Architecture for public service chatbots

DIGIT
Directorate-General for Informatics

ISA² Programme
ec.europa.eu/isa2

ISA²

DISCLAIMER

This document is for informational purposes only and the Commission cannot be held responsible for any use which may be made of the information contained therein. References to legal acts or documentation of the European Union (EU) cannot be perceived as amending legislation in force or other EU documentation.

The document contains a brief overview of technical nature and is not supplementing or amending terms and conditions of any procurement procedure; therefore, no compensation claim can be based of the contents of the present document.

The information and views set out in this publication are those of the author(s) and do not necessarily reflect the official opinion of the European Commission. The European Commission does not guarantee the accuracy of the data included in this document. Neither the European Commission nor any person acting on the European Commission's behalf may be held responsible for the use which may be made of the information contained therein

EUROPEAN COMMISSION

Directorate-General for Informatics
Directorate B — Interoperability Solutions for public administrations, businesses and citizens
Unit B6 — ISA^2 Programme

Contact:: isa2@ec.europa.eu

European Commission
B-1049 Brussels

DOCUMENT METADATA

Property	Value
Release date:	*04/09/2019*
Status:	*Published*
Authors:	*PwC EU Services*
Keywords (to increase the SEO presence)	*Architecture, chatbot, CPSV-AP, Core Vocabulary*

Table of Contents

List of Figures

List of Tables

1. EXECUTIVE SUMMARY

To facilitate the access of citizens and business to public services, a large number of e-catalogues and e-Government portals have been implemented throughout Europe. In many cases the development of these catalogues has not been harmonised. This makes it difficult for public administrations, citizens and businesses to have an integrated view on life events, business events and the public services provided within a specific country or region. The ISA[2] Action "Catalogue of Services"[1] supports an integrated view on existing life events, business events and related public services by:

- Analysing life events and business events and the related public services within the EU, in order to identify challenges and good practices to be shared;
- Building consensus on common ways of representing life events, business events and public services, across borders and sectors. This resulted in the creation of the Core Public Service Vocabulary Application Profile or CPSV-AP;
- Developing a number of reusable building blocks for implementing catalogues of services at local, regional, national and European level;
- Supporting the Single Digital Gateway action by providing tools starting from the description of public services to the practical provision of public services; and
- Creating guidelines and practices for the management of portfolios of public services.

In this context, this document aims to identify how public administrations can best provide public services via chatbots. The scope of this study is limited to a desk research on the usage of chatbots in the public sector and includes input from:

- the 2018 report "Report on the promotion for the uptake of the tools"[2], where a CPSV-AP based chatbot PoC was developed;
- the 'Catalogue of Services' workshop[3], which took place on 16th May 2019 in Brussels; and
- existing public administration chatbots[4].

This document contains the components of the high-level architecture for public service chatbots. Additionally, it aims to provide an overview of possible functionalities the chatbot could have.

---

[1] https://joinup.ec.europa.eu/solution/core-public-service-vocabulary-application-profile/about

[2] Limited access: https://webgate.ec.europa.eu/CITnet/confluence/display/COPS/D03.03%3A+Report+on+the+promotion+for+the+uptake+of+the+tools

[3] https://joinup.ec.europa.eu/collection/semantic-interoperability-community-semic/event/technical-workshop-catalogue-public-services-vis-vis-sdg-implementation-16-may-2019-930-1730-cet

[4] The overview of researched chatbots for this document can be found in section 12.2.

Public administrations will need to make use of (a lot of) resources to develop chatbots while chatbots are not (yet) the solution that will guide all users to the information they need in an instant. However, it will most likely get there.

Global interest in chatbots is booming. Chatbots can be trained exponentially faster than humans can, they are 24/7 available and react instantly to user queries. On top of that, public administrations can save a huge amount of resources by the decrease in user queries to human operators (e.g. through the helpdesk). Another big plus for utilising this technology is making it easier for the elderly, the sick and the disabled people to have access to public services. Chatbots could lower the barriers to contact or ask public administrations for help. Taking all of this together, researching an architecture for public service chatbots offers a glimpse of what the future could be for public service (information) delivery.

As there is a steep rise in the use of chat messaging services and voice assistants, public service chatbots could have the momentum they need to grow strongly in development as well as in usage.

This document researches different types and domains of chatbots, architectural components, chatbot functionalities, security, and so on. Figure 1 shows a high-level chatbot architecture, which will be used to summarise the needed technologies, when building a (public service) chatbot.

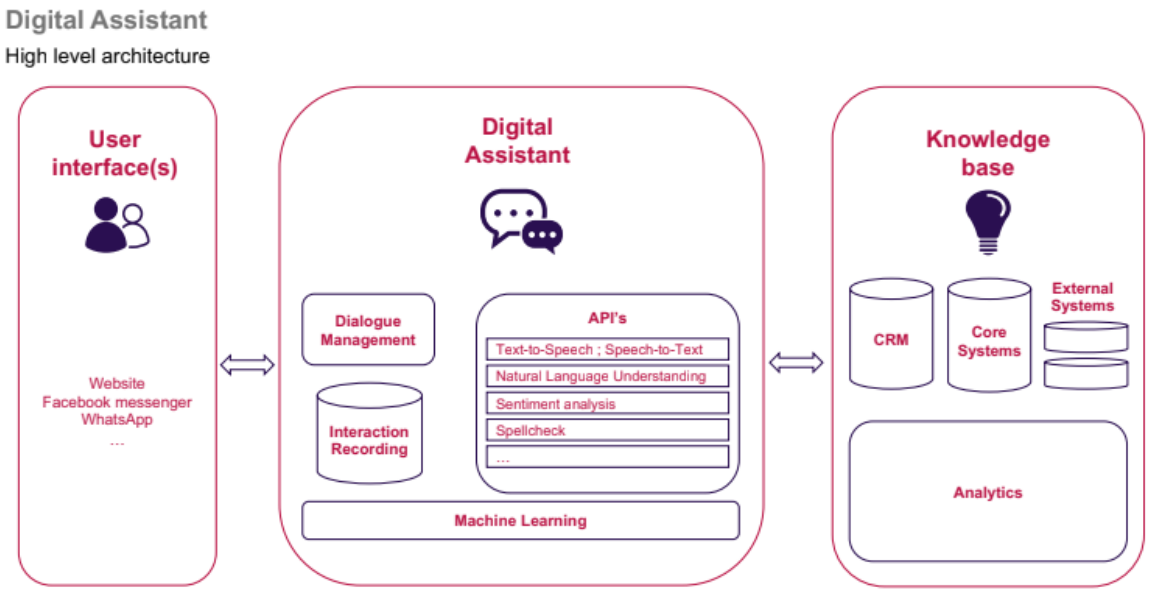


Figure 1: High-level chatbot architecture.[5]

Based on Figure 1, an overview is provided on all recommended components for public administration chatbots. This overview is provided in Table 1.

[5] Provided by Reimagine, working on the Woluwe-Saint-Pierre communal chatbot.

Table 1: Overview of public administration chatbot components.

Architectural main component	Architectural component	Section(s)
User interface	Bot Development Framework API	5.1.1 6.1.1
Dialogue Management	Preprocessing NLP pipeline	5.2.2
	Pattern matching	5.2.3
	NLP	5.2.3
	Response selection	5.2.4 6.2.2 6.2.3 6.2.4
Interaction Recording	Data (storage, real-time, …)	5.3
APIs	Text and speech	5.1.2 5.2.3 6.2.1
	NLU	5.2.3
	Authentication	5.1.2 6.4.5
	Language	5.1.3 6.1.3
	Redirection	6.3.1 6.3.2
	Requesting user input	6.2.5
	Payment service	6.4.4
	Sentiment analysis	7.1.1
Machine learning	Artificial Intelligence	7.2
Knowledge base	Filtering	6.4.1
	Bookmarking	6.4.2
	Feedback	6.4.3 7.1
Analytics	Bot Development Framework	7.3

Transversal to the components in the table above, there is a need to tackle some other important components, which can be found in Table 2.

This document recommends public administrations to pay a lot of attention to governance and security of a chatbot. Monitoring the usage and utilising user feedback are key features to create a qualitative chatbot. Users could also provide personal data through the chatbot, so a public administration should think about chatbot security and privacy.

Table 2: Overview of public administration chatbot components, part 2.

Architectural main component	Architectural component	Section(s)
Infrastructure	Bot Development Framework Cloud	5.1.1 5.1.4
Security	Multiple components	5.4
Privacy	Data	5.1.2 6.1.4 5.4

Next, it is recommended to use a thorough, cloud-based platform (see section 9.2), as a lot of architectural components that are needed to build a comprehensive chatbot are easily accessible on these (big) platforms. Furthermore, this document describes the possibility to provide multiple public services and the possibility to search for multiple topics in chatbots (see section 4.2.3).

To keep this feasible on a technical, economical and an organisational level, this document recommends to use a network of chatbots[6], where chatbots can redirect users to other chatbots (that are trained on different topics) in one integrated chat box.

This architecture could moreover work well on a European level: it fits perfectly in the Single Digital Gateway action. A European, federated chatbot network could provide autonomy to the Member States and the regions and could also support the public service portal Your Europe[7]. Providing a platform would make it possible to discuss standardising some design principles and intelligent layers to connect a European chatbot network.

[6] For an example, see http://startingupsmoothly.fi/

[7] https://europa.eu/youreurope/

2. INTRODUCTION

Over the past years, the increasing demand of information from citizens towards public administrations oblige the latter to go a step further in digital transformation. While public administrations are expected to know the nature of the public services they provide in details and their intended audience, they have also understood that they can only be effective and efficient by providing digital access to public services. The latest frontier of innovation in providing public services is by means of chatbots which have to be designed from a to z in order to satisfy citizens' needs.

2.1. Scope and objectives

The present document aims at identifying how public administrations can best provide public services via chatbots. The scope of this study is limited to a desk research on the usage of chatbots in the public sector and includes input from the 'Catalogue of Services' workshop, which took place on the 16th May 2019 in Brussels. The study is not intended to guide the reader into building a particular type of chatbot, nor to help the reader decide which tools he should work with, although some recommendations will be made for chatbots on a European level.

2.2. Structure of this document

The remainder of this document develops as follows.

Section 3 aims to explain the needs for using chatbots when providing public services.

Section 4 provides a short overview of different types of chatbots and the domains the chatbots can be active in.

Section 5 intends to provide awareness to public administrations regarding the aspects that need to be considered before starting a chatbot project. This section is divided into four distinct parts: (i) understanding the intended audience; (ii) understanding the user input data; (iii) understanding the data provided by the public administration; and (iv) securing your chatbot.

Section 6 focuses on how public administrations should think about user interactions when implementing a chatbot.

Section 7 intends to support public administrations in defining the governance of the chatbot.

Section 8 elaborates on different SAT views that should be taken as reference in the context of providing public services via chatbots.

Section 9 provides some recommendations based on all previous sections.

Finally, section 10 provides next steps for Member State public administrations and the European Commission to further develop chatbot solutions on all levels.

3. THE NEED FOR A CHATBOT (SEARCHING FOR PUBLIC SERVICES)

A chatbot is a computer program based on artificial intelligence that has the ability to conduct conversations via auditory or textual methods.

Chatbots are used in a wide variety of categories in the private sector like e-commerce, analytics, customer support, education, etc. In the public sector, chatbots provide several benefits for citizens and public administrations. Figure 2 shows the risen interest (through Google Search) in chatbot technology over a time span of ten years.[8]

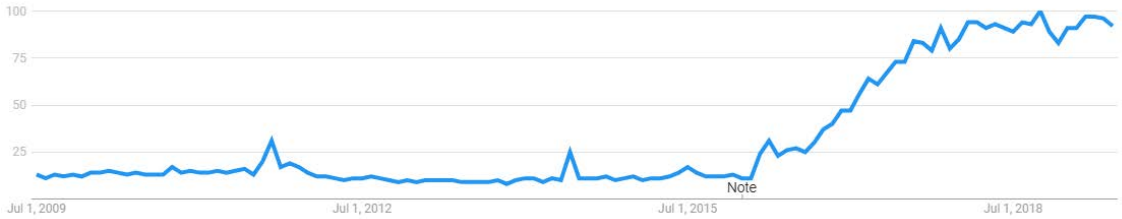


Figure 2: Interest in chatbots globally over a ten-year period.[9]

Citizens can make use of chatbots to find information easily, get assistance in their native language, save time and be able to get tailored services 24/7. Public administrations can benefit from using chatbots as they decrease workload and time for responses, deliver public services and address citizens issues easily, provide high availability and could provide multilingual support.

Furthermore, the data gathered by the chatbot can be analysed and public services provided by the government can be adapted or prioritised based on the needs of the citizens. If there is a clear need from citizens, as in a question that is asked often, but not yet captured in a public service, the public administrations can decide to adjust policy and cover those needs.

Developing a high quality chatbot needs resources. That is why it is important to identify real added value for public administrations when implementing this technology. A chatbot is an additional search tool next to a normal search bar on a website. Many chatbot benefits are present on a normal website (e.g. 24/7 availability, possibility to handle many users at once in contradiction to a helpdesk, etc.). Even the fact that chatbots easily integrate feedback in the chat, is or can be mirrored on public administration websites. The real differentiators are the decrease in time a user needs to search for (some) information or service and the chatbot analytics (e.g. what are the most used, specific user questions?). A user does not need to navigate a public administration's website anymore if he could find the information through a few questions to the chatbot. It is up to the public administrations to decide if the advantages of chatbots justify the costs of developing one.

---

[8] https://trends.google.com/trends/explore?date=2009-06-10%202019-06-10&q=chatbot

[9] A value of 100 is the peak popularity for the term.

3.1. Context

3.1.1. Public administrations

Local and national public administrations in the EU and in the rest of the world have adopted policies for the provision of public services. In particular, digital public services reduce the administrative burden on citizens and business and increase public administration efficiency as well. This has led to a global push to digitise public services.

In the EU for example, the Annex II of the proposal for regulation of the European Parliament and of the Council on "*establishing a single digital gateway to provide information, procedures, assistance and problem solving services and amending Regulation (EU) No 1024/2012"* [10], referred in Article 6(1), contains the list of public service procedures detailed in Table 16. The Regulation "*pursues a threefold* [objective][…]:

- *reducing additional administrative burden on citizens and businesses that exercise or want to exercise their internal market rights, including the free movement of citizens, in full compliance with national rules and procedures;*
- *eliminating discrimination; and*
- *ensuring the functioning of the internal market with regard to provision of information, procedures and assistance, and problem solving services*".

In addition, the European Union encourages the use of open data portals, starting from the implementation of the European Data Portal[11] to foster transparency of the Public Sector, regulated by the PSI Directive[12], which entered in force in December 2003, was revised in 2013 and was reached an agreement on for a new revision on 22nd January 2019.

Moreover, with the eGovernment Action Plan[13], the EU aims to improve the quality of public services while increasing public sector efficiency. It reaches this goal by removing existing digital barriers to the Digital Single Market and by joining efforts at EU level to support coordination and collaboration between Member States.

Not only the provision of information and public services by public administrations is being modernised around the world, but the provision of information by citizens and businesses is being modernised as well. The European initiative 'the Once-Only Principle (OOP)'[14], which aims to have citizens and businesses supply diverse data only once to a public administration, is one of these modernisations.

---

[10] http://data.consilium.europa.eu/doc/document/ST-10069-2018-INIT/en/pdf

[11] https://www.europeandataportal.eu/en/homepage

[12] https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32013L0037&from=FR

[13] https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52016DC0179&from=EN

[14] http://toop.eu/once-only

In the UK, the national government created a Government Transformation Strategy[15]. The strategy has three main components: transforming citizen-facing services, transforming full departments (to deliver public services in a more flexible way) and transforming the internal government to improve collaboration and deliver digitally enabled change more effectively.

The United States CIO Council on the other hand, provides a strategy[16] and 13 best practices with the aim of providing information and services to the American people anytime, anywhere and on any platform or device.

The Digital Government Strategies for Transforming Public Services in the Welfare Areas, an OECD comparative study[17], provides principles to guide the process of setting and implementing digital government strategies. The provision of digital government services will play a key role, as this will increase productivity and inclusiveness of service production and delivery.

3.1.2. Users

The 2018 State of Chatbots Report[18] researched and ordered the problems users have with traditional online experiences.

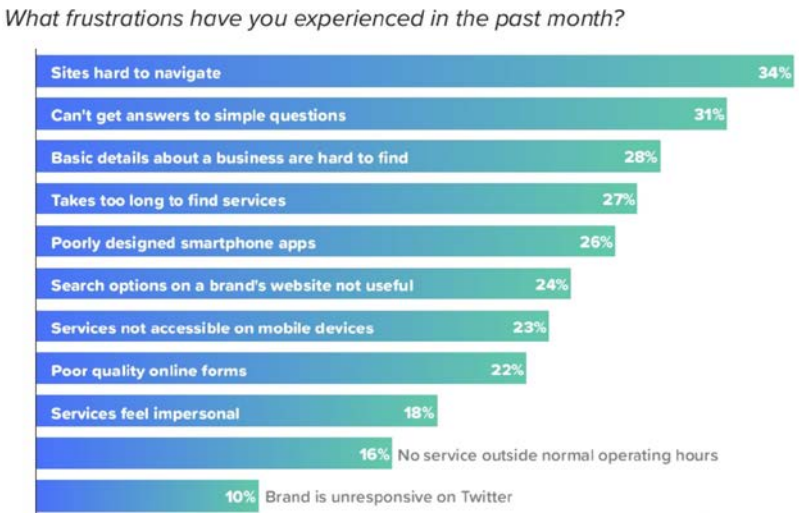


Figure 3: User problems with online experiences.[19]

As can be seen from Figure 3, the first reason of user frustration is that sites are hard to navigate. Because of this and the other user problems mentioned above, citizens might abandon their search. Therefore, a solution for these problems could be for

[15] https://www.gov.uk/government/publications/government-transformation-strategy-2017-to-2020/government-transformation-strategy
[16] https://www.cio.gov/fed-it-topics/modernization/digital-strategy/
[17] http://www.oecd.org/gov/digital-government/Digital-Government-Strategies-Welfare-Service.pdf
[18] https://www.drift.com/blog/chatbots-report/
[19] https://www.drift.com/blog/chatbots-report/

citizens to use interactive applications, which already exist in the market as messaging apps, to communicate with governments. The implementation of chatbots can be a possible answer.

Traditionally, chatbots have been implemented with the pure scope of information retrieval typically based on existing rules and standardised processes and domains (e.g. finding a flight for a precise day). However, nowadays, more advanced chatbots exist that can leverage machine learning capabilities to improve the communication with citizens. Moreover, chatbots have many other advantages like reducing human intervention, whilst having more customizable responses in comparison with text search. Figure 4 provides a short overview of chatbot advantages.

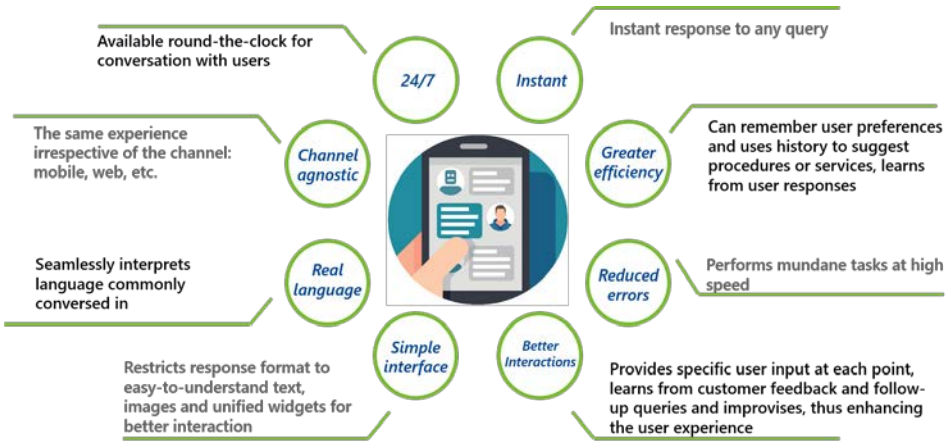


Figure 4: Chatbot advantages.

When people were questioned on these potential benefits of chatbots, some functionalities appear to be more useful than others, as can be seen in Figure 5.

