

ENRICH4ALL Action

(**E**-gover**N**ment [**RI**] **CH**atbot for **ALL**)

ELG Workshop on Resources for Luxembourgish and Flemish

—
Dr. Dimitra Anastasiou
IT for Innovation Services (ITIS)
Luxembourg Institute of Science and Technology (LIST)

08.07.2021

- Corporate video:
https://intranet.private.list.lu/fileadmin/files/3_Communication/Luxembourg_RTO_-_Where_tomorrow_begins.mp4
- LIST has **617 employees** (65% men, 35% women) over 46 nationalities and is organized into 3 main research departments dedicated to ecological innovation (ERIN), material innovation (MRT) and digital innovation (ITIS).
- The “IT for Innovative Services” (ITIS) department of LIST is home to 110 IT scientists and engineers performing research and innovation activities in computer and data science topics, applied, in particular, to the design, engineering and assessment of novel approaches to information systems and services. These activities are supported by a Data Analytics/AI technological platform including **HPC, Visualization Wall and a Cognitive pillar**. ITIS expertise relates to:
 - *Trusted Distributed and Optimized Systems*
 - *Data Science and Explainable Analytics*

R&D projects in progress in 2019

100

competitive
national
projects

25

competitive
European
projects

63

collaborative
projects
& similar

1

international
competitive
project

7,87%

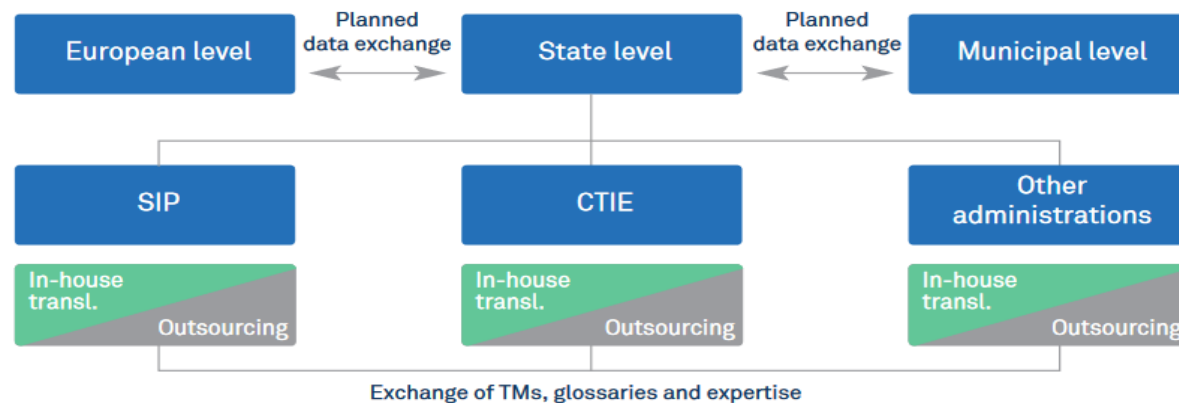
main service
contracts

Σ 297

total of RDI projects & contracts

Luxembourg & Language Technology

- Highly multilingual country (3 official languages: **Luxembourgish, French, German**)
 - Luxembourgish is not an official EU language
 - <https://luxembourg.public.lu/en/publications/ap-langues.html>
 - Translation management lacks a systematic workflow, since the translation is made internally
 - Low R&D
 - LSPs / Translation companies: Alphatrad, Inlingua, Lingua Custodia, Pythagoria, Semantica, Tradulux, ViaVerbia, Wordbee



ELRC White paper: https://www.lr-coordination.eu/sites/default/files/ELRC_Conference/ELRCWhitePaper.pdf

Which language and where?

- **Home:** Lëtzebuergesch is the most widely spoken language (74%), followed by French (32%) and Portuguese (15%).
- **Workplace:** 84% use primarily French, followed by Lëtzebuergesch (73%), German (51%) and English (37%).
- **Social context:** French (81%) narrowly outstrips Lëtzebuergesch (77%): the latter is the preferred language among young people aged 15 to 24 (92%) and those aged 65 and above (80%) in the context of their free time.

