to support military missions in previously uncovered regions or domains.

- Develop interfaces to existing systems OSINT, WEBINT, SOCMINT, SIGINT, CYBERINT, as well as to new and evolving social media platforms.
- Interface to current efforts applying brute-force (or other) decryption of encrypted data, primarily encrypted social media, web-sites, video-channels, etc.

MOBILE & AUGMENTED COMMUNICATION

THE CHALLENGE: Ensure that military communication does not fall behind civilian communication

Rapidity of action and communication is vital for deployed forces. Agents on the field should be able to efficiently interact with their equipment, their teams and with HQ, whether they need to send out reports, ask for information or immediate actions. Besides, more and more teams are equipped with mobile devices, and expect them to offer same services and ease of use as publicly available smart devices which rely on GAFA cloud applications like Google voice, Siri, Amazon Alexa, etc. For obvious reasons, these public services cannot be activated on troops dedicated devices, but robust and secure equivalent should be made available and tailored to specific troops needs

THE SOLUTION: Develop augmented communication solutions

- Develop mobility solutions for field operations: ability to use voice commands (a European military version of Siri/Alexa) and voice dictation for reports and memos.
- Develop augmented communication solutions for interactions between field and HQ, why not using assistant bots (converges with section 4 “simulation”).

PROPOSED ACTION

- Identify relevant scenarios with end-users
- Set up hybrid architecture: application embedded on device / services running on troop secure private network
- Work on standard commands and adaptation toolkits
- Train high accuracy troops memo and report dictation application
- Define augmented communications models and applications

4. Enabling Combatants through Simulation and Training

THE CHALLENGE: Augmented aerial, naval, land and cyber systems

The role of speech technologies to secure military advantage in operations has been overlooked for decades. With the advent of successful consumer-based speech technology devices like Alexa, military operators worldwide are now considering how speech technologies might just solve some workload challenges at an operational level for the soldier. Whether it is shaving seconds off information sharing, readying a weapon or assessing competency in a simulation and training scenario, speech technologies now present a real advantage that can no longer be ignored. Given there are some legacy perception issues around accuracy levels of speech technology, the obvious sandbox for speech technologies today is within LVC (Live, Virtual and Constructive), also known as simulation and training. Here there is scope for rapid capability development, testing and validation that maintains a comfortable safety margin.

THE SOLUTION: Immersive, high and low fidelity simulators

Simulation and training is the use of immersive, high and low fidelity simulators to enable the solider to prepare for operations. While aircraft pilots are still the focus of most dynamic simulation and training systems, naval, land and cyber systems are being augmented to ensure the trainee understands their role in a battlespace and can execute on the plan to achieve a successful outcome on the mission. A new wave of technologies is emerging that are collectively known as ‘multi-modal’ technologies; these are being used to understand the soldier and their performance in both simulated scenarios and in time in operations itself. The primary use of speech technologies in simulation and training is at an early but significant stage. Some significant challenges among military operators around access to instructors and the challenge of maintaining standards of training outputs has them investing in evidenced based systems that are underpinned by data driven technologies, including speech technologies.

PROPOSED ACTION

Additional R&D is recommended in the following areas:

- Human machine interaction (HMI) studies of workload in simulated normal and non-normal operations; blending speech technologies that are looking at both content and social signalling between participants.
- Standard operating procedure (SOP) adherence and modification studies (particularly for multi-cultural and inter-service operations);
- Avatar based procedural training that enable the military operator to provide personalised and adaptive training; and enable instructors to deploy their knowledge more accurately.
- Closing the operations and training gap; we need greater use of audio capture in operations and training to enhance the fidelity of training, gain understanding of workloads and accuracy and suitability of procedures in a battlespace.
- Data driven training for coalition forces across synthetic training will provide more accurate understanding of mission readiness; as the use of this synthetic (virtual) training becomes more common place, there will be the need to utilise speech technology to understand readiness of a force and their accuracy in executing on a plan of operation.