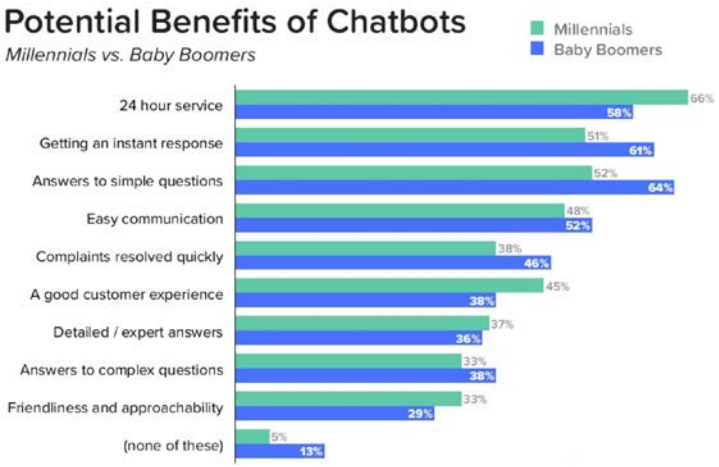


Figure 5: Millennials and baby boomers on potential benefits of chatbots.[20]

[20] https://www.drift.com/blog/chatbots-report/

3.2. From public service catalogues to chatbots

3.2.1. Catalogues of public services

Nowadays, several public administrations have implemented public service catalogues in order to speed up the search of public services by their citizens.

In Europe, there already exist different examples of public service catalogues:

- In Finland, the Finnish Service Catalogue[21], which is mandatory to use by law (the law on common administrative e-service support services 571/2016[22]) and by all public sector organisations, thus containing the information of all public sector services and the service channel. Such catalogue can be searched by keywords and can filter out the intended audience (citizens, businesses and public administrations);
- The French Government has created an open data catalogue[23] and an API[24] portal[25] as well, which can be used by digital administrations to speed up the provision of public services by searching for keywords;
- The Government of Norway created a central catalogue of public services, which includes datasets, APIs, concepts and information models[26];
- The UK Government created a catalogue of public services, where users can dig in all the details of public services, discovering by which organisation the service is delivered or which duties are required by means of Linked Data (e.g. civil marriage and civil partnership venue licence[27]).

Despite that the information provided by the catalogues can be accurate, citizens are required to search through the catalogues, which might be a tedious activity. Additionally, in many cases, the development of these catalogues has not been harmonised inside the country and between different countries.

3.2.2. CPSV-AP

As explained previously, the Core Public Service Vocabulary Application Profile[28] was developed. CPSV-AP is a data model that has been developed in the context of the ISA² Catalogue of Services Working Group for describing public services.

A Core Vocabulary is a simplified, reusable and extensible data model that captures the fundamental characteristics of an entity in a context-neutral fashion. Core Vocabularies[29] are the starting point for agreeing on new semantic interoperability assets and defining mappings between existing assets to guarantee a level of cross-

[21] https://www.suomi.fi/frontpage
[22] http://www.finlex.fi/fi/laki/ajantasa/2016/20160571
[23] https://www.data.gouv.fr/fr/
[24] Application Programming Interface
[25] https://api.gouv.fr/
[26] https://fellesdatakatalog.brreg.no/
[27] https://standards.esd.org.uk/?uri=service%2F1296&tab=details
[28] https://joinup.ec.europa.eu/solution/core-public-service-vocabulary-application-profile/about
[29] https://joinup.ec.europa.eu/solution/e-government-core-vocabularies

domain and cross-border interoperability that can be attained by public administrations.

An Application Profile is a specification that re-uses terms from one or more base standards, adding more specificity by identifying mandatory, recommended and optional elements to be used for a particular application, as well as recommendations for controlled vocabularies to be used.

Using CPSV-AP would be a necessary step to develop public service chatbots as it links different information concerning public services that is important in a chatbot. An example of how a public administration chatbot could look like can be found in section 5.2.4. Moreover, the model would be useful on a European level, as harmonising the data is an absolute minimum requirement for a European chatbot to work efficiently and effectively. The next section provides a short overview of European chatbot possibilities.

3.3. Single Digital Gateway

As mentioned in section 3.1, the EU is digitising public services through the Single Digital Gateway Regulation. Your Europe[30], the EU citizens and businesses portal has a search function which will give access to:

- Information: Citizens will be able to easily find reliable, qualitative information on EU and national rules that apply to them when they want to exercise their Single Market rights;
- Procedures: Citizens will find out exactly how to carry out administrative procedures and what steps they need to follow. The EU Member States will have to provide the 21 procedures (see Table 16) digitally by 2023;
- Assistance services: If users are still confused about which rules apply or have trouble with a procedure, they will be guided to the EU or national assistance service most suited to address their problem.

This subsection considers how chatbot technology could support the SDGR[31] action. Figure 6 shows the current Your Europe portal. The portal provides a clear overview of themes for the citizen to choose from. As more information will be included on this portal in the future, this clear overview might be lost and will cause users to spend more time searching for the right public service or information.

[30] https://europa.eu/youreurope/index.htm

[31] Single Digital Gateway Regulation

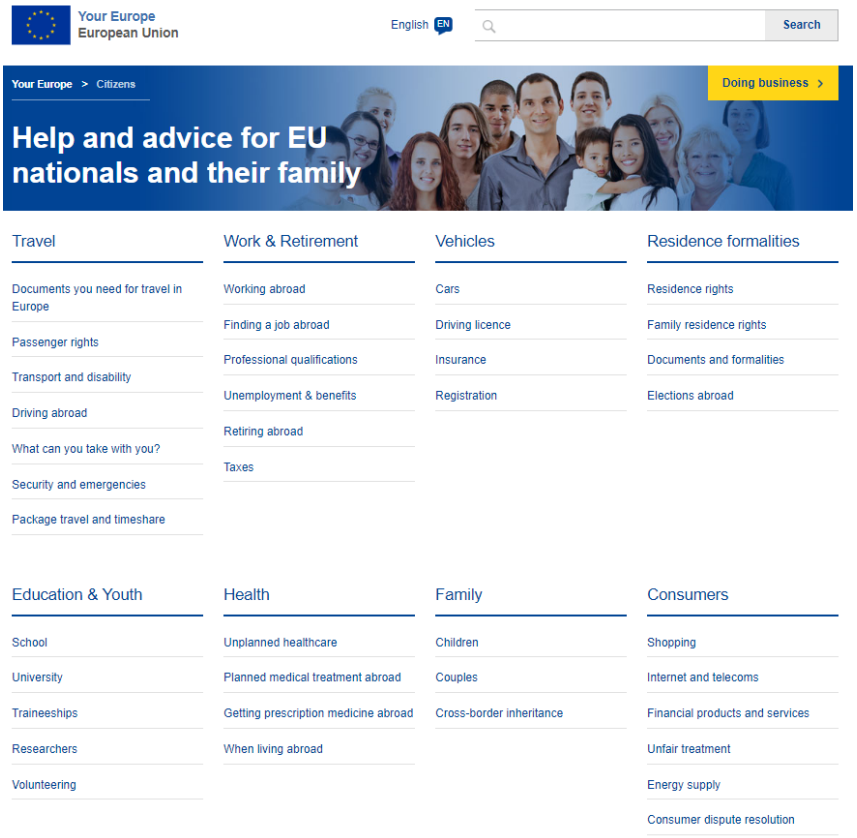


Figure 6: Your Europe portal, citizens' view.[32]

Chatbot technology could be a way to decrease search time here. Two specific solutions will be described in the next paragraphs: one general chatbot that could answer any question and a network of public service chatbots.

3.3.1. One chatbot to rule them all

One possibility is to develop a chatbot that can answer any question on public services of any EU Member State. In section 4.1.3, this 'generative' chatbot is described as complex and costly to build. During the workshop on the Catalogue of Services of 16th May in Brussels, a clear opinion was voiced by Member State representatives: one European chatbot is not a preferred solution as it would decrease the independency of the EU Member States, the coordination effort to create such a chatbot would be huge and the cost to develop it would be enormous.

3.3.2. Network of public service chatbots

The other chatbot solution option was found more plausible by the workshop participants. The European Commission could provide a chatbot that connects member state chatbots through an intelligent top layer. In section 4.1.2, this document describes the existing Finnish chatbot network. A European chatbot network would be more complex and could need some extra functionalities to manage the additional complexities, e.g. a top level to choose a specific country.

[32] https://europa.eu/youreurope/citizens/index_en.htm

One way to implement this chatbot network would be to set up some basic design principles (e.g. same interface or chatbot names) to easily integrate the chatbots and refer to other chatbots without the need for changing a webpage.

3.3.3. Cooperation

During the workshop, the demand for cooperation arose. Especially when an initiative on a European level is considered, there will be a need for close cooperation. A proposition came forward to (1) create a European platform to bring European Member State chatbot stakeholders together, (2) exchange best practices while building a chatbot and (3) reach consensus on chatbot design principles in case a European chatbot network would be considered.[33]

[33] Slides from the workshop and a summary of the results can be found on Joinup: https://joinup.ec.europa.eu/collection/semantic-interoperability-community-semic/event/technical-workshop-catalogue-public-services-vis-vis-sdg-implementation-16-may-2019-930-1730-cet

4. SCOPE OF THE CHATBOT

There are different types of chatbots to cater specific needs of the provider. To decide what type of chatbot a public administration should work with, different perspectives should be kept in mind. The organisation's resources and the amount of domains and/or services an organisation wants to provide via chatbot(s) are the most important perspectives.

4.1. Type

This section will describe five types of chatbots. Types could be combined, e.g. a 'goal-oriented' 'conversational' chatbot.

4.1.1. Goal-oriented chatbot – Retrieval-Based[34]

This architectural model of a chatbot is easier to build and much more reliable than a generative chatbot. Though there cannot be 100% accuracy of responses, the possible types of responses are known and ensure that no inappropriate or incorrect response is delivered by the chatbot.

Currently, retrieval-based models are more in use than the generative models. Several algorithms and APIs are readily available for developers to build chatbots on this architectural model. This bot[35] considers the message and context of the conversation to deliver the best response from a predefined list of messages.
One limitation of using a goal-oriented chatbot is that it only addresses one topic: users still have to find the right service (chatbot) before diving deeper in the questions.

4.1.2. Network of goal-oriented chatbots

A possible way to fill this functionality gap of a goal-oriented chatbot is to refer users to another specific chatbot if a question is asked about another topic. A benefit of a chatbot network is that it makes things easier for a user who may not know the organisation he should get in touch with. Figure 7 shows how Finland uses a network of three chatbots for residence permits, setting up a company and an information chatbot on business and work-related taxes.

[34] https://dzone.com/articles/understanding-architecture-models-of-chatbot-and-r
[35] Short for chatbot.

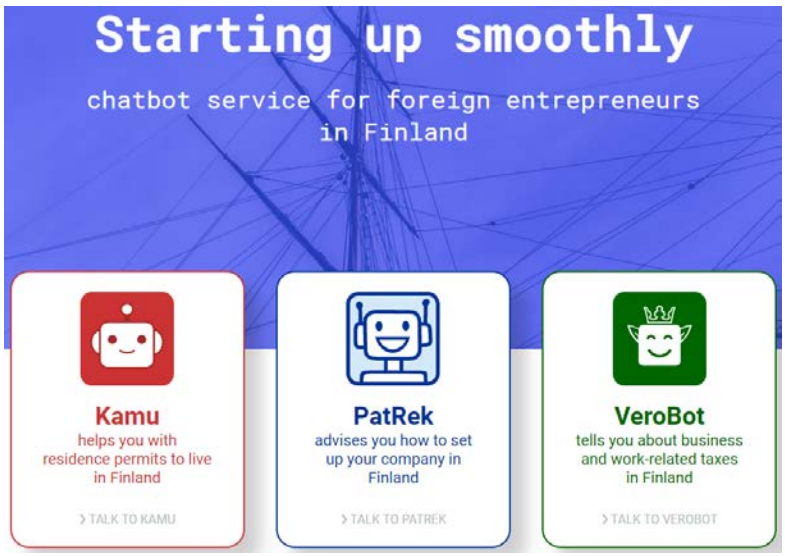


Figure 7: Network of three chatbots[36], Finland – Homepage

Users can start a chat with one of the chatbots. The chatbots can refer the user to another chatbot if necessary. These referrals are called transfers. There are currently three types of transfers.

- The **reactive transfer** is something that is triggered when the current bot does not have an answer but based on the key words in the user's question, the AI predicts the right organisation and chatbot.
- The **proactive transfer** is something that is built into the dialogues based on the understanding of the user needs. If there is more information the user might need and another organisation is responsible for that information, the bot offers that transfer proactively.
- Finally, **the manual transfer** is something that the users can trigger themselves by typing the name of the bot or the organisation.

In Figure 8, chatbot PatRek from the Finnish Patent and Registration office, using the reactive transfer protocol, asks if the user wants to be redirected to the VeroBot chatbot of the Finnish Tax Administration. If the user agrees, the webpage and interface stay the same, except a change of color and a change in chatbot name, so the user knows he is talking to another chatbot.

[36] http://startingupsmoothly.fi/

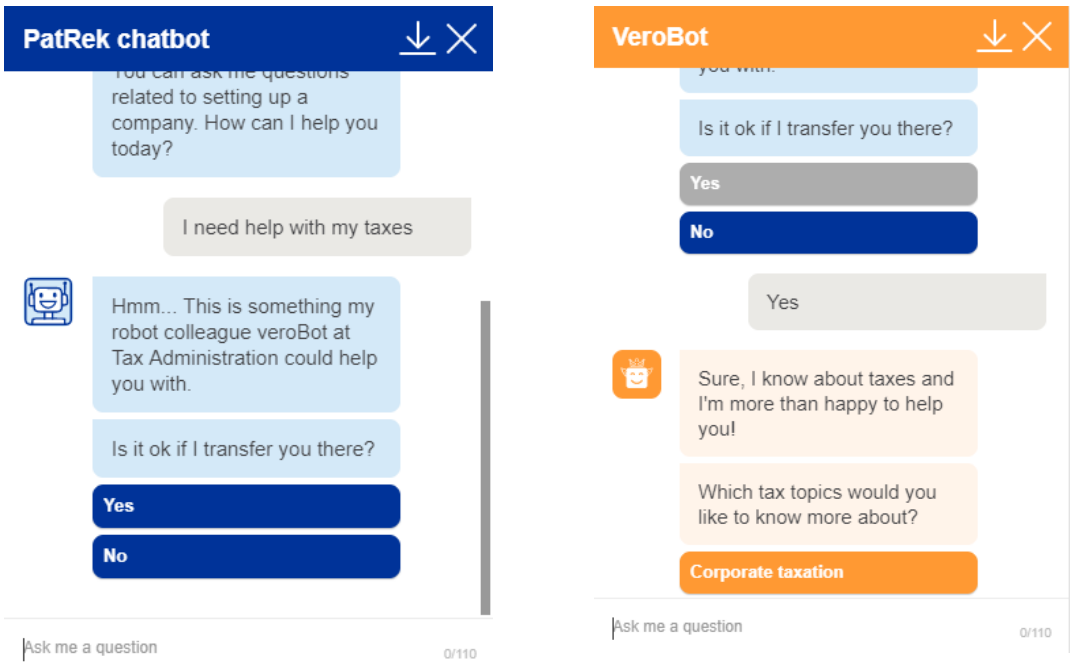


Figure 8: PatRek Chatbot, Finnish patent and registration office, switching to the VeroBot chatbot, Finnish Taxation Administration.

Starting up smoothly[37], the Finnish chatbot platform, is a collaboration between the Finnish Immigration Service (Migri), the Finnish Tax Administration (Vero) and the Finnish Patent and Registration Office (PRH), to pilot a common chatbot service, which answers questions related to starting a company, taxation and immigration to Finland.

Each organisation has its own independent chatbot: Kamu (Migri), VeroBot (Vero) and PatRek (PRH). Each organisation creates its own content and trains their chatbot independently (see Figure 9). These independent chatbots are linked to each other through an intelligent layer in the background. Collaboration between organisations is needed, but less than if all the content creation and training would happen in one chatbot.

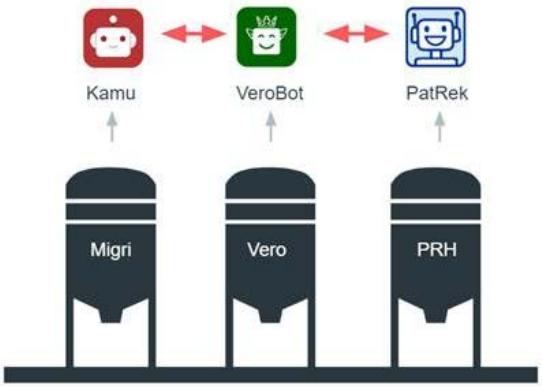


Figure 9: Finnish chatbot network architecture.

[37] http://startingupsmoothly.fi/

4.1.3. Generative chatbot

Generative chatbots are advanced in nature. This type of chatbot is very rarely used, as it requires the implementation of complex algorithms. Generative models are comparatively very difficult to build, develop and operate. Training this type of bot requires investing a lot of time and effort by giving millions of examples. This way, the deep learning model can engage in conversations. However, it is still unsure what responses the model will generate. An example of this kind of chatbot is Google Assistant.

4.1.4. Scripted / sequential chatbots

Scripted chatbots are pre-defined with a conversational flow. When a user throws a query, the chatbot responds with a pre-defined script from the library. The chatbot does not and cannot go out of this scope, meaning a user cannot ask questions like 'How are you?' and 'What's the weather like today?' A scripted chatbot is generally useful when the maker wants to automate tasks like filling a grievance, logging feedback[38], etc. Figure 10 shows chatbot Rammas, an example of a scripted chatbot.

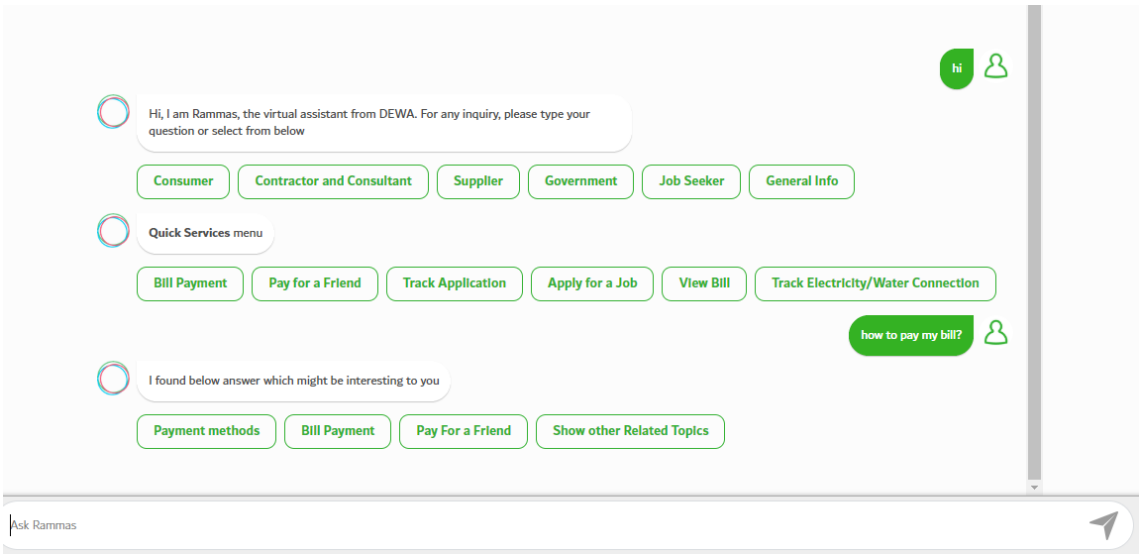


Figure 10: Chatbot Rammas, Dubai Electricity & Water Authority – Example of a scripted chatbot.

4.1.5. Conversational chatbots

Conversational chatbots use NLP[39] (see section 5.2.3) to answer user queries. They provide users with the opportunity to ask any question. This kind of chatbot is trained to user's queries and makes a pattern to responds back to them with the most reliable answer possible, see e.g. Figure 11.

[38] User feedback will be an important functionality of the SDG. A short example on user feedback is given in section 6.4.3. A detailed report on chatbot feedback possibilities can be found in section 7.1.