UNE POPULATION PETITE ET DIVERSE

APPRENEZ-EN PLUS SUR LA COMPOSITION DE LA POPULATION LUXEMBOURGEOISE



FR DE EN

 Guichet.lu



CORONAVIRUS

Procedures

Appointment Large Scale Testing



The MyGuichet.lu app:
my administrative procedures
**where I want, when I want,
how I want!**

Available on



 Homepage
Citizens

 Homepage
Business

POPULATION



Today, Luxembourg has a population of 626,000 people. Albeit small, it boasts an incredible diversity. In fact, almost 50% of the population doesn't have Luxembourgish nationality! And this does not even take into account the 185,000 cross-border employees who work in Luxembourg. This mix of languages and cultures from all these communities is rewarding for life in Luxembourg and gives it a cosmopolitan characteristic.



Minister

- Xavier Bettel

Minister Delegate

- Marc Hansen

Contact

Remit

The government program presented by Prime Minister Xavier Bettel to the Chamber of Deputies on 11 December 2018, reflects the government's desire to place digitalisation at the center of its concerns.

Thus, for the first time in its history, was created a Ministry for Digitalisation headed by the Prime Minister, Minister for Digitalisation, Xavier Bettel, and Minister Delegate for Digitalisation, Marc Hansen.

[Read more](#)

Luxembourg's strategic vision for AI

This strategic vision for AI in Luxembourg articulates our activities, ambitions and intentions related to the role of AI, both here and across the EU.

- files:



a strategic vision for
Luxembourg



une vision stratégique pour le
Luxembourg

Agenda

- **Context of ENRICH4ALL**
- Benefits of a government chatbot
- eTranslation
- Domains

ENRICH4ALL: Chatbot project

- European project CEF (Connecting Europe Facility) Telecom 2020 (Automated Translation): <https://ec.europa.eu/programmes/horizon2020/en/h2020-section/fet-open>
- Deploying / **Integration projects**
- Project duration: 24 months – Start: 01.06.2021
- 4 Partners / Coordinator: LIST
- Total project budget: ca. 803,858 € (CEF 75%: 671,177 €)