[39] Natural Language Processing

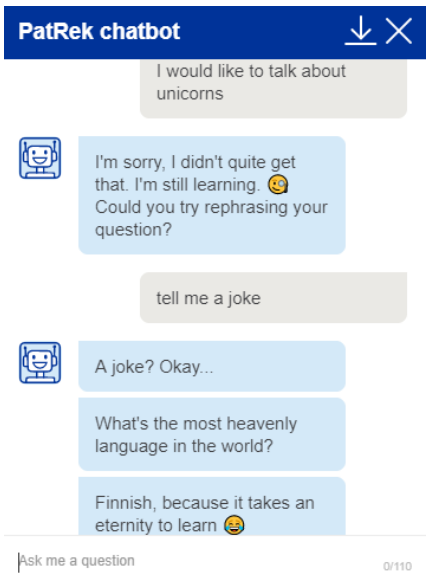


Figure 11: Chatbot PatRek – Conversation.

4.2. Domain

4.2.1. Information

Most existing public administration chatbots[40] provide information about public services to the user. Some chatbots, like Noa (see Figure 12), combine the knowledge of multiple public services and procedures on starting a business in one chatbot.

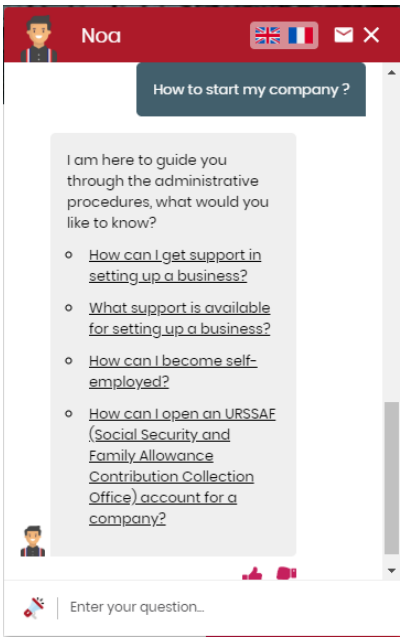


Figure 12: Chatbot Noa, Préfecture d'Île de France - Providing information on administrative procedures and public services.[41]

[40] See Table 15, which provides an overview of researched chatbots.

[41] https://french-tech-central.com/

4.2.2. Service delivery

Some public administrations are experimenting to deliver services with chatbots. For Portuguese residents, it is possible to change their official address via chatbot Sigma. In Woluwe-Saint-Pierre, a payment service is included in a chatbots, for a specific public service. Information delivery as a service is possible as well, as shown in Figure 13.

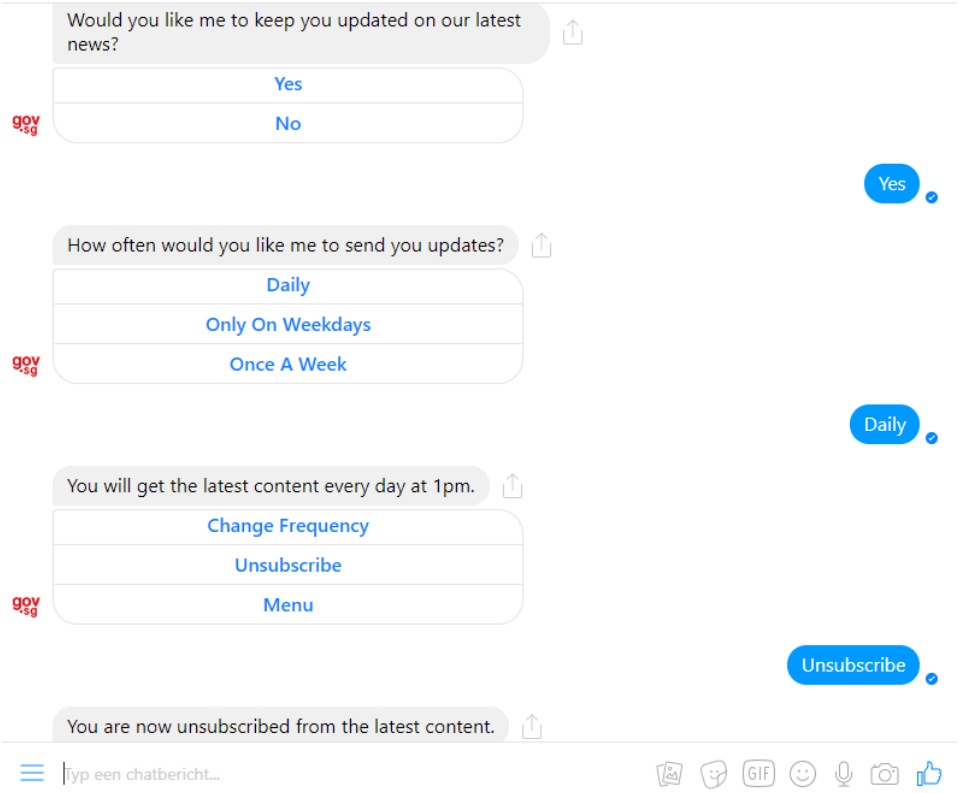


Figure 13: Gov.sg chatbot, Singapore Government - Example of providing an information subscription service via a chatbot.[42]

4.2.3. Combination

Combining information provision and service delivery within one chatbot is possible, but makes the chatbot much more complex as it moves focus from a goal-oriented chatbot to a more generative chatbot (see section 4.1). A solution would be to refer to other (integrated) chatbots whenever another functionality is needed (see section 4.1.2).

[42] https://www.facebook.com/pg/gov.sg/about/?ref=page_internal

5. BUILDING UPON

This section intends to provide awareness for public administrations before implementing a chatbot in terms of:

- Leveraging knowledge about the audience of a chatbot in order to customize it to user needs.
- Being mindful when providing and using data via chatbots.

5.1. Know your users

5.1.1. Platform

To know what platform(s) to use for a public administration chatbot, the foremost question that should be asked is “How will the users access the chatbot? Is the bulk of the users active on mobile, laptop, tablet or voice assistant?”
Chatbots can be built on multiple platforms. Platforms who provide a messaging app as a service might have a slight advantage, as the amount of users on the big four messaging apps, see Figure 14, is growing very fast. A messaging app is thus a great location to reach users.

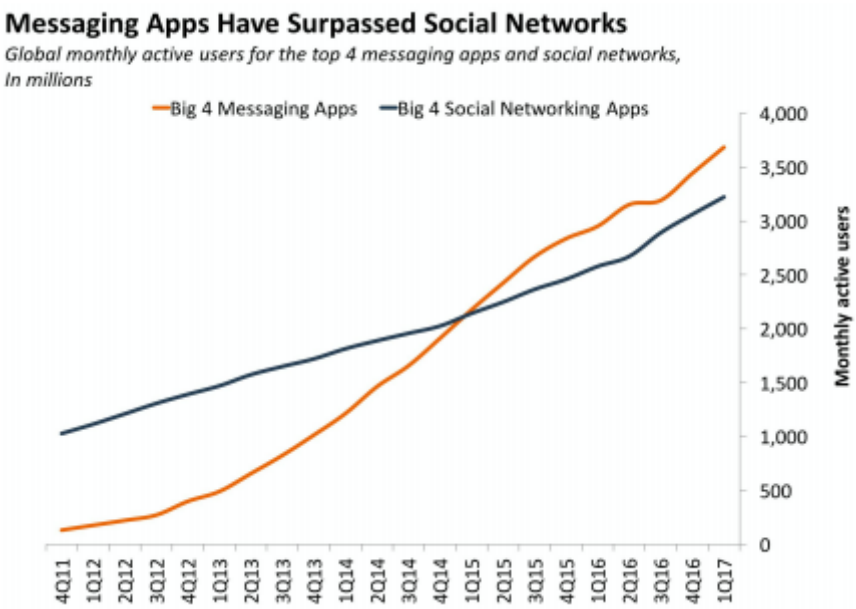


Figure 14: The four most used messaging apps in the world.[43]

[43] Source: Companies, apptopia, TechCrunch, BI Intelligence estimates, 2017 (WhatsApp, WeChat, Facebook Messenger and Viber).

- Facebook Messenger

Facebook provides a functionality to integrate chatbots on its platform. This functionality is also integrated in the Facebook Messenger app, making the chatbot easy accessible online and via mobile. Figure 15 shows a screenshot of using two public administration chatbots on the Facebook platform.

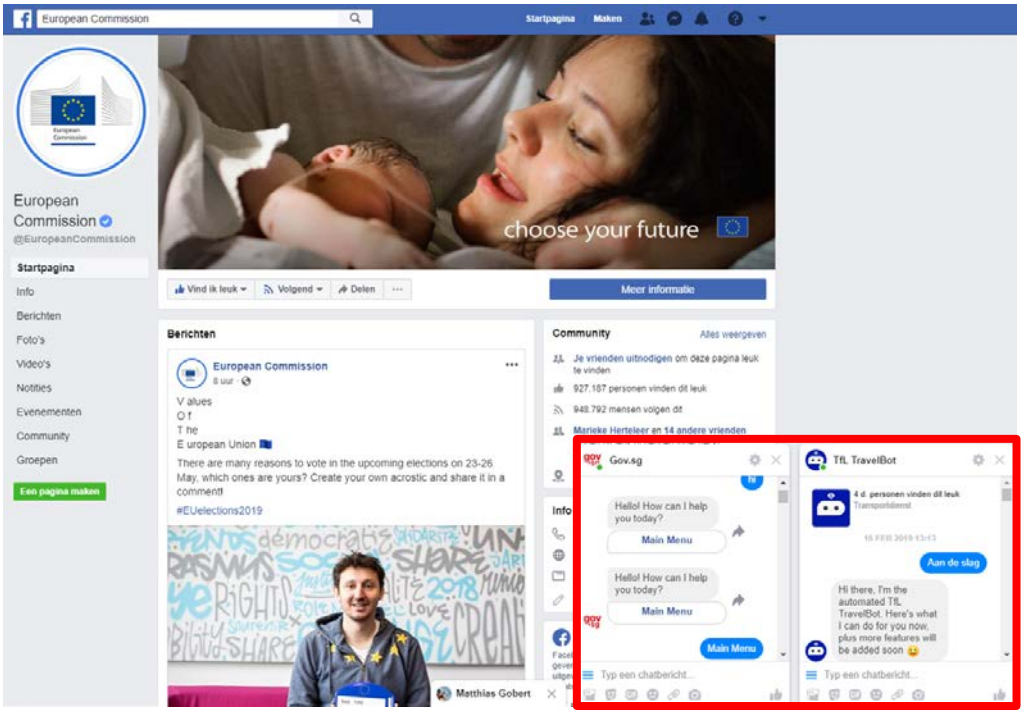
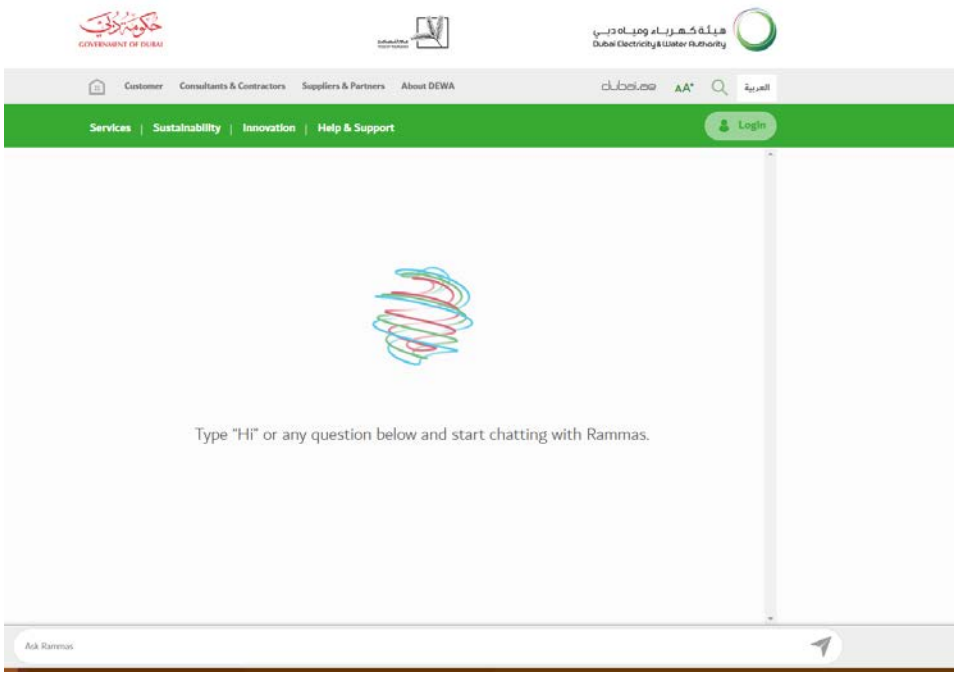


Figure 15: Chatting with two public administrations via Facebook.

- Custom page, full page

Some public administrations provide their chatbot through a dedicated page on their platform, e.g. Figure 16.

Figure 16: Chatbot Rammas[44], Dubai Electricity & Water Authority – Homepage.



- Custom page, multiple chatbots

Very similar to the previous example, the Finnish chatbot network uses a custom page with an overview of all three chatbots (Figure 17). When a user clicks on one chatbot, he is redirected to the relevant public administration website with a chatbot window open.

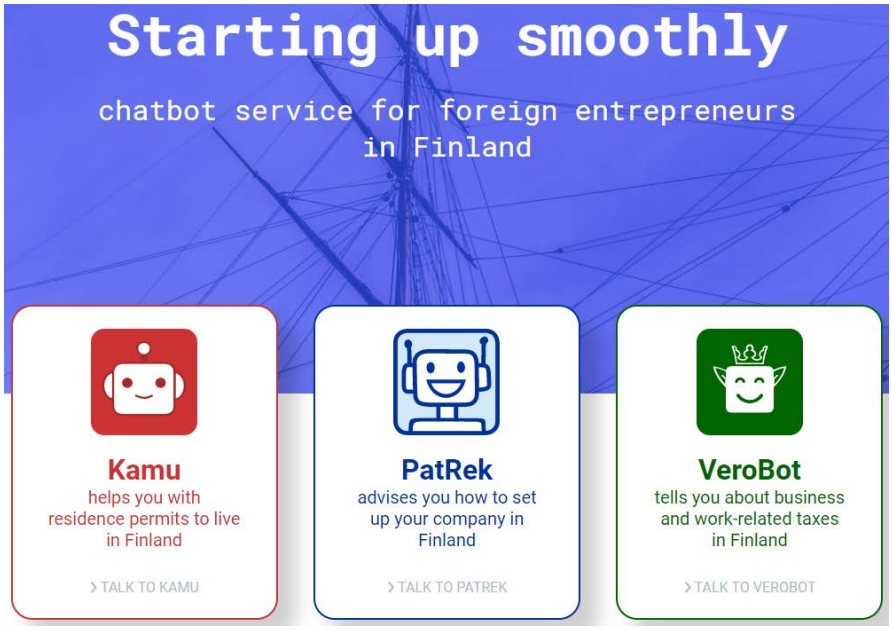


Figure 17: Network of three chatbots, Finland – Homepage.[45]

[44] https://www.dewa.gov.ae/en/rammas
[45] http://startingupsmoothly.fi/

- Normal website, bar

  Most researched public administrations use some kind of pop-up visual to provide access to the chatbot. Chatbot Alex (Figure 18) uses a bar, similar to a minimised chat window in a messaging app.

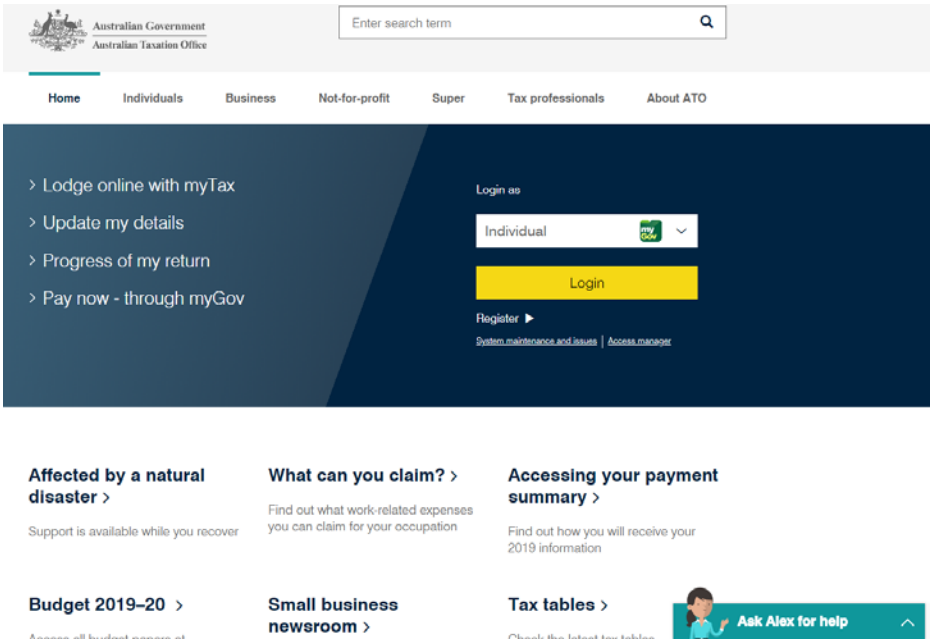


Figure 18: Chatbot Alex, Australian Taxation Office – Homepage.[46]

- Normal website, icon

  The use case, shown in Figure 19, is very similar to the previous example, but shows an icon to open the chat and not a chat bar.

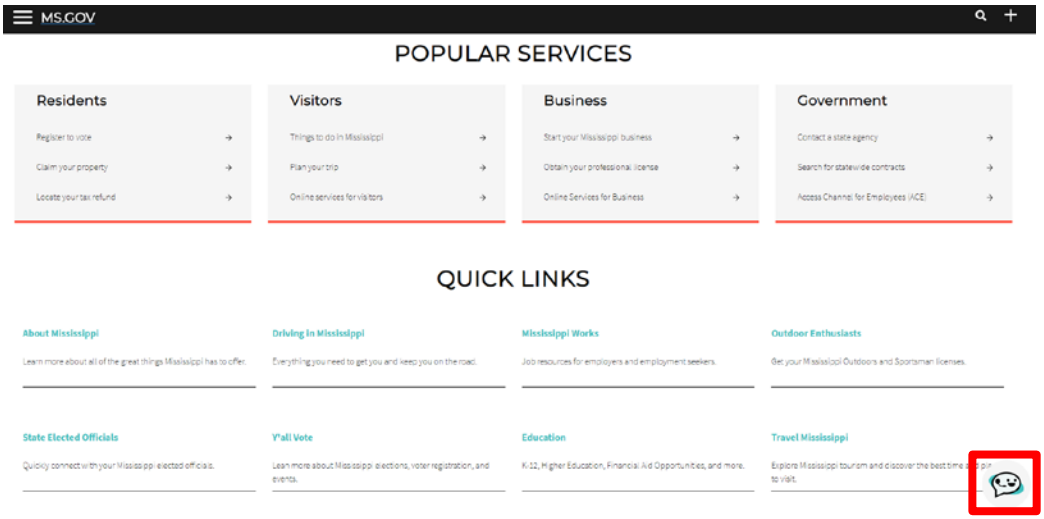


Figure 19: Chatbot MISSI, Mississippi Government – Homepage.[47]

[46] https://www.ato.gov.au/

[47] https://www.ms.gov/Technology

- Phone

Next to Facebook, other public administrations use a mobile platform for their chatbot. Chatbot Citibot (Figure 20) uses an sms service to communicate to users through their phones.

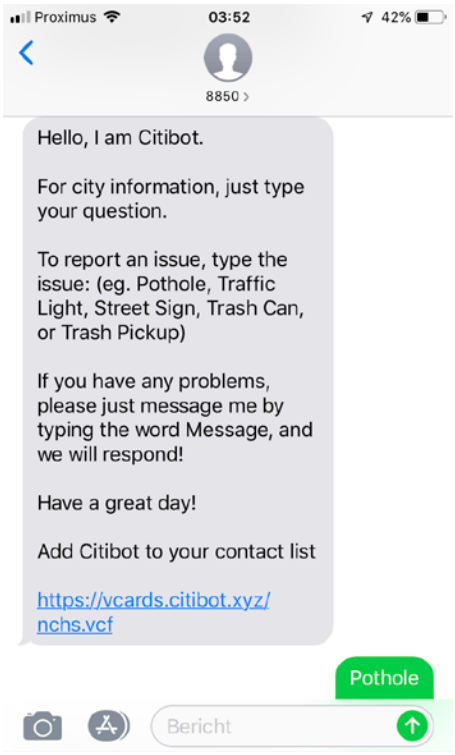


Figure 20: Chatbot Citibot, City of North Charleston – Opening message.[48]

- Voice Assistants

Smart assistants like Google Assistant, Cortana and Amazon Echo could be used when the audience prefers voice-enabled services. Public administration chatbot MISSI has integrated functionalities with the two major voice assistants of the moment: Alexa and Google Assistant. On the Mississippi government site, it is possible to integrate the MISSI functionalities with the voice assistants as shown in Figure 21.