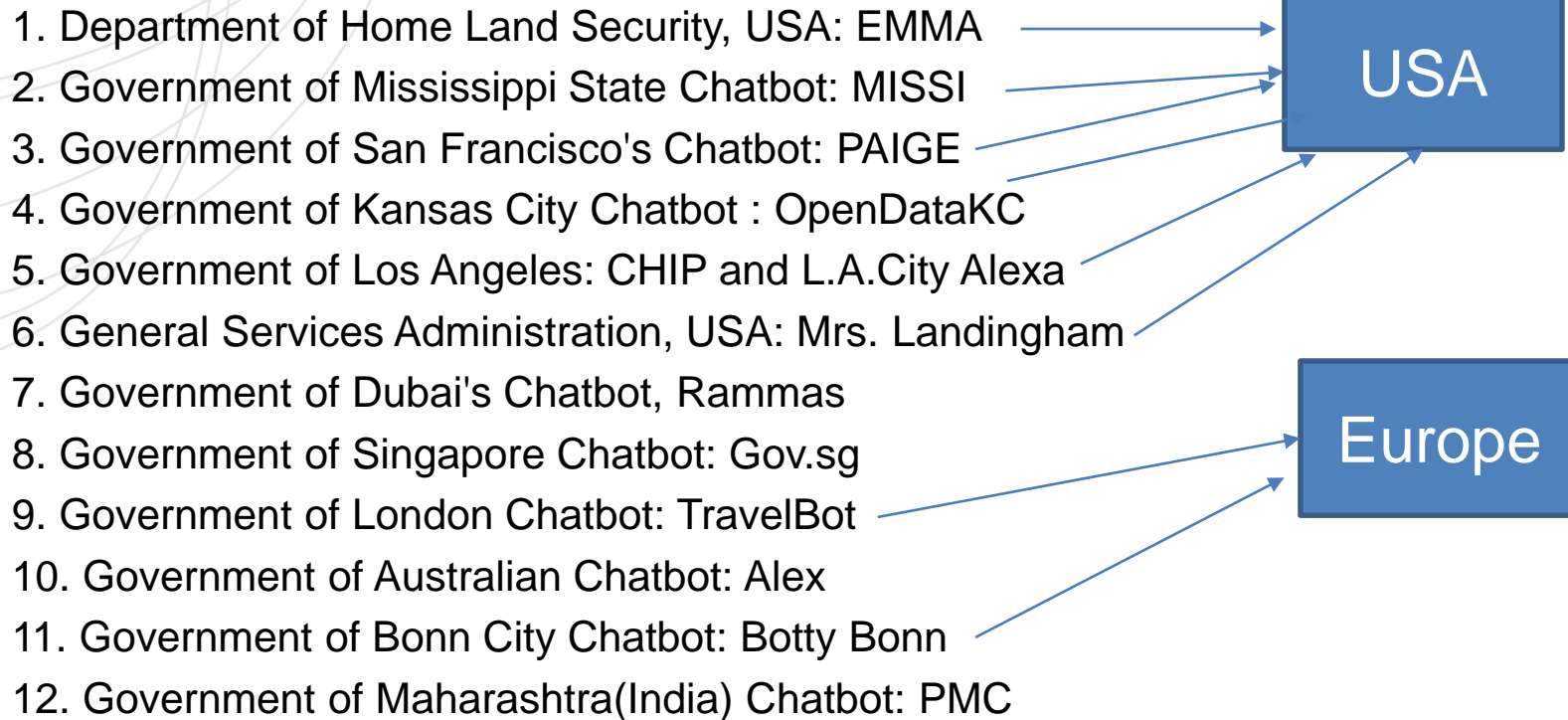
ENRICH4ALL Summary

- We aim to develop an **e-government chatbot (eChat)** which lowers the language barriers in the EU and the CEF Associated countries. Since it is a virtual assistant, it is continuously available and can instantly reach large amounts of people, irrespective of age, gender, educational background, and geographical barrier.
- **Multilingual bot:** The Machine Translation tool **eTranslation** will be integrated in the **eChat** bot, so that the user can interact with the **eChat** in their own mother tongue.
- Deployment of the eChat in **Luxembourgish, Romanian, and Danish** public administration
- ENRICH4ALL is an initial Action towards a **unified, broad EU-wide eTranslation-based bot**, which is currently missing in public administration in many countries.

Chatbot statistics

- eCommerce chatbot statistics show that businesses spend around **\$1.3 trillion** on customer requests per year. With the assistance of chatbots, this **expense could be reduced by 30%** (*Chatbots Life*).
- According to *Acquire*, 1.4 billion people are using chatbots now on a fairly regular basis. *IBM* stated that **chatbots can answer 80% of standard questions**.
- **US** is home to the largest portion of the world's chatbot users (36%), followed by **India** (11%); **Germany** is at third place (4%) – it's the highest-ranked non-English-speaking country on the list (*Chatbots Life*).

Government agencies that use chatbot



E-Estonia: <https://e-estonia.com/ai-chatbot-to-replace-and-improve-governmental-e-services/>

Source: <https://blog.vsoftconsulting.com/blog/15-governments-agencies-that-use-chatbots>

PARTNERS

- **BEIA Consult International** (Romania) <https://www.beiaro.eu/>
 - Key supplier for ICT solutions/services in Romania, with experience in 10000+ commercial projects since 1991.
 - Has been working on many speech, text and data processing projects.
- **SupWiz**: <https://www.supwiz.com/> (Denmark)
 - University-spinout founded in the summer of 2017
 - AI product suite comprises chatbots, live chat, a ticket analyser, etc.
 - Consists of 20 multi-disciplinary experts
 - Nominated “Most Promising Start-Up” at the SIA 2018 (Search Industry Awards 2018), and in 2019, SupWiz was awarded the largest innovation prize by the Danish Minister of Science and Education
- **RACAI** (Research Institute for Artificial Intelligence, Romania) <http://www.racai.ro/en>
 - Established in 1994 as a small research centre in AI and became an institute in 2002
 - Largest group at RACAI (10 persons) is doing research and development in (multilingual) natural language and speech processing.