[48] https://www.northcharleston.org/connect/

Figure 21: Chatbot MISSI – Integrating chatbots functionalities in voice assistance providers. [49]

Finally, Table 3 provides a list of chatbot development frameworks, which can be used to integrate a chatbot on a public administration portal (e.g. Figure 22: Chatbot framework by Dialogflow).

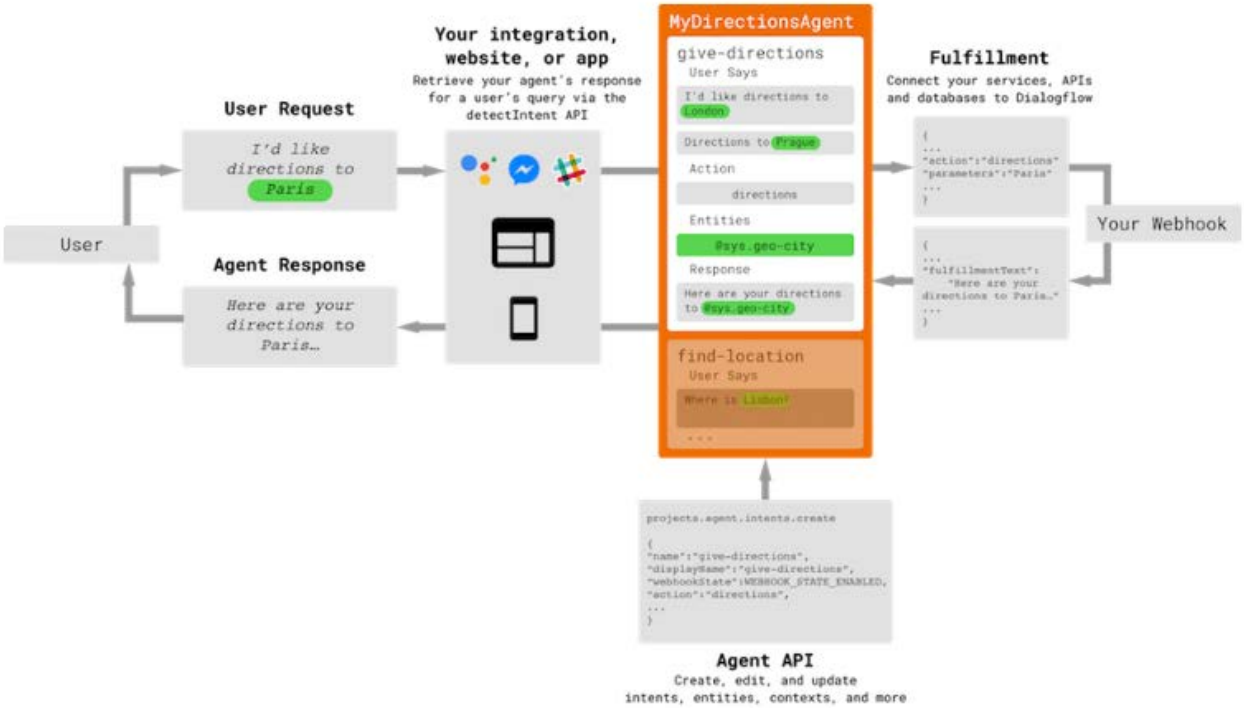


Figure 22: Chatbot framework by Dialogflow.

Table 3: Non-exhaustive list of chatbot development frameworks.[50]

Framework	Link
Microsoft Bot Framework	https://dev.botframework.com/
Wit.ai (Facebook)	https://wit.ai/
Dialogflow (Google)	https://dialogflow.com/
IBM Watson	https://www.ibm.com/watson/how-to-build-a-chatbot

[49] https://www.ms.gov/Technology

[50] https://geekflare.com/chatbot-development-frameworks/

Pandorabots	https://home.pandorabots.com/home.html
Botpress	https://botpress.io/
Botkit	https://botkit.ai/
RASA Stack	https://rasa.com/
ChatterBot	https://github.com/gunthercox/ChatterBot

5.1.2. User input data

This section aims to provide a short overview on what sorts of input data a user can provide in chatbots. Later, in section 6.2, different possibilities are given concerning text and voice options.

The main types of input data given by a user are:

- Text and voice

  Users can create data input in different ways. The most used input for chatbots is input via text. As mentioned above, it is also possible to have a voice input via a voice assistant. This complicates the chatbot but is especially useful for public administrations to increase the accessibility of public services for people with disabilities (e.g. blind people). For a voice-enabled chatbot, there is a need for:
  - Speech recognition (if a custom voice assistant is used)
  - Speech-to-text software
  - Text-to-speech software

- Structured vs unstructured data

  A (scripted) chatbot can provide a range of possible answers via buttons, which are linked to structured data. A conversational chatbot uses NLP (see 5.2.3) to translate the unstructured user input into structured data.

- Authentication data

  For public administration chatbots, authentication could be useful to deliver personalised services and personalised information. There exist multiple solutions, e.g. eID in Belgium, tokens, mobile authentication app linked to a user's eID (e.g. itsme[51]), etc. To develop such a functionality, the public administration could base itself on the eIDAS[52] principles. In Figure 23, the public administration chatbot Sigma notices that the user who wants to change his address is not authenticated. To proceed with this action in the chatbot, the chatbot asks the user to authenticate himself in the ePortugal portal with a Citizen Card or with the Digital Mobile Key.

[51] https://www.itsme.be/en/lang-splash

[52] https://ec.europa.eu/digital-single-market/en/trust-services-and-eid

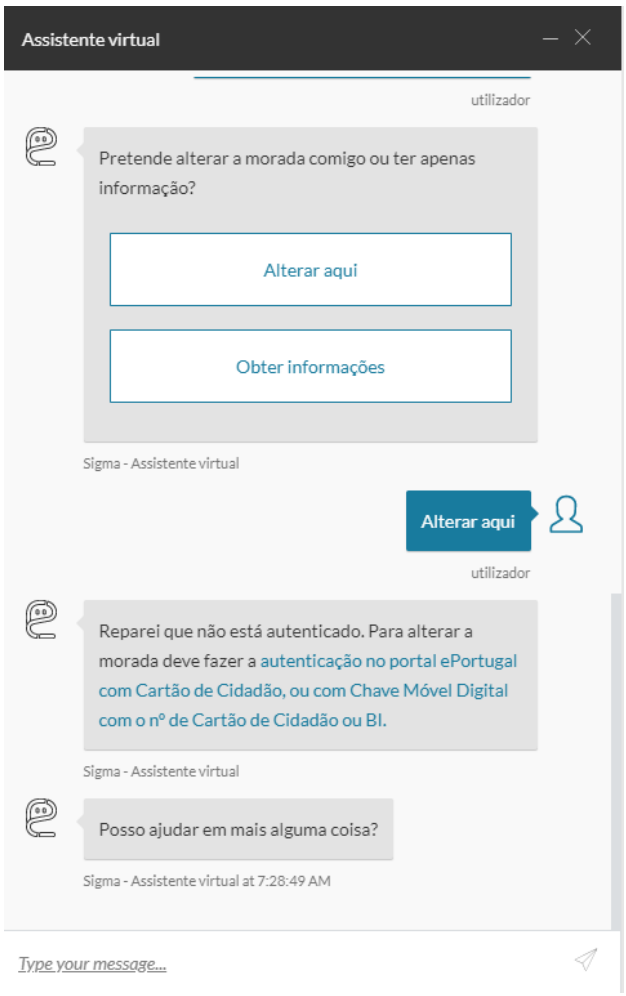


Figure 23: Chatbot Sigma, ePortugal – Requesting authentication.[53]

- Personal data

  In line with authentication, users can provide additional personal data in some cases. Travelbot can request a user's location for example (see Figure 24).

[53] https://eportugal.gov.pt/

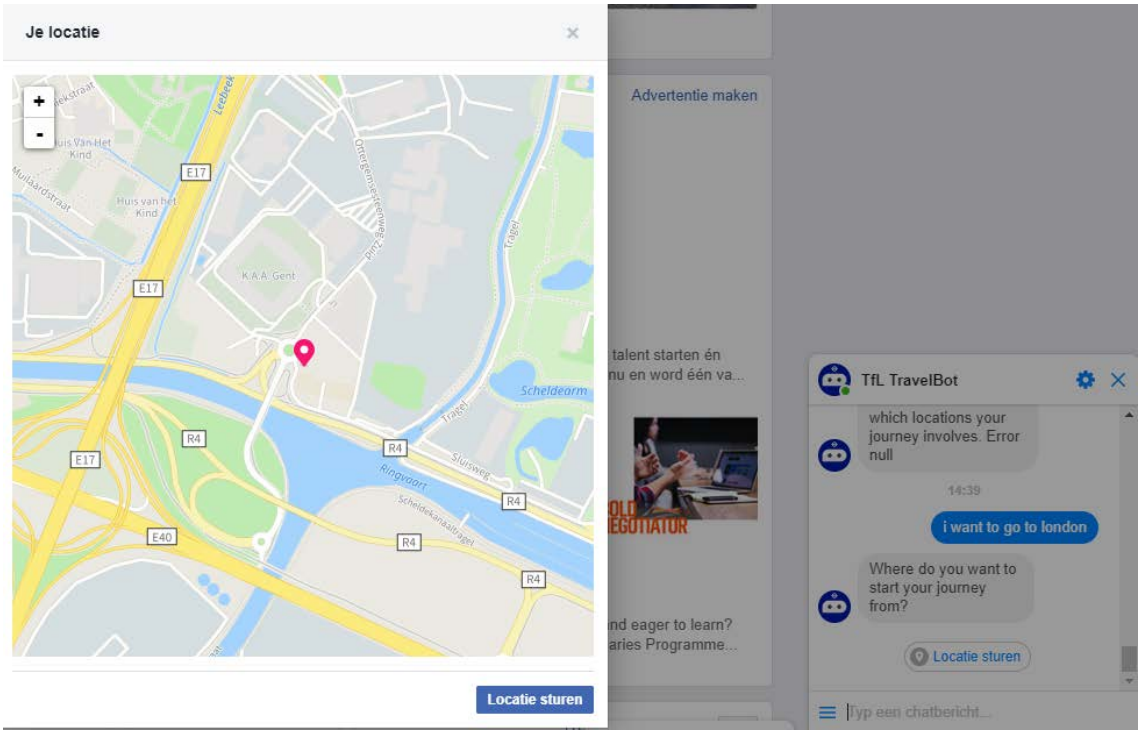


Figure 24: Chatbot Tfl, Transport for London - Travelbot Tfl requesting location data of the user.

When personal data is in play, a chatbot developer should take the European GDPR[54] into account. Personal data should be handled and processed correctly to comply with this regulation.

GDPR & IDPR
The principles of data protection should be kept in mind from the moment of conceptualization, through the implementation and end of the chatbot. The most relevant principles are: • Data protection by design o Limit the collection and processing of personal data to what is relevant and strictly necessary for the purpose(s). (**data minimization**) o Limit the purpose for which the chatbot collects and processes the personal data to what is strictly necessary for the chatbot. Make sure the purpose(s) are specified, explicit and legitimate. (**purpose limitation**) o Delete or anonymise personal data as soon as they are no longer necessary for the purposes. (**storage limitation**) o Implement the appropriate security measures[55]. (**integrity and confidentiality**)

[54] https://eur-lex.europa.eu/eli/reg/2016/679/oj

[55] See section 5.4 for more details.

- Data protection by default
  - The strictest privacy settings should automatically apply. Data subjects can actively choose (give consent) to allow their data to be processed for more purposes, to enhance the capabilities of the chatbot.

If the European Commission would like to develop a chatbot (e.g. on the Your Europe portal) it should comply with the IDPR[56], which is very similar to the GDPR.

5.1.3. Language

Depending on the user base of the chatbot and the resources available, different language options should be explored:

- Start the conversation with a language option menu when having a mixed language audience;
- Start the conversation in the original language with an option to switch to English;
- Start the conversation in the original language; or
- Start the conversation in English.

5.1.4. Timing

As seen in Figure 5, one of the greatest benefits of a chatbot is the 24/7 availability. This means there will be peaks and troughs in the usage of the service. To handle peaks, enough service capacity needs to be available. A cost-efficient way to have enough server capacity is to use cloud-based resources, which would scale depending on the usage.

5.2. Understanding input text and providing responses

5.2.1. Introduction

Chatbots can handle messages and give responses very differently. This depends on a range of factors: the purpose of the chatbot, the user input data, the way of classifying data, etc.
Figure 25 provides a general overview of the different interactions in a chatbot architecture.

[56] https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32018R1725

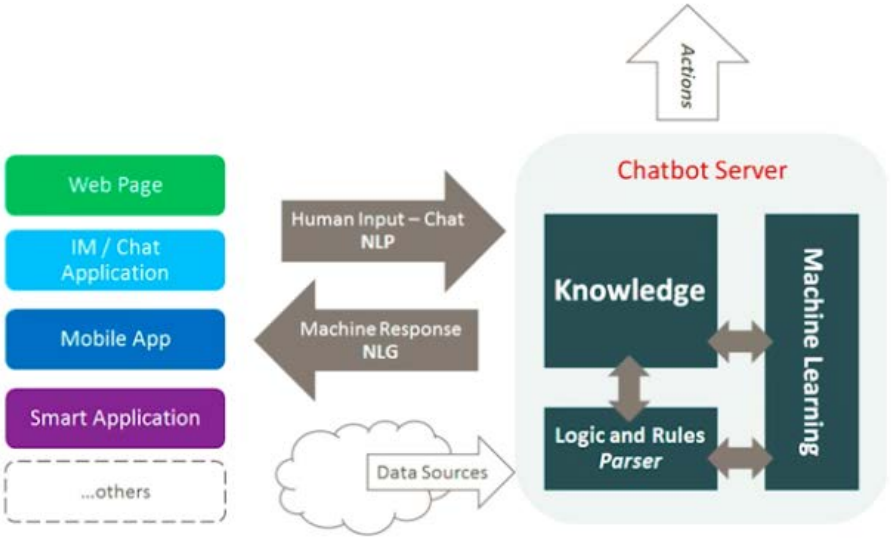


Figure 25: Anatomy of a chatbot.[57]

This section provides an overview of the different steps needed in the chatbot process to come to a useful response. Figure 26 provides an example using Natural Language Understanding and generated responses.

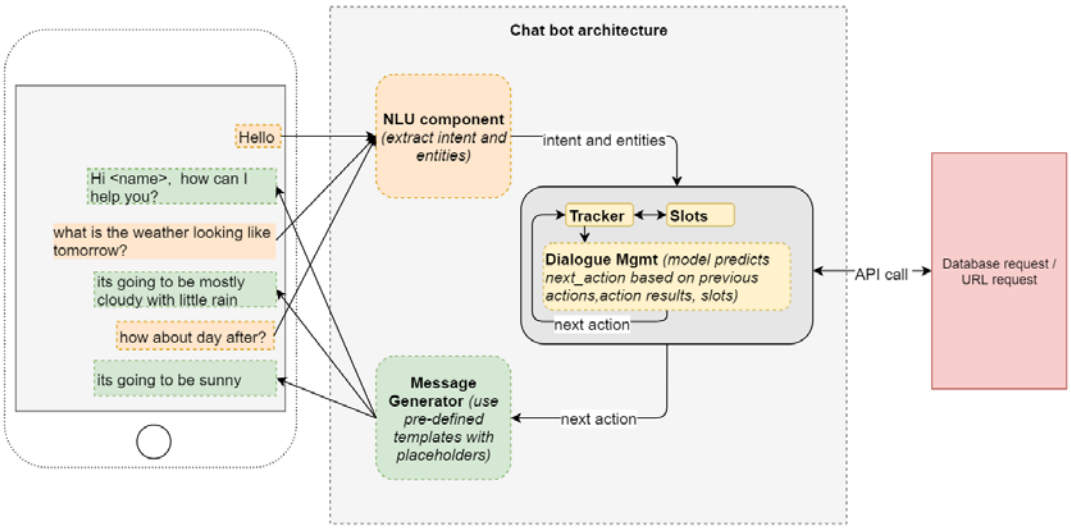


Figure 26: Example of a chatbots architecture handling messages.[58]

5.2.2. Preprocessing

A chatbot needs a preprocessing NLP pipeline to handle typical errors. Table 4 below describes the nine steps[59] to preprocess data.

---

[57] https://medium.com/analytics-vidhya/building-a-simple-chatbot-in-python-using-nltk-7c8c8215ac6e

[58] https://towardsdatascience.com/architecture-overview-of-a-conversational-ai-chat-bot-4ef3dfefd52e

[59] https://medium.com/@surmenok/natural-language-pipeline-for-chatbots-897bda41482

Table 4: Chatbot preprocessing steps.

Preprocessing steps	
Spellcheck	Raw user input data can contain spelling errors. A spell checker can be used here.
Split into sentences	It could be helpful to analyse every sentence separately. NLP libraries (e.g. NLTK[60], StandfordNLP[61], SpaCy[62]) can be used to split text into sentences.
Upper/lower case	Converting the entire text into uppercase or lowercase, so that the algorithm does not treat the same words differently.
Removing noise & stop words	Everything that is not a standard number or letter should be removed. Sometimes, some extremely common words, which would appear to be of little value in helping select information matching a user need, are excluded from the vocabulary entirely. These words are called stop words.
Split into words	Hardcoded rules typically operate with words. This can be done with the NLP libraries, mentioned above.
POS tagging	Some words have multiple meanings. Knowing more context can help to disambiguate the meaning. The same NLP libraries can be used, or Google SyntaxNet[63], that is a little bit more accurate and supports multiple languages.
Lemmatise words	One word can have many forms: e.g. ‘pay’, ‘paying’ and ‘paid’. In many cases, the exact form of the word is not important for writing a hardcoded rule. If preprocessing code can identify a lemma, a canonical form of the word, it helps to simplify the rule. Lemmatisation, identifying lemmas, is based on dictionaries, which list all forms of every word. The most popular dictionary for English is WordNet[64]. NLTK and some other libraries allow using it for lemmatization.
Entity recognition	Dates and numbers can be expressed in different formats: “3/1/2016”, “1st of March”, “next Wednesday”, “2016–03–01”, “123”, “one hundred”, etc. It may be helpful to convert them to unified format before doing pattern matching. Other entities which require special treatment: locations (countries, regions, cities, street addresses, places), people, phone numbers, …
Find concepts / synonyms	E.g., when searching for a breed of a dog, not all the dog breeds should be listed in the rule, as there are hundreds of them. It is a good practice if preprocessing code identified a dog breed in the message and marked the word with a special tag. Then it becomes

[60] https://www.nltk.org/
[61] https://nlp.stanford.edu/software/
[62] https://spacy.io/
[63] https://opensource.google.com/projects/syntaxnet
[64] https://wordnet.princeton.edu/

	easy to look for that tag when applying the rule. WordNet can be used to identify common concepts. There might be a need to add domain specific concept libraries, e.g. a list of drug names when building a healthcare bot.

After preprocessing is done, the result will be a list of clean sentences and words inside each sentence. Each word is marked with a part of speech and concepts, and a lemma for every word. The next step is to define a classification.

5.2.3. Classification

Any intent classification code can make errors of two types:

- False positives: the user does not express an intent, but the chatbot identifies an intent.
- False negatives: the user expresses an intent, but the chatbot does not find it.

There are different methods to find out intent classifications:

- Pattern matching[65]

  Bots utilize pattern matches to group text and produce an appropriate response from the clients. "Artificial Intelligence Markup Language" (AIML) is a standard structured model for these patterns. A simple example of Pattern matching is shown in Figure 27.

```
<aiml version="1.0"encoding="UTF-8"?>
        <category>
                <pattern> WHO INVENTED EMAIL</pattern>
                <template>according to google Ray Tomlinson invented email.</template>
        <category>


        <category>
                <pattern> DO YOU KNOW WHO * IS</pattern>
                <template>
                        <srai>WHO IS <stat/></srai>
                <template>
        </category>
</aiml>
```

Figure 27: Using AIML to match patterns.

The machine gives the following output:

Human: Who invented the email?

Robot: According to Google, Ray Tomlinson invented email.

The chatbot knows the appropriate answer because her or his name is in the related pattern. Similarly, the chatbots react to anything relating it

[65] https://bigdata-madesimple.com/how-do-chatbots-work-an-overview-of-the-architecture-of-a-chatbot/

to the correlate patterns. Nevertheless, it cannot go past the related pattern. To take it to a progressive stage, algorithms can help.

For each kind of question, a unique pattern must be available in the database to provide a suitable response. With lots of combinations in patterns, it creates a hierarchical structure. Algorithms can be used to reduce the classifiers and generate a more manageable structure. Computer scientists call it a “Reductionist” approach - in order to give a simplified solution, reduce the problem.

Multinational Naive Bayes is the classic algorithm for text classification and NLP. For instance, assume a set of sentences are given that belong to a particular class. With a new input sentence, each word is counted for its occurrence and is accounted for its commonality, and each class is assigned a score. The highest scoring class is most likely to be associated with the input sentence.[66]

- Natural Language Understanding (NLU)

  NLU has three specific concepts[67]: entities, intents and context.
  An **entity** represents a concept in the chatbots. E.g. a payment system in an e-commerce chatbot.

  Intents are the actions the chatbot should perform when the user writes something. Different user interactions may have the same intent. “I want to purchase a pair of red shoes” and “Do you have red shoes” will trigger the same command in a chatbot.

  When an NLU algorithm analyses a sentence, it does not have access to the history of the user conversation. It means that if it receives the answer to a question it has just asked, it will not remember the question. For differentiating the phases during the chat conversation, its **context** should be stored, e.g. “ordering pizza”. With the context, chatbots can easily relate intents without the need to know what the previous question was.

  Figure 28 provides a chatbot architecture example of how NLU could be used.

[66] https://www.marutitech.com/chatbots-work-guide-chatbot-architecture/

[67] https://bigdata-madesimple.com/how-do-chatbots-work-an-overview-of-the-architecture-of-a-chatbot/

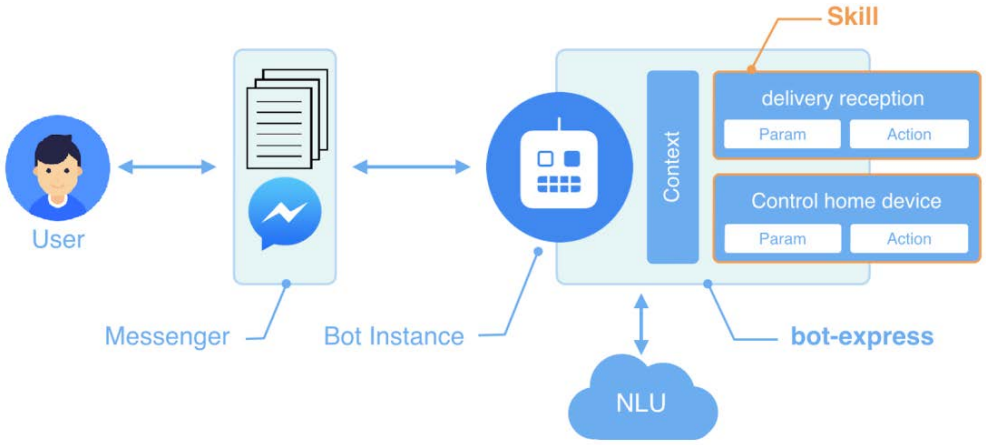


Figure 28: Use of NLU in chatbots architecture.[68]

- Natural Language Processing (NLP)

  The NLP component has the ability to derive meaning from complex input provided by the user. As a user has the ability to type any question that he wants, it is impossible to grasp everything in the logic of the chatbot. The NLP-component ensures that any input can be processed by the chatbot. The component processes the questions and derives "intents" or follow up questions from it (an example of a chatbot architecture using NLP can be seen in Figure 29). The NLP component is a learning component, meaning that it will be able to provide better answers and process complex questions the more it is used. Table 5 describes five commonly used NLP steps in chatbot technology.

Table 5: NLP steps.

NLP steps[69]	
Sentiment analysis[70]	Tries to learn if the user is having a good experience or if, after some point, the chat should be forwarded to a human.
Tokenisation	The NLP divides a string of words into pieces or tokens that are linguistically symbolic or differently useful for the application.
Named Entity Recognition	The chatbot program model looks for categories of words, like the name of the product, the user's name or address, whichever data is required. (See Figure 30 for an example of NER[71] using DialogFlow[72].)
Normalisation	The chatbot program model processes the text in an effort to find common spelling mistakes or typographical errors that might

[68] https://bigdata-madesimple.com/how-do-chatbots-work-an-overview-of-the-architecture-of-a-chatbot/
[69] https://www.marutitech.com/chatbots-work-guide-chatbot-architecture/
[70] More info on sentiment analysis can be found in section 7.1.1.
[71] Named Entity Recognition
[72] https://dialogflow.com/

	convey the user intent. This gives a more human like effect of the Chatbot to the users.
Dependency parsing	The chatbot looks for the objects and subjects, verbs, nouns and common phrases, in the user's text to find dependent and related phrases that the user might be trying to convey.

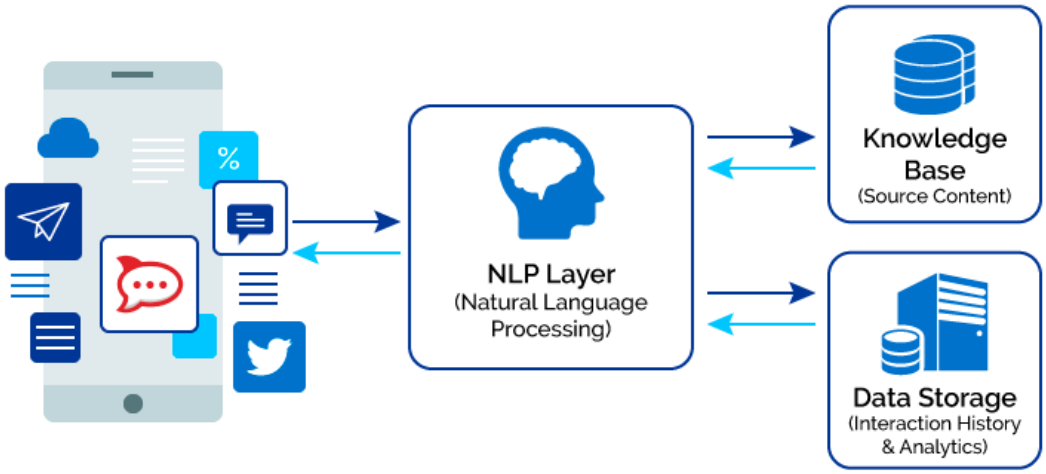


Figure 29: Handling chatbots messages with NLP.[73]

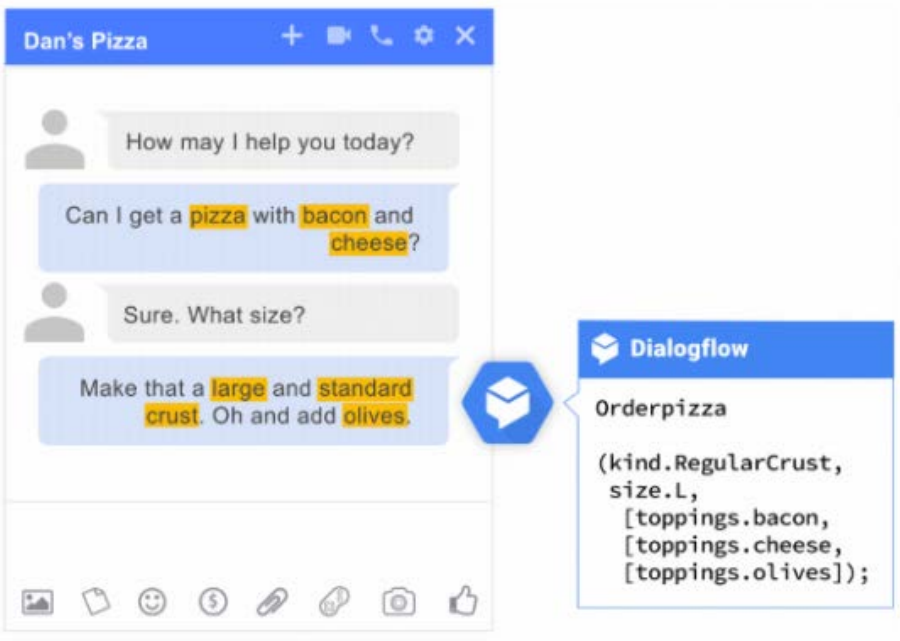


Figure 30: Example of the use of NLP with Dialogflow.

[73] https://bigdata-madesimple.com/how-do-chatbots-work-an-overview-of-the-architecture-of-a-chatbot/

- Natural Language Generation (NLG)

  Natural Language Generation is the process of generating text from a meaning representation. It can be taken as the reverse of the NLU. NLG systems provide a critical role for text summarization, machine translation and dialog systems. In NLG, the system responds as a semantic frame; it maps back to a natural language sentence, understandable for the end user.

  Machine learning based (trainable) NLG systems are common in today's dialog systems. Such NLG systems use several sources as input such as: content plan, representing meaning representation of what to communicate with the user, knowledge base, structured database to return domain-specific entities, user model, a model that imposes constraints on output utterance, dialog history, the information from previous turns to avoid repetitions, referring expressions, etc.

  Trainable NLG systems can produce various candidate utterances (e.g., scholastically or rule based) and use a statistical model to rank them. The statistical model assigns scores to each utterance and is learnt based on textual data.

  On the other hand, if NLG is based on a semantically controlled Long Short-term Memory (LSTM) recurrent network, it can learn from unaligned data.

5.2.4. Response selection

Once the chatbot understands the user's message, the next step is to generate a response. One way is to generate a simple static response. Another way is to get a template based on intent and put in some variables. The chatbot development company chooses the method for generating the response depending on the **purpose** for which the chatbot is employed.[74] The following diagram (Figure 31) is the representation of separate response generation and response selection modules:

[74] https://dzone.com/articles/understanding-architecture-models-of-chatbot-and-r

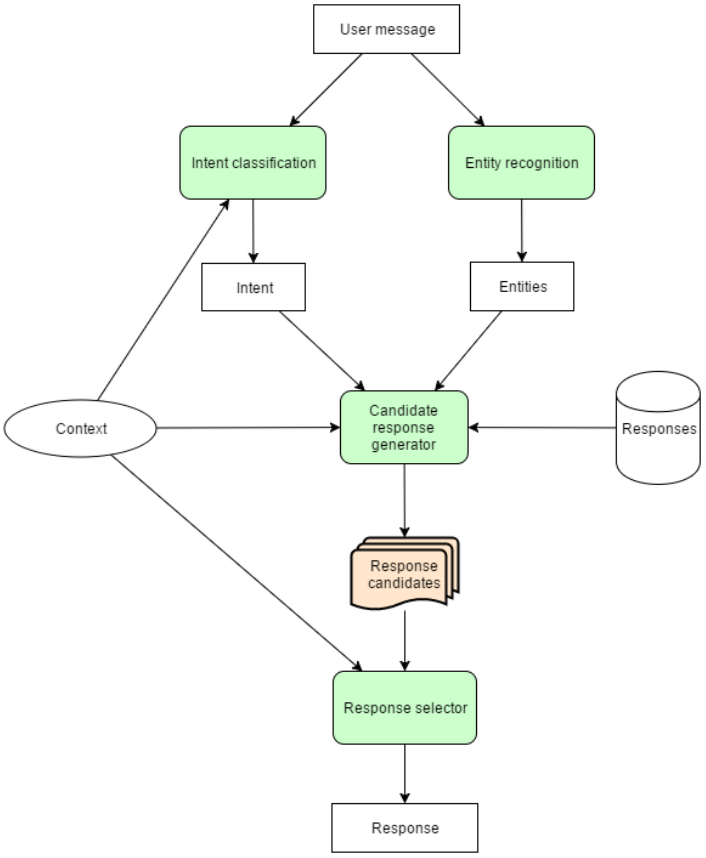


Figure 31: Separate response generator and response selector modules, using NLU.[75]

The candidate response generator is doing all the domain-specific calculations to process the user request. It can use different algorithms, call a few external APIs, or even ask a human to help with response generation. The result of these calculations is a list of response candidates. All these responses should be correct according to domain-specific logic; it cannot be just tons of random responses. The response generator must use the context of the conversation as well as intent and entities extracted from the last user message, otherwise, it cannot support multi-message conversations.

The response selector scores all the response candidates and selects a response, which would work best for the user. The chatbot can express the same message using different words. A weather bot can say, "It's rainy", or "Probability of rain is 80%" or "Please carry an umbrella today". Which one will work the best for the user? Different users prefer different styles of response; these styles are described in section 6.1.1.

The bot can analyse previous chats and associated metrics[76] (length of the conversation, probability of sale, rating of customer satisfaction, etc.) to tailor

[75] https://dzone.com/articles/understanding-architecture-models-of-chatbot-and-r
[76] See section 7.2 for an overview of chatbot metrics.

responses for the user.[77] This document provides three types of responses in Table 6.

Table 6: Response types.

Response types[78]	
Static responses	Here the bot selects from a predetermined range of answers. For instance, “you have these many items left in your cart.”
Dynamic responses	The bot may use a knowledge base to score a right answer. This is appropriate mostly when the bot will function as an FAQ system.
Generated responses	The bot uses deep learning algorithms to curate better answers instead of canned responses. It will need to employ more time and data for maturity.

Public service chatbots could make use of the CPSV-AP[79] data model to optimise their responses. Figure 32 and Figure 33 show a chatbot pilot using data provided in the CPSV-AP model.

[77] https://medium.com/@surmenok/chatbot-architecture-496f5bf820ed

[78] https://www.newgenapps.com/blog/complete-guide-on-chatbot-architecture
Section 6 provides public administration chatbot examples of responses.

[79] https://joinup.ec.europa.eu/solution/core-public-service-vocabulary-application-profile/about

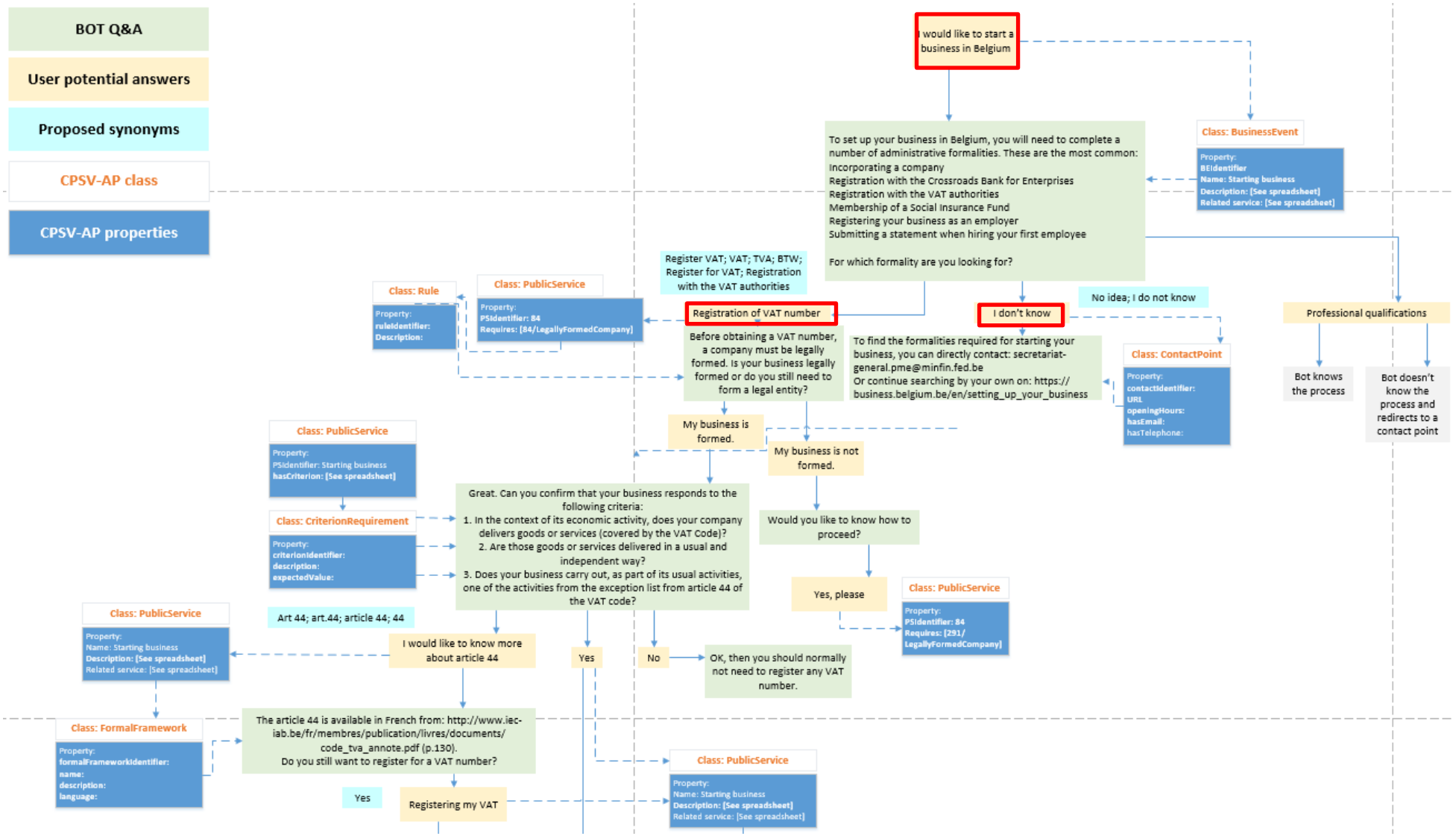


Figure 32: Extract I of the chatbot pilot using the CPSV-AP data model.

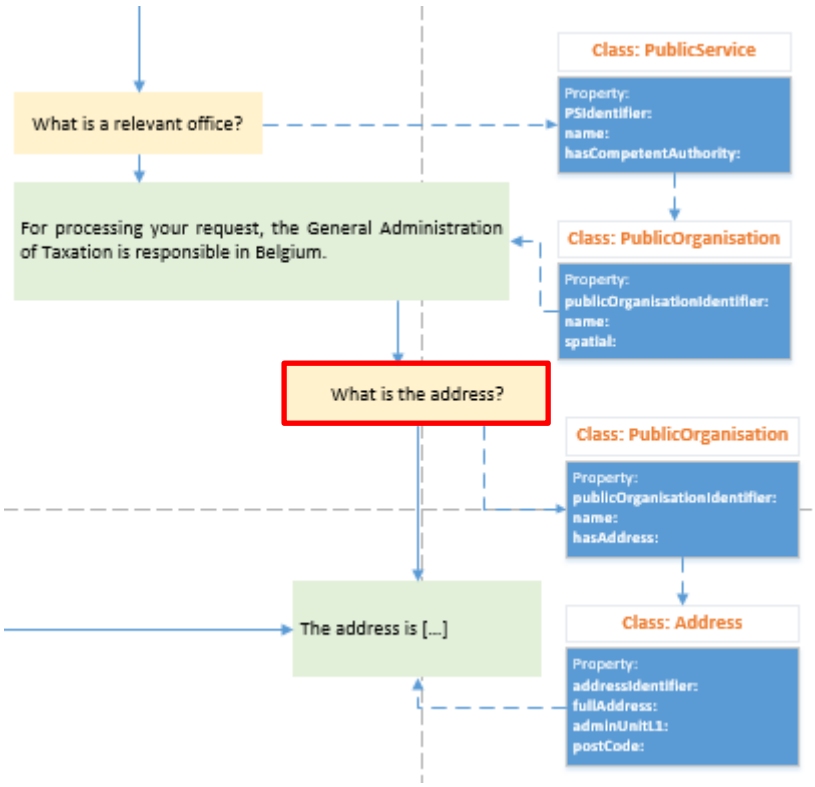


Figure 33: Extract II of the chatbot pilot using the CPSV-AP model.[80]

Looking closer at Figure 32 and Figure 33, multiple use cases of a chatbot using CPSV-AP are distinguished. This document provides a few concrete examples on how this works in Table 7.

Table 7: Examples of a chatbot pilot using the CPSV-AP data model.