Agenda

- Context of ENRICH4ALL
- **Benefits of a government chatbot**
- eTranslation
- Domains

Government chatbot



Benefits:

- Process service request in huge numbers
- Work 24/7
- Target large amounts of people irrespective of language, geographical barrier, COVID situation
- Lower cost for public administration
- Always timely and up-to-date information

Personal data security is an essential requirement for the deployment and viability of e-government chatbots.

Government chatbot: How chatbots can improve citizens' experience: <https://www.virtualspirits.com/government-chatbot-improve-citizens-experience.aspx>

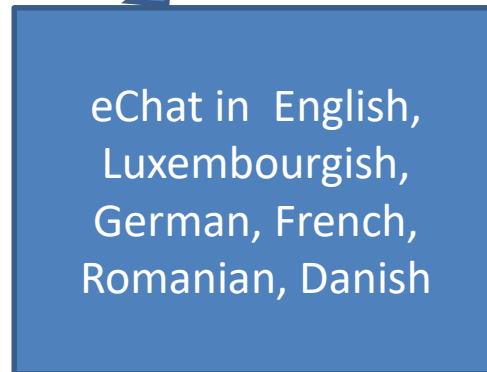
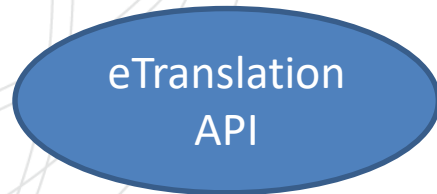
Agenda

- Context of ENRICH4ALL
- Benefits of a government chatbot
- **eTranslation**
- Domains

CEF eTranslation

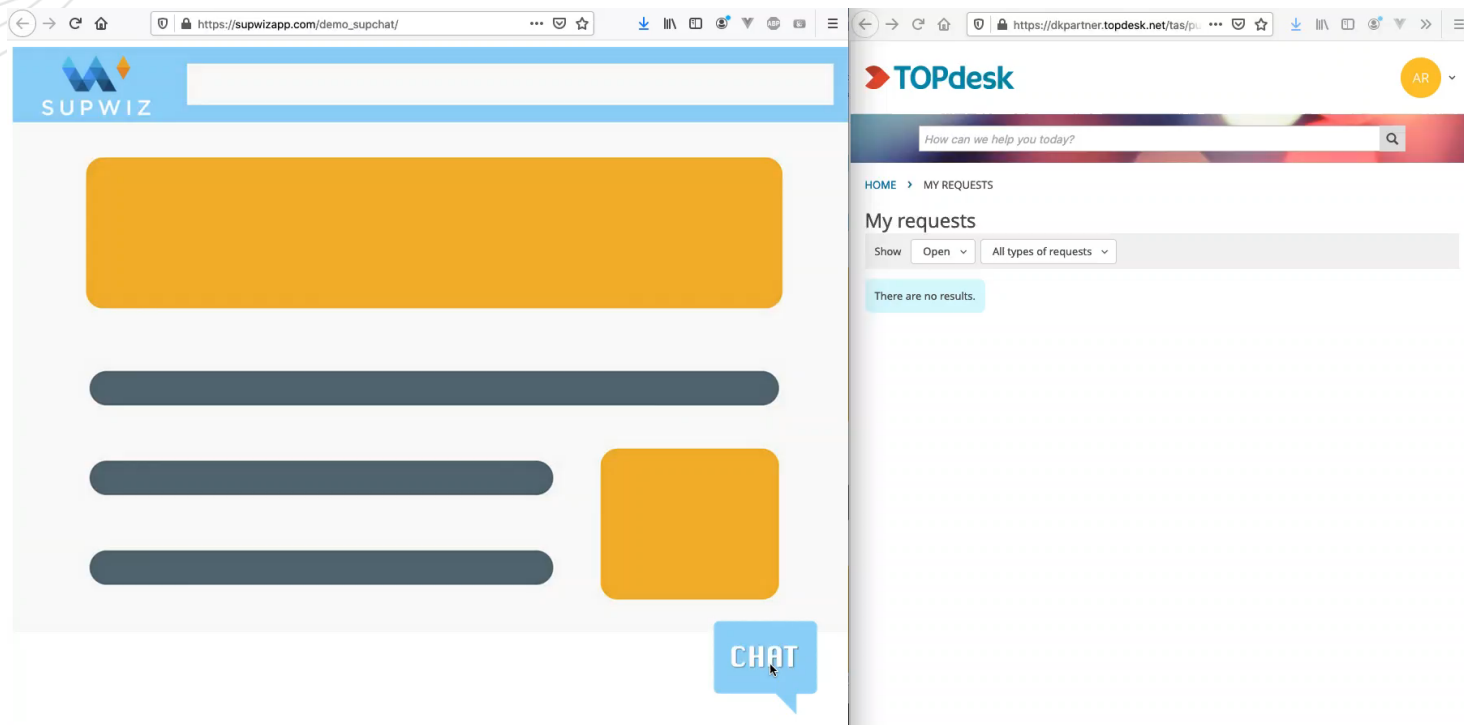
- Access: <https://ec.europa.eu/cefdigital/wiki/display/CEFDIGITAL/eTranslation>
- CEF eTranslation helps public services being multilingual:
 - **Neural machine translation** tool provided by the European Commission to all EU bodies but also public services and public administrations across Europe
 - for 24 official languages of the EU (24 <-> 24), Chinese, Russian, Turkish, Arabic
 - Secure & scalable
 - Facts and figures:
 - 70+ systems connected to eTranslation
 - 7.000 SMEs registered
 - More than 40 Mio. pages translated in 2020
- Domains coverage:
 - **EU formal language**
 - General text
 - Court of Justice Case Law
 - Cultural
 - Deutsche Bundesbank
 - IP Case Law
 - Ministère des Finances (France)
 - Public Health
 - Technical Regulation Information Systems
 - Valtioneuvoston Kanslia