User input	CPSV-AP class[81]	Class elements used
"I would like to start a business in Belgium"	This user input was linked to CPSV-AP Class:BusinessEvent through the list of class names (here: 'starting a business').	A list was provided to the user containing the most common **descriptions** of public services linked to this CPSV-AP class name.
"I don't know"	This user is redirected (through the identifier) to the relevant information found in Class:ContactPoint.	In this example, the URL and email (hasEmail) were provided.
"Registration of VAT number"	The chatbot uses the identifier to look up the public service in Class:PublicService, where a	Identifier to link user input, public services and rules. Description of the rule.

[80] Pilot model and additional info to be found on Confluence (limited access).

[81] The CPSV-AP specifications can be found on Joinup: https://joinup.ec.europa.eu/release/core-public-service-vocabulary-application-profile/221

	connected rule is found, linked to the public service. Class:Rule provides this rule.	
"What is a relevant office"	The chatbot links this question again to a public service and then to a specific public organization through Class:PublicOrganisation.	Class:PublicService and Class:PublicOrganization are linked through the element 'has competent authority'.

5.3. What kind of data will the chatbot rely on?

The chatbot relies on data about the user needs and data that is needed for the training of the chatbot. This depends a lot on the chosen software and platform. The Finnish solution for example, is based on training data (example questions which users would ask) and test data which is used to test the model's performance.

5.3.1. Storing intents as structured data

The database behind the chatbot stores intents to be matched against answers, frequently asked questions and of course the information that should be provided to the user of the chatbot. What specific sort of database to use depends on the purpose of the chatbot. Structured data is typically stored in databases that conform to a data model or schema.

Most relational databases use SQL (Standard Query Language). The chat patterns modelled in the pattern-template are stored in a relational database management system (RDBMS), e.g. MySQL[82].

5.3.2. Static Data vs Real time data

Most chatbots make use of static data, but it is possible to use real-time data as well. Figure 34 shows an example of a chatbot providing the real-time bus arrivals at a bus stop. Some platforms provide real time data streaming to facilitate this kind of services.[83]

[82] https://www.mysql.com/

[83] E.g. https://sendbird.com/ & https://www.pubnub.com/

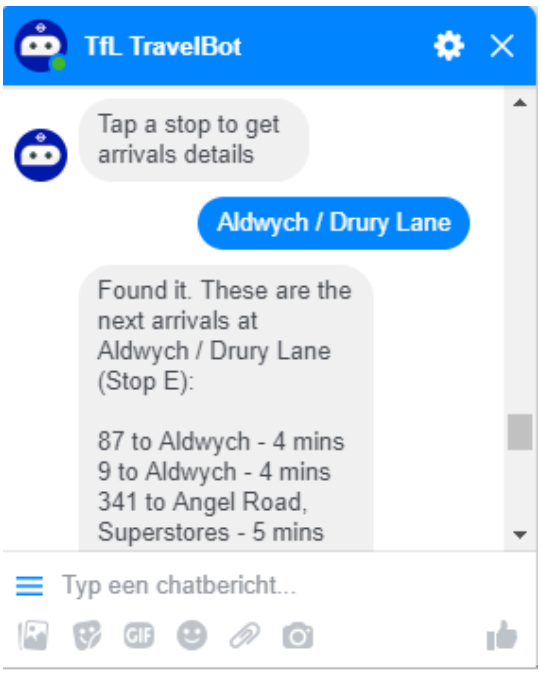


Figure 34: Chatbot Tfl, Transport for London - TravelBot provides the next arrivals at a bus stop.[84]

5.3.3. Unstructured data

Many processes are involved in preprocessing and processing user input data, which is typically unstructured. To handle this, the NLP engines call APIs to process data in different ways.

5.3.4. History

Conversational chatbots might need to store the conversation, which could be useful:

- For users, to recall the last conversation and find back what the chatbot provided; and
- For the chatbot, to keep track of user needs and improve the service provided.

5.4. Chatbot security

Just like any new technology, chatbots come with cyber-security risks. For this reason, it is very important to know the security practices that are commonly implemented when working with chatbots. For the most part, chatbots do not present security issues that have not already been discovered and properly mitigated. The next subsections provide chatbot security concerns and mitigating strategies.

5.4.1. Use of public chat platforms

Public platforms such as Facebook and Slack have become a popular medium for deploying chatbots. However, there are significant privacy concerns pertaining to the

[84] https://www.facebook.com/tfltravelbot/

data handled by such platforms and, at the moment, it is prudent to restrict these bots to informational services and redirect the user to a more secure interface when the input of sensitive details such as account numbers are required. Sensitive details should be sent only over platforms that support end-to-end encryption.

5.4.2. Encryption

Secure messaging protocols such as HTTPS should be implemented while transferring data. Further, end-to-end encryption ensures that only the two parties involved in the conversation have access to the contents of the conversation. Lack of encryption may leave information susceptible to eavesdropping from an attacker launching a Man-in-the-Middle (MitM). Often, the chat service provider itself may collect data being sent via chat services. Encryption must be implemented for data in transit as well as data at rest. Encryption must be implemented using secure encryption algorithms while maintaining perfect forward secrecy i.e. the compromise of one set of keys does not enable an attacker to decrypt all previous conversations.

5.4.3. Authentication

Multi-factor authentication must be implemented before any sensitive information is handled i.e. if the bot operates on Facebook messenger, the user's authorization through their Facebook login is insufficient and another layer of authentication such as the use of eID should be enforced through the chatbot. Upon successful authentication, secure authentication tokens may be generated which maintain the user's session for a fixed period, after which the authentication should time out and the user must be prompted to reauthenticate himself. Sufficient authorization controls must also be put in place before returning a user's data, to ensure that a user cannot retrieve data pertaining to another user.

5.4.4. Integrity check

The contents of the message can be verified at each end of the conversation to ensure there is no tampering/corruption of data. For example: Facebook implements the X-Hub-Signature header as part of each request, which is a SHA-1 signature of the request payload. The bot can then re-compute the header value and match it with the one sent in the request header.

5.4.5. Social Engineering attacks

Bots are capable of impersonating humans, and vice versa. Customers may be tricked by malicious bots impersonating a company/business and asking for sensitive information. The key to combat these attacks is increased user awareness initiatives. Mechanisms should be in place to verify the identity of the application that the user is communicating with.

5.4.6. Infrastructure security

The bot can be compromised when attackers take advantage of vulnerabilities in the technologies used. Hence, it is imperative to ensure secure infrastructure with the latest patches put in place.

5.4.7. Data storage and handling

The more data a chatbot collects, the greater the chatbot owner's liability in case of a breach. It is important to ensure that data collected is being stored securely and handled with proper care. Additionally, the application must only store data when necessary. A data retention period must be in place beyond which copies of the data must be destroyed. The data must be stored in an encrypted format and access to this data is to be carefully monitored. Further, if this data is being used for analytics and research purposes, the user's identities must be protected through anonymisation/pseudonymisation.

5.4.8. Self-Destruction of messages

Messages sent and received via the chatbot should automatically disappear after a stipulated time period.

5.4.9. Application security

The chatbot application must handle data input and output in a secure manner. A fundamental tenet of secure coding principles is to "never trust user controlled input". Insecure handling of user input can result in vulnerabilities such as code injection and client side scripting attacks.

5.4.10. Pollution in the communications channel

Self-learning chatbots, which utilize machine learning to improve accuracy of responses over time, can be trained to generate inappropriate/inaccurate responses. These chatbots usually parse assertions made in the communication channel. If enough users make the same statement, it assumes it is true. Chatbots should be programmed to detect and disregard anomalous behaviour of this kind.

5.4.11. Use of external NLP services

If a chatbot platform relies on external libraries or services to analyse the user text, e.g., extract a date or a phone number, then this communication must be secured. A reasonable approach is to never send any personally identifiable information (e.g., name or address) or any session information to such services, so they cannot associate messages with users. Another approach is to only use libraries inside the Conversation Management Engine, which does not communicate with outside entities.

5.4.12. Logging and access rights

The chatbot platform should log all actions. The chatbot platform follows strict policies on who is granted access to what data; in general employees should only be given access to sensitive data if they sign the right forms and have a real need to access this data. Logs should be encrypted and stored in a secure location.

A more detailed overview of these security concerns (a security testing guide) can be found in Annex III.

6. THINKING ABOUT USER INTERACTIONS

A great benefit of chatbots is the wide range of possibilities to interact with users. Depending on the audience and purpose of the chatbot, it is possible for public administrations to select multiple kinds of interaction possibilities.

6.1. Introduction

6.1.1. Chatbot persona

A chatbot could be neutral or have a personality. Different reasons exist to add a persona to a chatbot, e.g. adding a human face to the interaction could increase familiarity for the user. Researched public administration chatbots had five different categories of avatars:

- Human cartoon;
- Robot cartoon;
- Human picture (e.g. Figure 35);
- Shape; and
- Shape gif (e.g. moving circles).

Figure 35: Chatbot Jamie, IRAS[85] - Introduction.[86]

The chatbot's personality differs not only in a visual way, but also content-wise (e.g. a sarcastic chatbot, a chatbot using emoticons, etc.). The personality of a chatbot is present in every stage of the conversation. Chatbot Kamu opens friendly and professional, as it provides information on immigration and citizenship (Figure 36).

[85] Inland Revenue Authority of Singapore

[86] https://www.iras.gov.sg/irashome/default.aspx

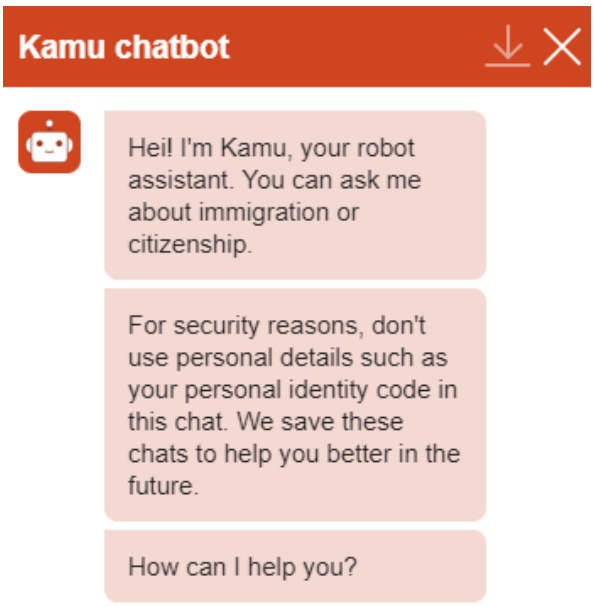


Figure 36: Chatbot Kamu, Finnish Immigration Service – Introduction

Architecture components: This information could be stored in a database and retrieved by means of an API.

6.1.2. Introducing itself as a bot

Some public administration chatbots welcome users and explain that they are chatbots (e.g. Figure 37). As the difference is sometimes not clear if one is talking to a chatbot or a human, this could be a useful functionality.

Figure 37: Chatbot Bonn, City of Bonn – Welcoming the user. [87]

[87] https://govbot.bonn.de/

Architecture components: This static information could be stored in a database and retrieved by means of an API.

6.1.3. Welcoming users in multiple languages

Multiple chatbots have multilingual functionality. Chatbot Emma (see Figure 38) starts the conversation with giving two language options.

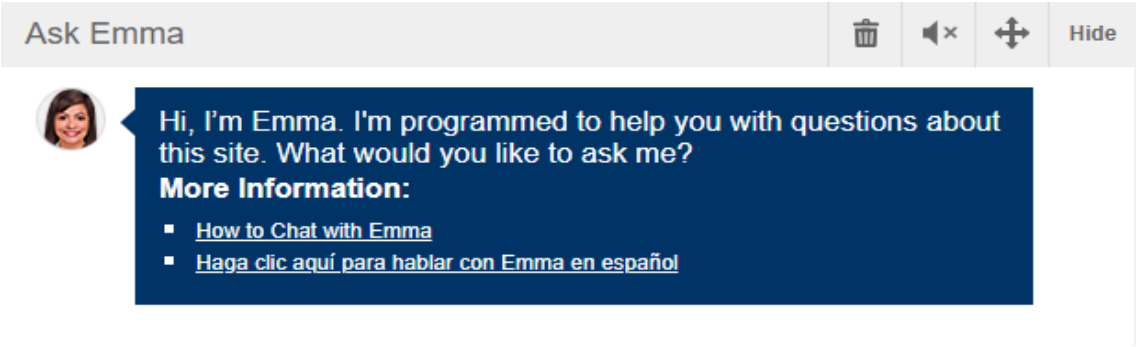


Figure 38: Chatbot Emma, U.S. Citizenship and Immigration Services - Opening sentence.[88]

Other public administration chatbots tackle this functionality in other ways.

- Chatbot Rammas (Dubai Electricity & Water Authority) does not have the functionality to switch languages, but it is possible to pick a language on the site where the chatbot is embedded.
- Chatbot Chip gives the user the option to write in any language (61 supported languages) and has language recognition software (see Figure 39 and Figure 40).
- Chatbot Bonn's (Figure 41) language options are available through a menu.
- Chatbot Noa (Figure 42) provides the language choice directly through the interface.
- Berlin's Chatbot (Figure 43) provides an explanation on how the translations work.

[88] https://www.uscis.gov/emma

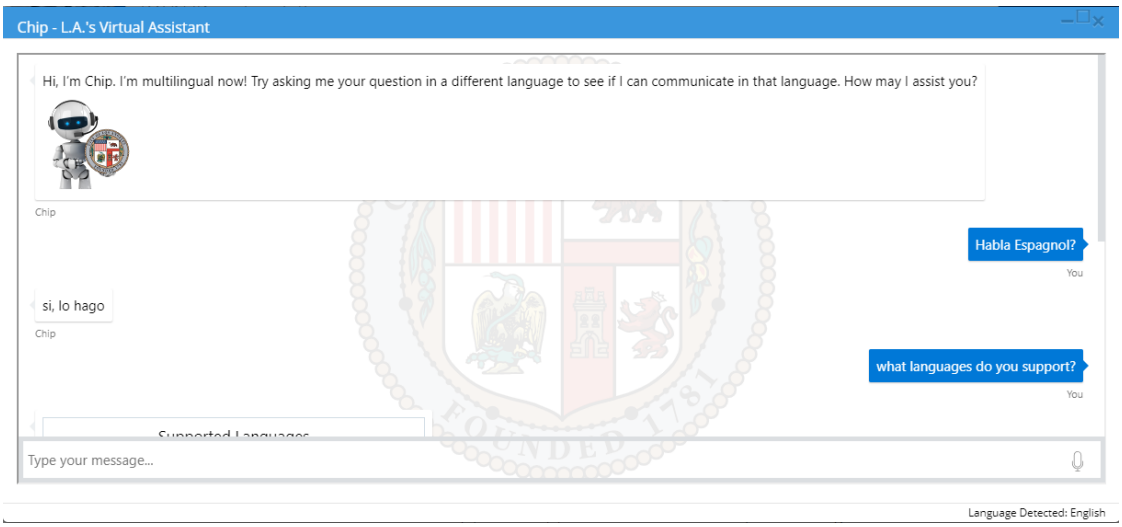


Figure 39: Chatbot Chip, Los Angeles City - Opening sentences.[89]

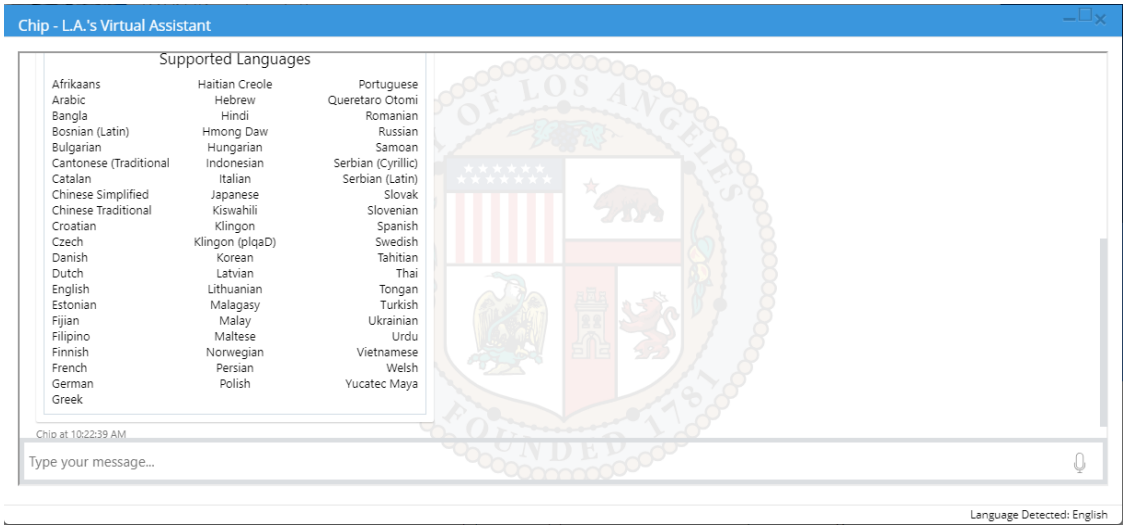


Figure 40: Chatbot Chip, Los Angeles City - Supported languages.

[89] https://www.labavn.org/#

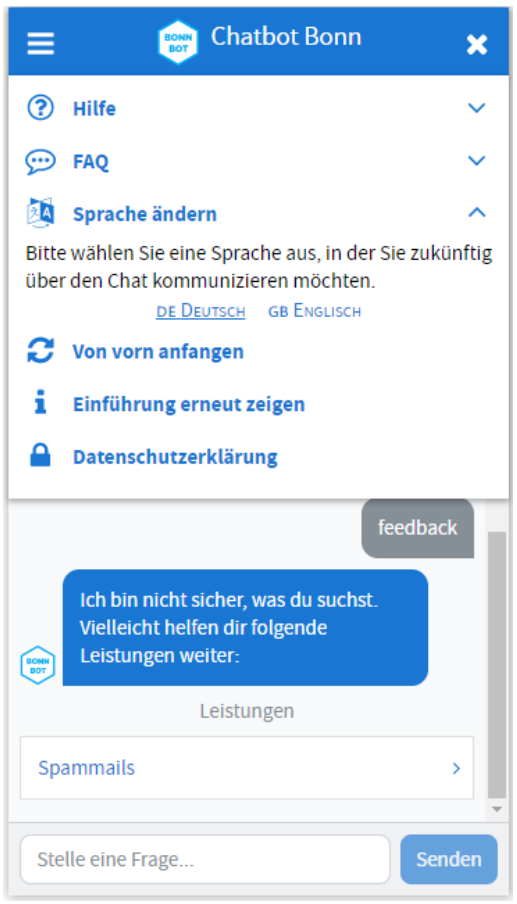


Figure 41: Chatbot Bonn – Changing language via a menu.

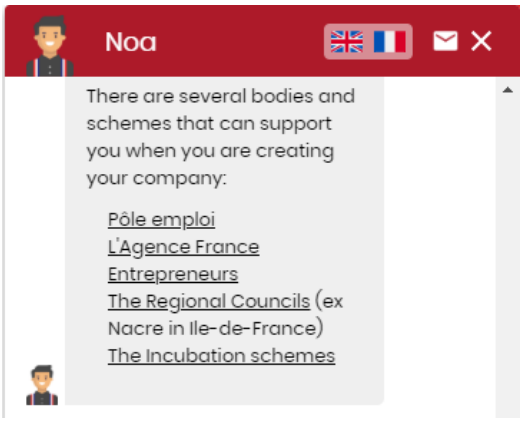


Figure 42: Chatbot Noa – Choosing between two languages directly in the interface.

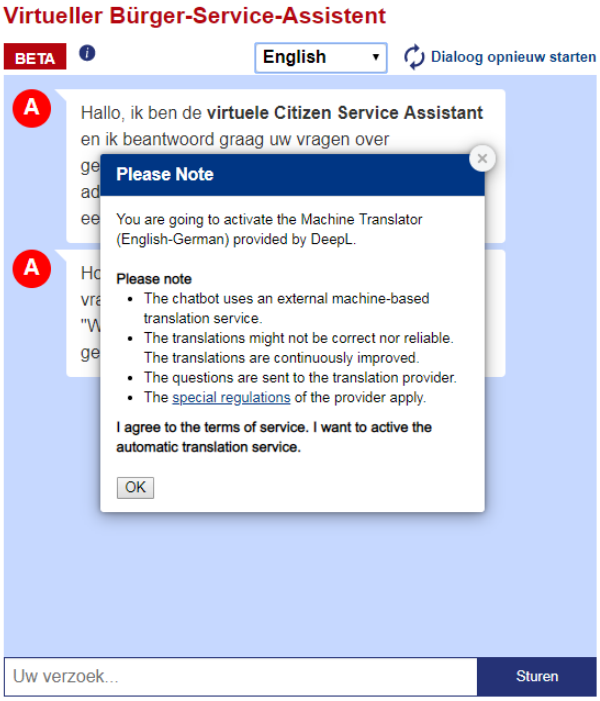


Figure 43: Virtual assistant, Berlin – Explaining how the translation service works.[90]

Architecture components: The multilanguage support could be provided by means of API coupling: one for the original content retrieval and one for the pure translation. The first could be connected to an internal database, the second a third party service such as Google Translate or the CEF eTranslation service[91].

6.1.4. Including privacy policy details in introduction

Some existing public administration chatbots introduce themselves and add information on privacy details such as the use of personal data (see Figure 44) and the storage of conversations (see Figure 44 and Figure 45).[92]

[90] https://service.berlin.de/virtueller-assistent/virtueller-assistent-606279.php

[91] https://ec.europa.eu/cefdigital/wiki/display/CEFDIGITAL/eTranslation

[92] This can be linked to the class Legal Resource in CPSV-AP. The CPSV-AP specifications can be found on Joinup: https://joinup.ec.europa.eu/solution/core-public-service-vocabulary-application-profile/releases

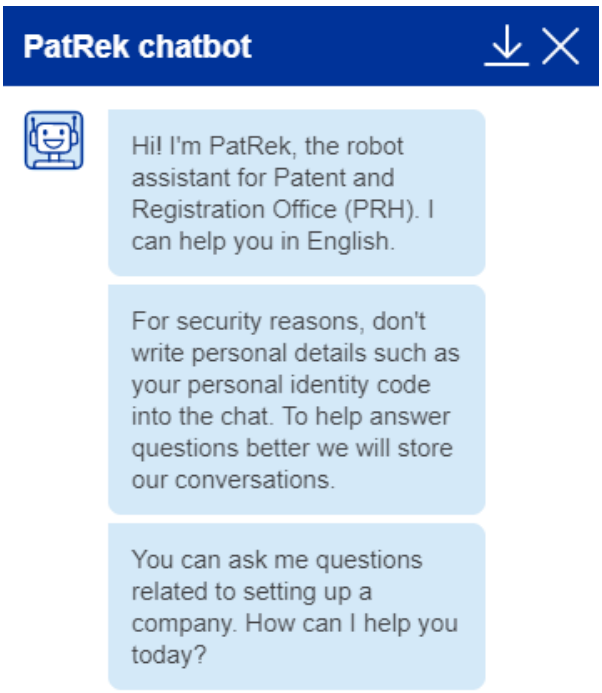


Figure 44: Chatbot PatRek, Finnish Patent and Registration Office – Explaining the privacy policy.

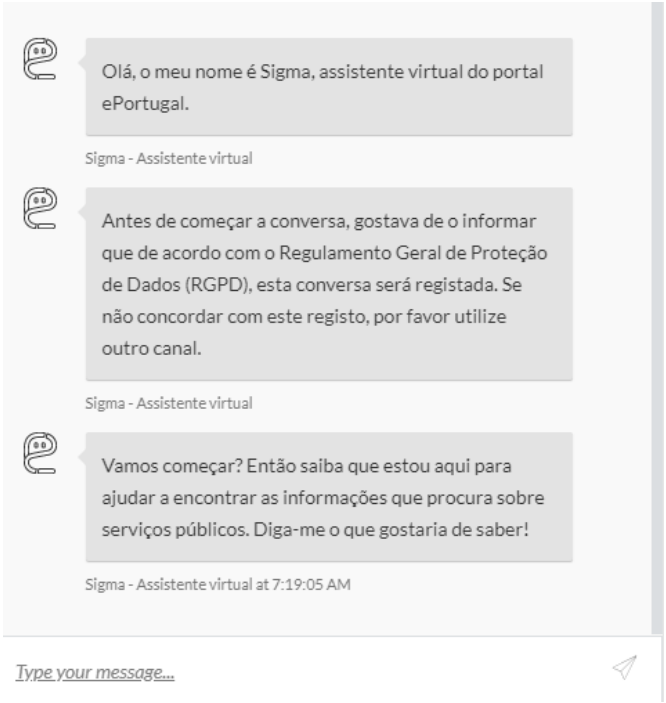


Figure 45: Chatbot Sigma, ePortugal – Introducing itself.[93]

Architecture components: This static data could be stored in a database and retrieved by means of an API or linked to a website where data privacy is discussed.

[93] https://eportugal.gov.pt/

6.2. Basic interaction

6.2.1. Providing different channels

Chatbots can use different channels to provide information to the user. Most public administration chatbots - researched in this study - provide the classical text-to-text interaction, but as seen in Figure 46, public administration chatbots exist that provide voice-to-text and text-to-voice assistance.

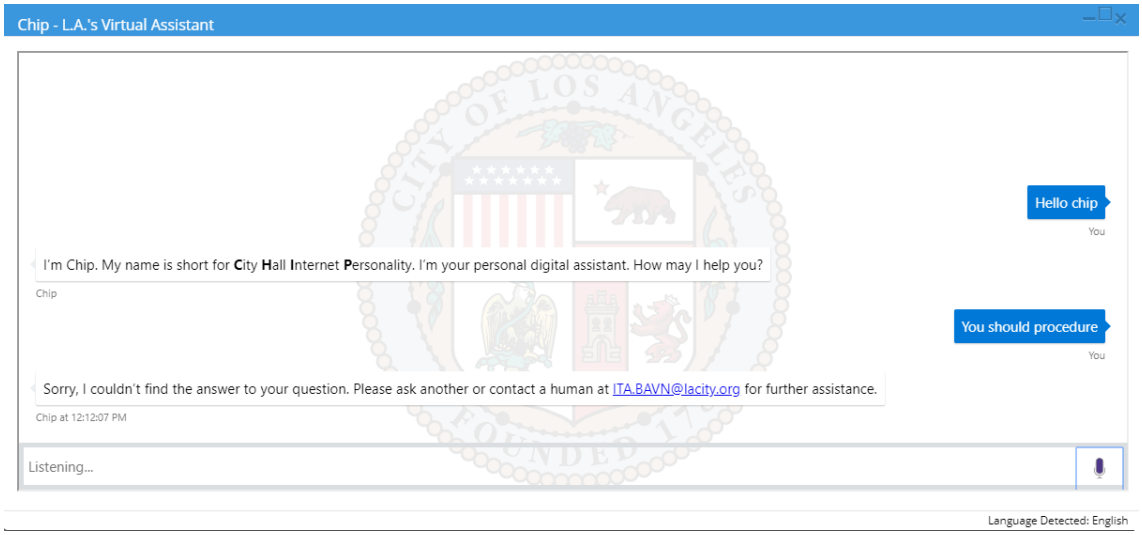


Figure 46: Chatbot Chip[94], City of Los Angeles - Using the voice-to-text and text-to-voice functionality.

Architecture components: Third party libraries or APIs could support audio encoding.

6.2.2. Menu-based / quick reply chatbots

Chatbots offer different ways of providing options to choose from to the user (see Figure 47). Some chatbots solely use options and do not respond in sentences. Most chatbots offer multiple options after each question, but can still reply without choosing a specific option (so with unstructured data). Some chatbots ask questions back before giving options.

94 https://www.labavn.org/#

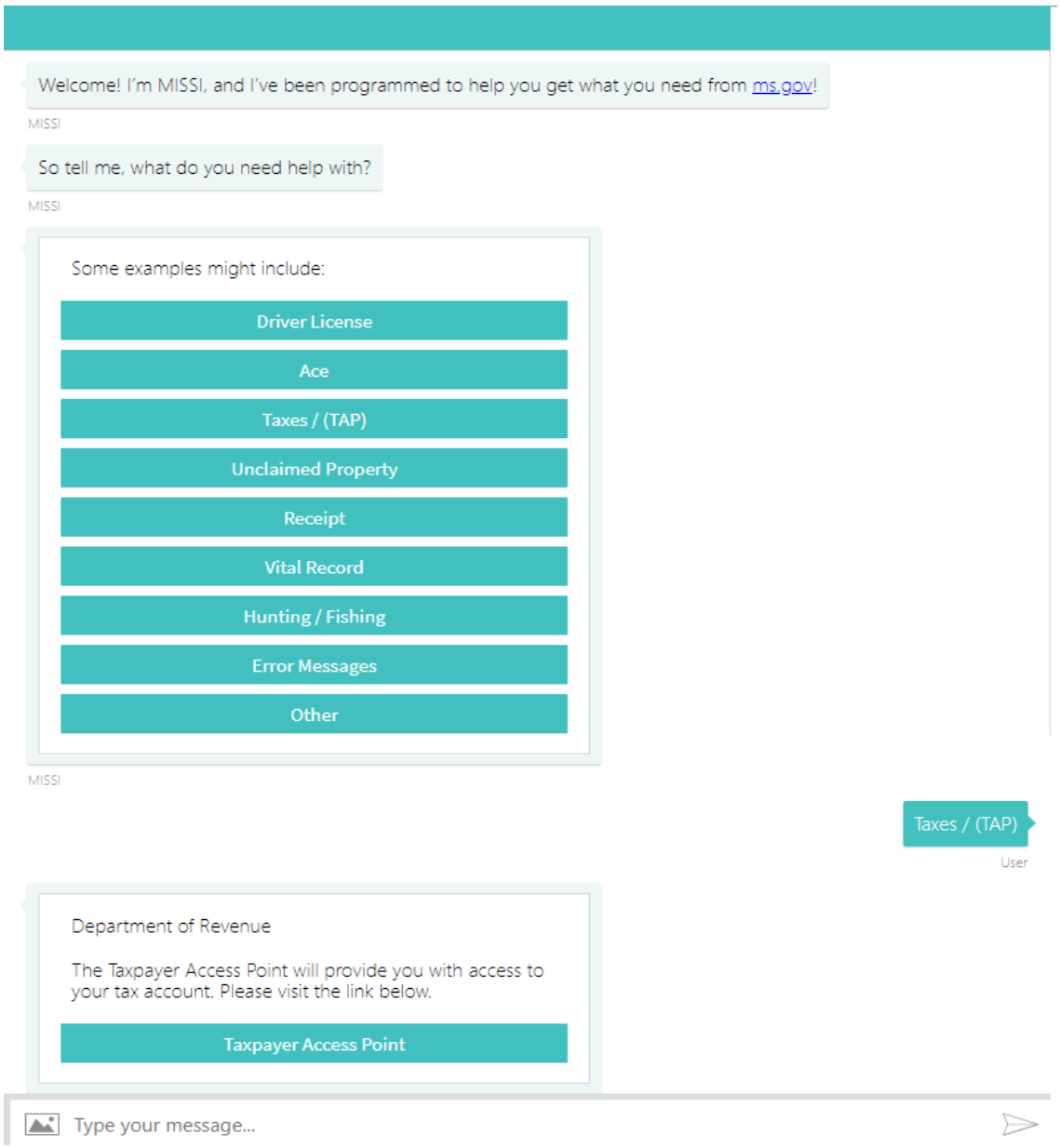


Figure 47: Chatbot MISSI - Opening the chat by presenting a list of possible topics to choose from. [95]

Architecture components: This information could be stored in a database.

6.2.3. Providing a selection menu where the user can confirm multiple answers at once

Some public administration chatbots have a functionality where the user can confirm multiple answers at once. This is especially interesting in public health chatbots as (mental) health issues can have multiple causes and outcomes (e.g. Figure 48).

[95] https://www.ms.gov/Technology

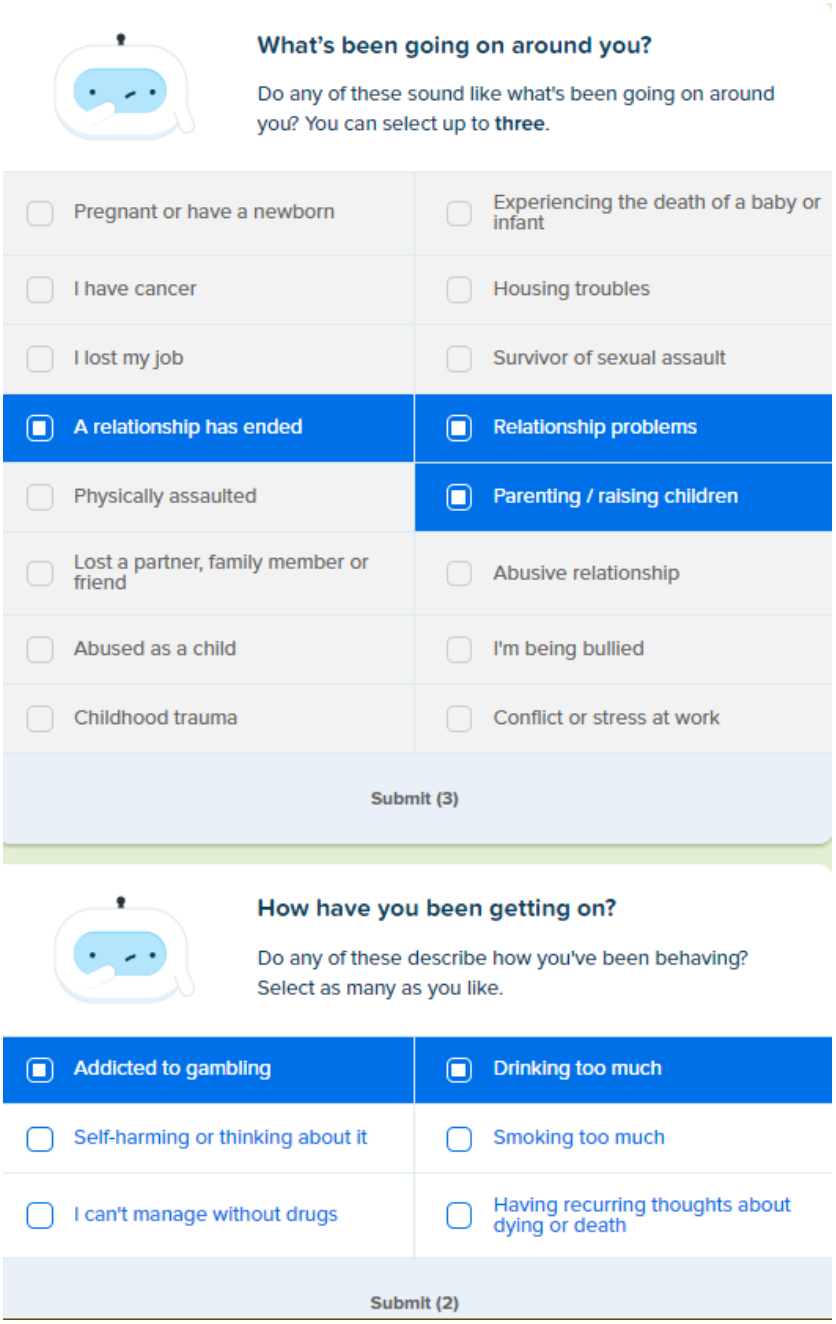


Figure 48: Chatbot Sam, Australian Government Department of Health – Filtering options.[96]

Architecture components: This information could be stored in a database.

6.2.4. Providing concise answers and links

Some public administrations respond to user queries by providing concise answers and links, e.g. chatbot Alex (Figure 49).

[96] https://headtohealth.gov.au/sam-the-chatbot

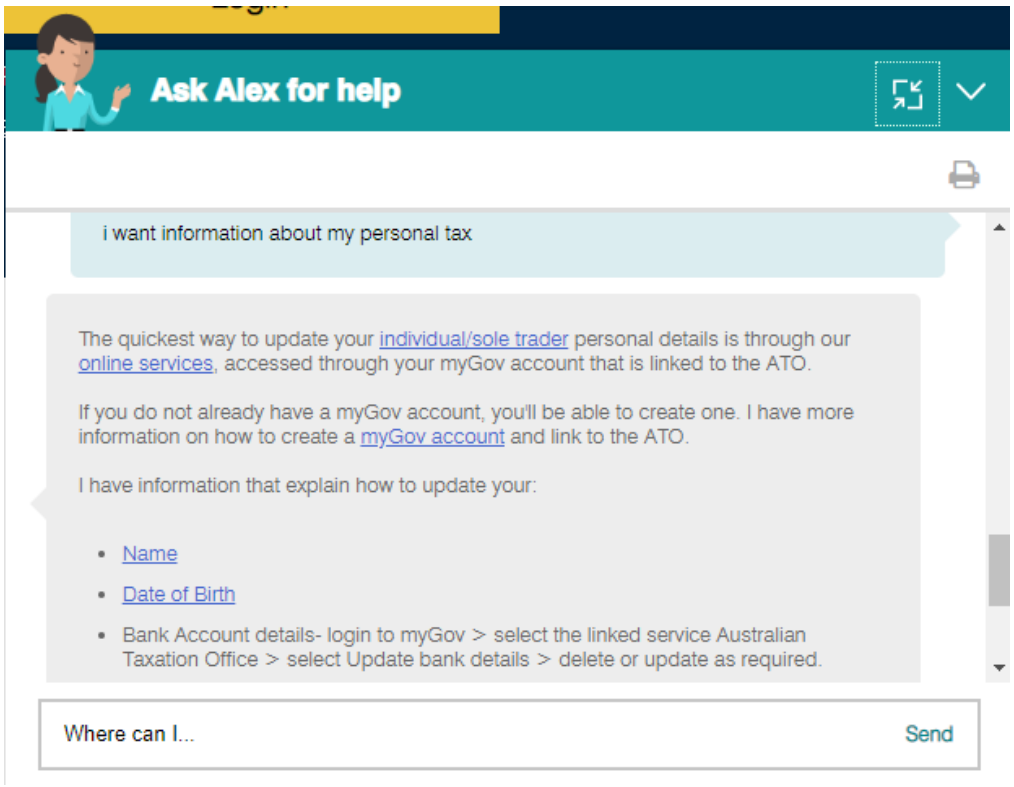


Figure 49: Chatbot Alex, Australian Taxation Office - Answering a question.

Architecture components: This information could be stored in a database or retrieved from web pages.

6.2.5. Requesting input from the user

Government chatbots request input from users ranging from the user's driving license number to the location of the user. In the Tfl TravelBot (see Figure 50), questions were asked to approximately locate the user to deliver the best transport offerings.

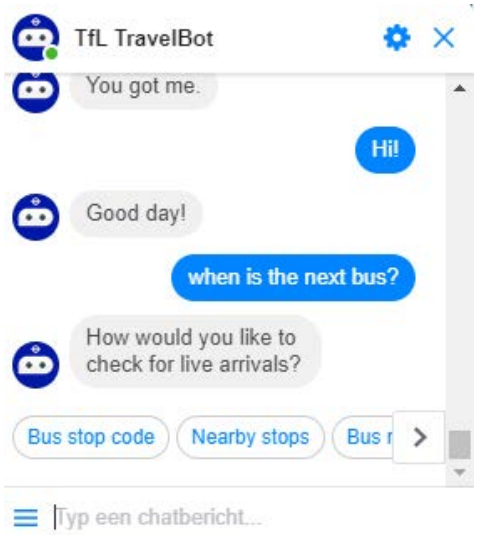


Figure 50: Chatbot Tfl, Transport for London - Presenting a list of options after a question is asked.

Architecture components: Using a location API and a real-time API to check the closest public transport in order to minimize the time needed for the transport.

6.3. Redirection

6.3.1. Redirecting to a human if the conversation cannot be managed

Multiple public service chatbots have the functionality to redirect the user if that chatbot cannot find an answer to a user query. Next figures (Figure 51, Figure 52 and Figure 53) show different ways on redirecting: through a video chat, phone, email and chat.