ENRICH4ALL Workflow



Ticket analyzer

- The Ticket Analyzer harnesses users' historic data to effectively predict the intent in tickets and automatically route them to the right agents.
- Integrate the Ticket Analyzer to an ITSM (IT service management), e.g. ServiceNow, BMC Remedy or Zendesk.



Natural Language Understanding

NLU

- NLU understands the *context*: no need to type in tons of examples of wording to manipulate the model.
- NLU understands *typos and synonyms*: users should not worry about typos, misspellings and synonyms for your specific keywords.
- *NLU* of Supwiz is based on BERT models.
 - *CamemBERT* for French language model: <https://camembert-model.fr/>



CamemBERT

A Tasty French Language Model

Facebook AI Research

Inria

ALMAAnaCH

CamemBERT

CamemBERT is a state-of-the-art language model for French based on the [RoBERTa architecture](#) pretrained on the French subcorpus of the newly available multilingual corpus [OSCAR](#).

We evaluate CamemBERT in four different downstream tasks for French: part-of-speech (POS) tagging, dependency parsing, named entity recognition (NER) and natural language inference (NLI); improving the state of the art for most tasks over previous monolingual and multilingual approaches, which confirms the effectiveness of large pretrained language models for French.

CamemBERT was trained and evaluated by [Louis Martin](#), [Benjamin Muller](#), [Pedro Javier Ortiz Suárez](#), [Yoann Dupont](#), [Laurent Romary](#), [Éric Villemonte de la Clergerie](#), [Djamé Seddah](#) and [Benoît Sagot](#).

Agenda

- Context of ENRICH4ALL
- Benefits of a government chatbot
- ELRC and eTranslation
- **Domains**

Domains of ENRICH4ALL

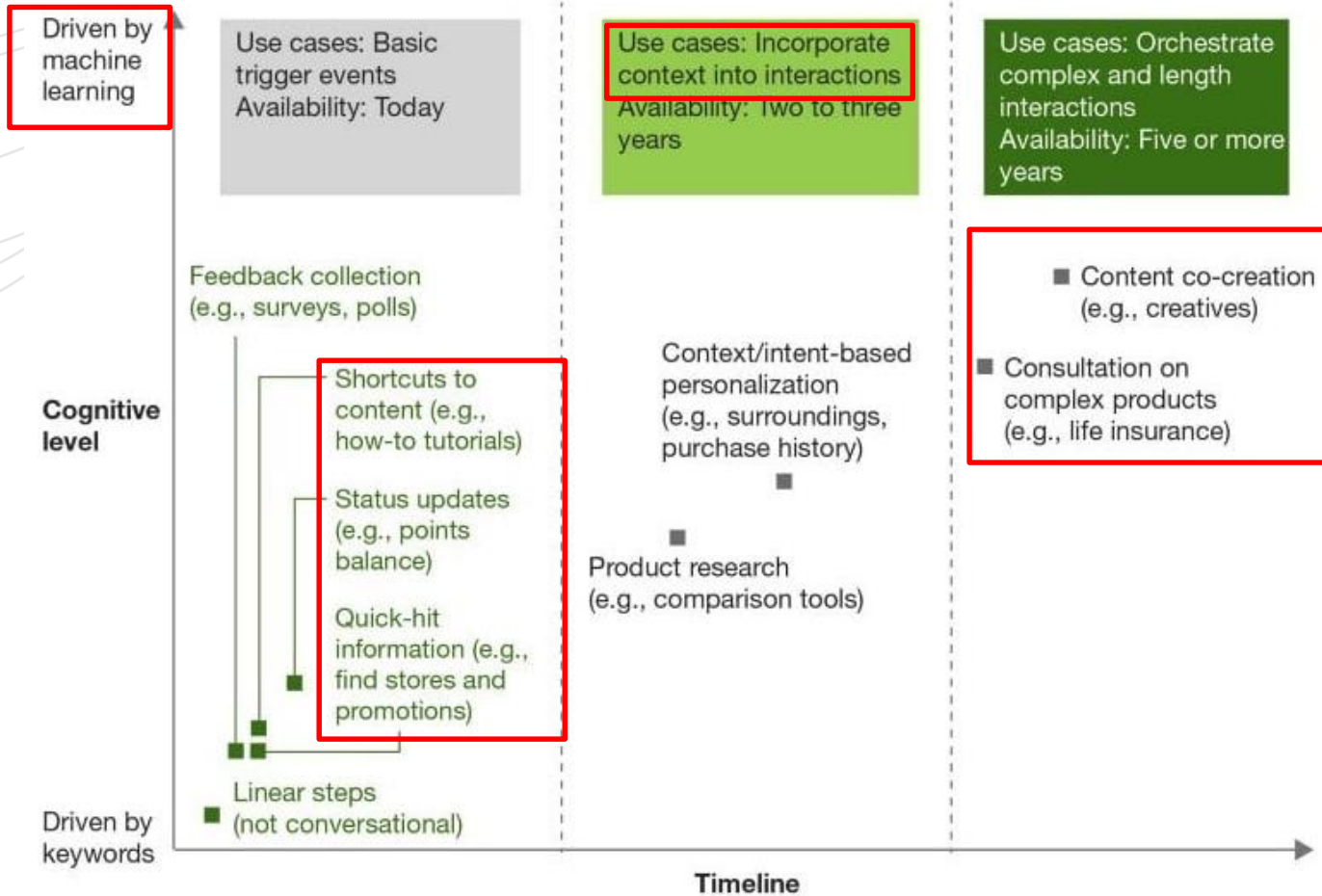
- General administrative questions (certificate of residence, marriage)
- Tax information
- COVID-19 /eHealth
- Visa applications
- Transport/mobility

Challenges

- Artificial agents are presumed to have sub-standard intelligence, people still apply certain social stereotypes to them (Heyselaar & Bosse, 2019).
 - More anthropomorphic cues->more likeable, sociable, friendly, and personal
- Chatbot behaviour: naturalness and content preservation.
- Chatbot's **credibility**: Credibility is presented as a rating of confidence in the accuracy of the information contained in the answer. According to Corritore et al. (2005), credibility consists of four factors: honesty, expertise, reputation, and predictability.
- Influence of **conversational style** on user experiences with chatbots. Elsholz et al. (2019) compared interactions with chatbots that use modern English vs. a Shakespearean language style. Users perceived the chatbot the former as **easy** to use, while the latter as more **fun** to use.

Future of chatbots

AI advancements move chatbot potential from “question and answer” to “human-like”





**EUROPEAN
LANGUAGE
GRID**



dimitra.anastasiou@list.lu

www.enrich4all.eu