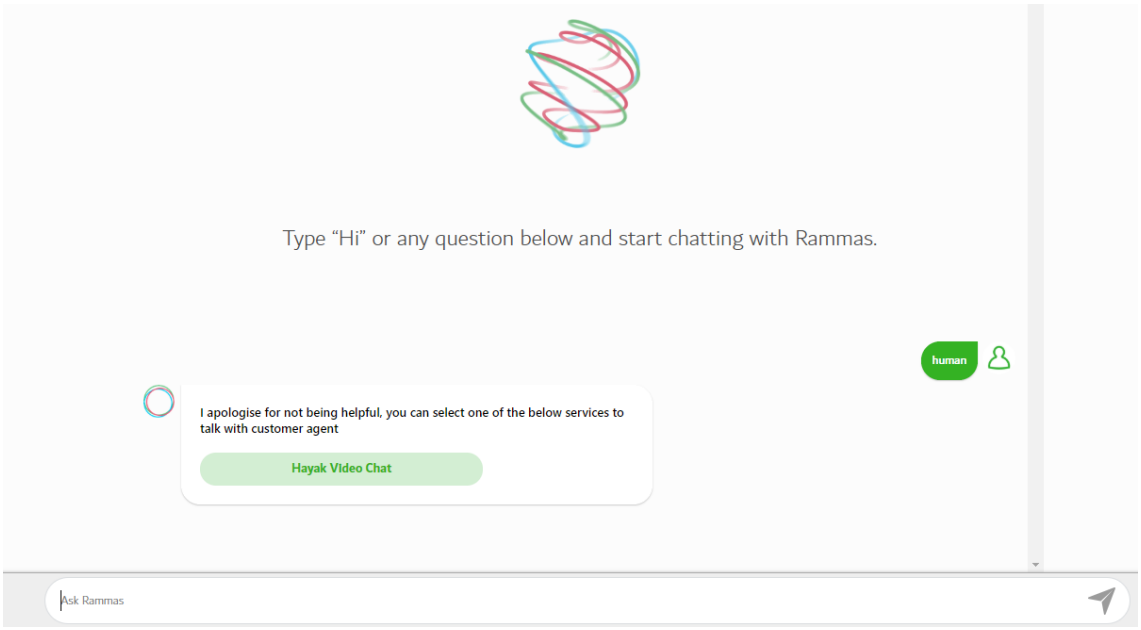


Figure 51: Chatbot Rammas, Dubai Electricity & Water Authority – Redirection to a human operator.

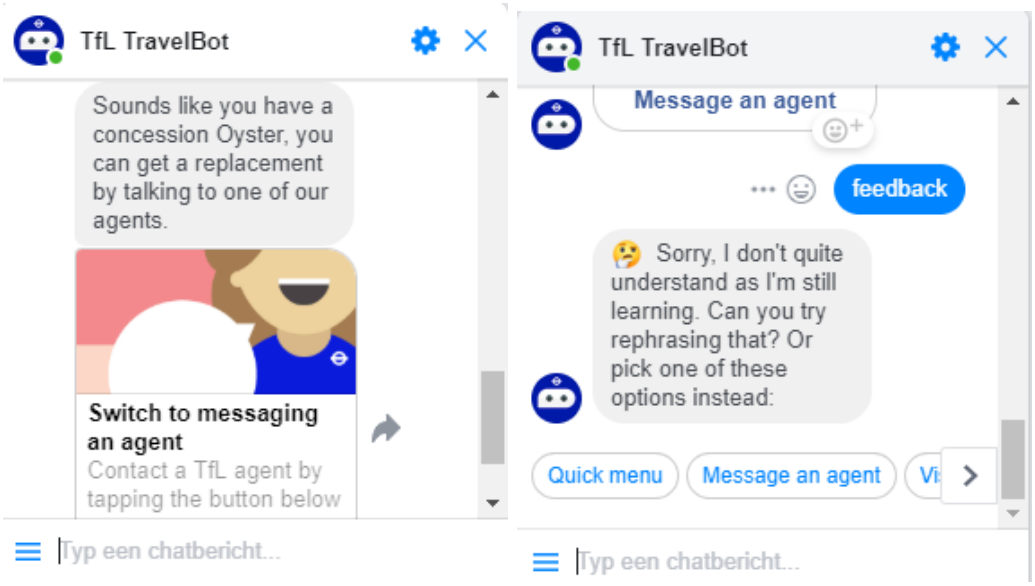


Figure 52: Chatbot TfL TravelBot – Redirection through another Facebook Messenger chat window.

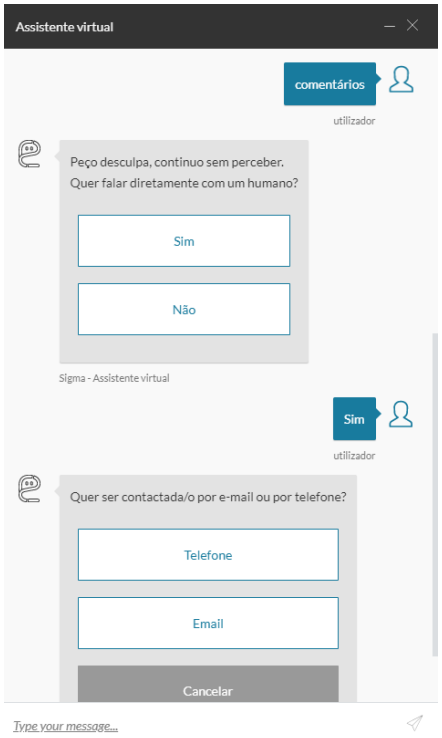


Figure 53: Chatbot Sigma – Redirection to a human through phone or mail.

Architecture components: A messaging API could connect chatbots, chat systems or a mail delivery system.

6.3.2. Redirecting to another chatbot

One researched chatbot has the functionality to transfer the user to another chatbot if the user's query is linked to a topic covered by that other chatbot (Figure 54). If a chatbot would be a public service provider, a 'related public service' of CPSV-AP could link to the next chatbot.

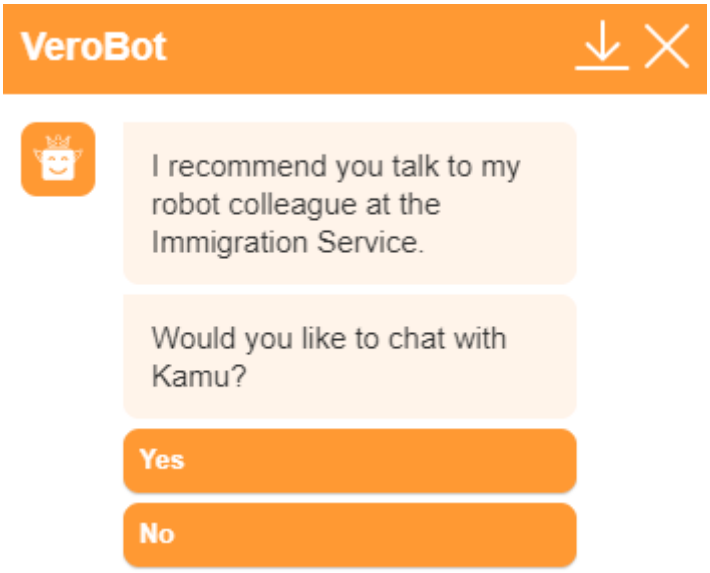


Figure 54: Chatbot VeroBot[97], Finnish Tax Administration – Redirection to another chatbot.

Architecture components: A chatbot API could be used, that can transfer context between the chatbots.

6.4. Other functionalities

6.4.1. Filtering

Public chatbots can have a filtering function. Chatbot Sam provides multiple outcomes and gives the user the opportunity to filter in three ways: format, age group and population (see Figure 55).

[97] https://www.vero.fi/en/individuals/

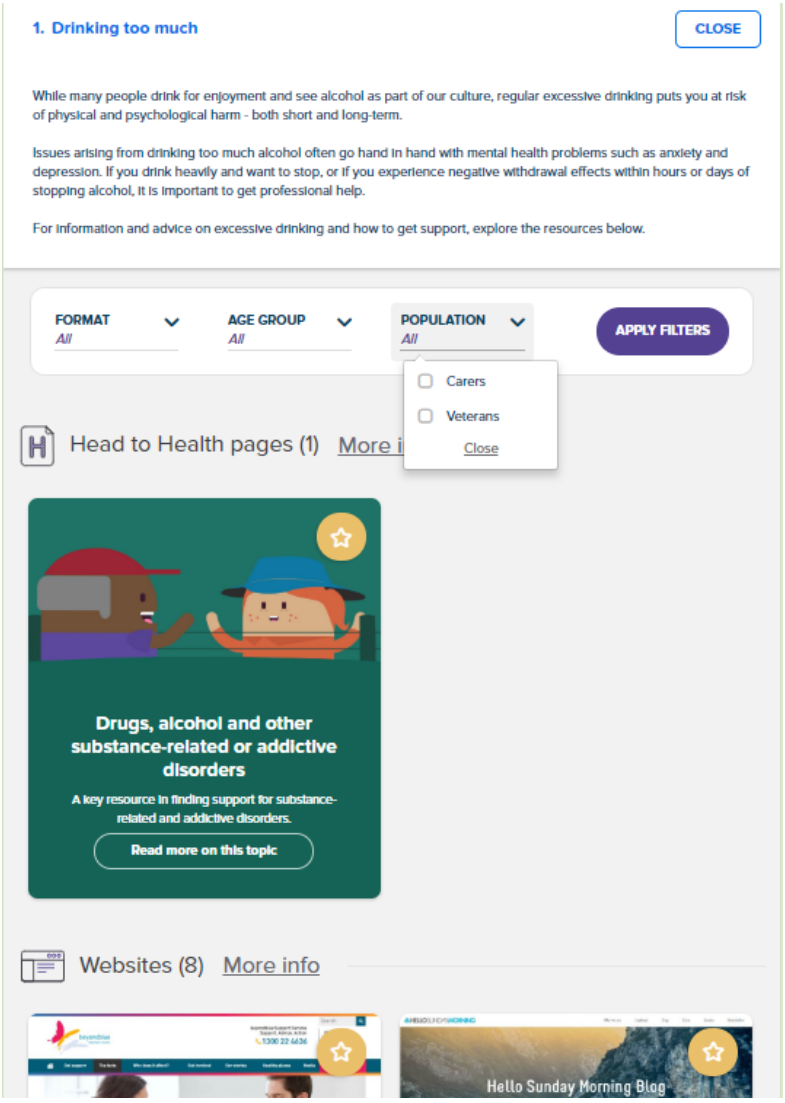


Figure 55: Chatbot Sam, Australian Government Department of Health – Filtering options.

Architecture components: Content can be retrieved from a database and be displayed at an application level.

6.4.2. Bookmarking

Another interesting functionality of public administration chatbots is the option to bookmark discovered resources. Especially when a chatbot provides multiple sources as an answer, this is a useful functionality for the user to find back the preferred answer (see Figure 56).

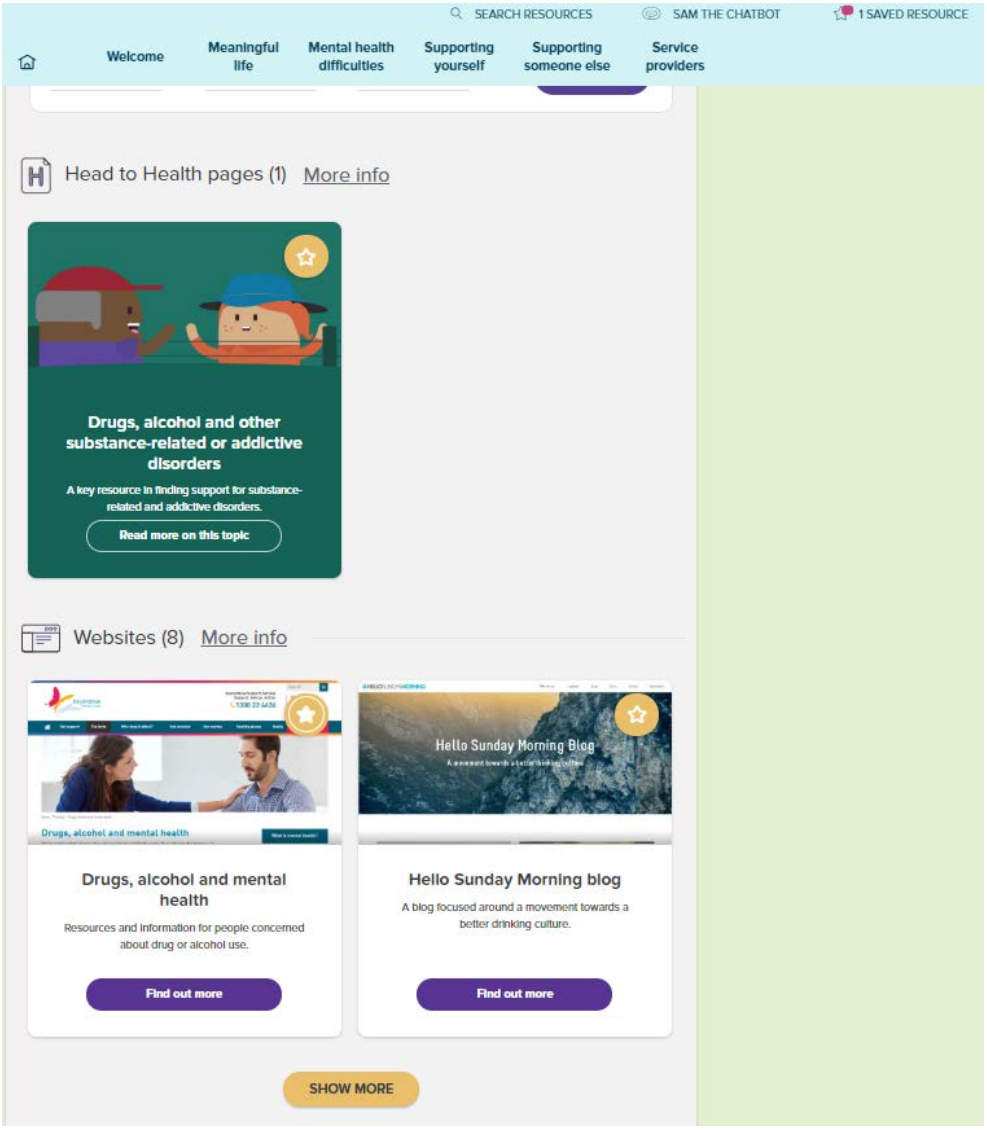


Figure 56: Chatbot Sam, Australian Government Department of Health – Bookmarking.

Architecture components: Bookmarks can be stored by means of cookies or through a local database and can be send over to the chatbot as to provide a better user experience.

6.4.3. Requesting feedback from the users

Some chatbots have the functionality to request feedback from the users. Chatbot Emma (Figure 57) uses this functionality by asking the yes/no question if the offered answer was helpful. The user can click on yes or no. More advanced feedback schemes were discovered during the desk research: section 7.1 dives deeper into the different feedback possibilities.

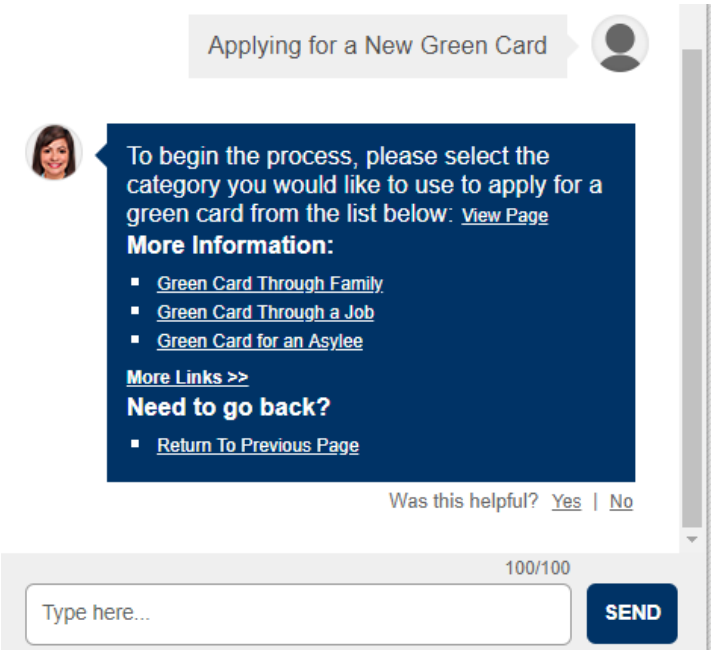


Figure 57: Chatbot Emma, U.S. Citizenship and Immigration Services – Requesting feedback.

Architecture components: Information provided by the user could be stored in the database, which could be useful to score results and monitor the chatbot.

6.4.4. Providing payment services

In the municipality Saint-Pierre-Woluwe, chatbot Pierre is in development. This chatbot will have a payment functionality.

Architecture components: An API that connects to the payment service could be used.

6.4.5. Providing authentication possibility

Another important functionality for public administration chatbots is the possibility for the user to authenticate himself. More information and examples can be found in section 5.1.2.

Providing authentication is linked to the CPSV-AP Evidence class, as it could be mandatory for the user to provide this authentication as evidence for using a specific public service.

Architecture components: An API that connects to the authentication service could be used.

6.5. Components overview

This subsection provides an overview (Table 8) of the different architectural components, discussed in the previous subsections.

Table 8: Overview of chatbot functionality architectural components.

User interaction	Component	Architectural needs
Introduction	Chatbot personality	This information could be stored in a database and retrieved by means of an API.
	Introduction as a bot	This static information could be stored in a database and retrieved by means of an API.
	Multilingual	The multilingual support could be provided by means of API coupling: one for the original content retrieval and one for the pure translation. The first could be connected to an internal database, the second a third party service such as Google Translate or the CEF eTranslation service.
	Privacy policy	This static data could be stored in a database and retrieved by means of an API or linked to a website where data privacy is discussed.
Basic interaction	Channels	Third party libraries or APIs could support audio encoding.
	Menu-based	This information could be stored in a database.
	Multiple answers	This information could be stored in a database.
	Concise answers and links	This information could be stored in a database or retrieved from web pages.
	Requesting user input	Using a location API and a real-time API to check the closest public transport in order to minimize the time needed for the transport.
Redirection	To a human	A messaging API could connect chatbots, chat systems or a mail delivery system.
	To another chatbot	A chatbot API could be used, that can transfer context between the chatbots.
Other functionalities	Filtering	Content can be retrieved from a database and be displayed at an application level.
	Bookmarking	Bookmarks can be stored by means of cookies or through a local database and can be send over to the chatbot as to provide a better user experience.
	Requesting user feedback	Information provided by the user could be stored in the database, which could be useful to score results and monitor the chatbot.
	Payment service	API that connects to the payment service.
	Authentication	An API that connects to the authentication service could be used.

7. GOVERNANCE

An important part of the architecture for public service chatbots is the governance. As public services keep evolving, so should the chatbot, as it would be outdated very soon if it would not be maintained.

7.1. Feedback

Feedback loops are crucial for public service chatbots. Because most chatbots are open-ended, users will say things that cannot be anticipated. Looking at what users are actually saying or asking, the public administration can figure out how to handle these unanticipated requests and adapt to evolving needs.
This document defines two feedback sources[98]:

7.1.1. In-conversation feedback sources

Logging complete user conversations is a great feedback source. Taking a conversation sample and scoring it via KPIs (e.g. “Did the chatbot understand the user’s intent?”) will quickly provide a grasp of how the chatbots is performing.

Another way to provide a simple, frictionless way for users to give chatbot feedback, is by simply introducing response ratings with every reply the chatbot gives. This allows users to flag a response with just a simple tap. When checking these ratings in the conversational analytics, there will be a clear view if a response was a negative or positive experience for the user. The downside to this method is that it does not provide specific feedback.

With sentiment analysis, chatbot analytics can automatically discover chatbot feedback without having to prompt the user at all. By analysing the tone and polarity of terms (whether a user’s vocabulary indicates a positive or negative experience), a feedback chatbot should understand whether it is creating a tense and frustrating experience. For this method, there is a need to label responses or keywords as positive or negative feedback. Table 9 provides examples of positive and negative feedback phrases.

Table 9: Sentiment analysis – Examples.

Sentiment analysis - Examples
Examples of negative feedback could be: “That’s not what I asked for.” “That’s incorrect.” Phrases or words indicating positive feedback could include things like: “Thank you!” “You’re a lifesaver.”

To make best use of this chatbot feedback method, the bot should respond appropriately to the user’s tone. For example, it should thank the user for positive

[98] https://chatbotsmagazine.com/amplifying-user-intelligence-with-chatbot-feedback-loops-b8e6ded391ec

feedback or indicate that it is always happy to help. Negative feedback should prompt the bot to give an apology and ask how its response could be improved.

After the feedback bot has funnelled the user feedback to the developer, it is possible to quickly gauge the user experience in their own words. Sentiment analysis can also present a general idea of what kind of responses from the bot triggers which moods, which is an invaluable addition to general conversational analytics.

Next figures (Figure 58, Figure 59, Figure 60 and Figure 61) are examples of methods to include a feedback functionality in a chatbot.

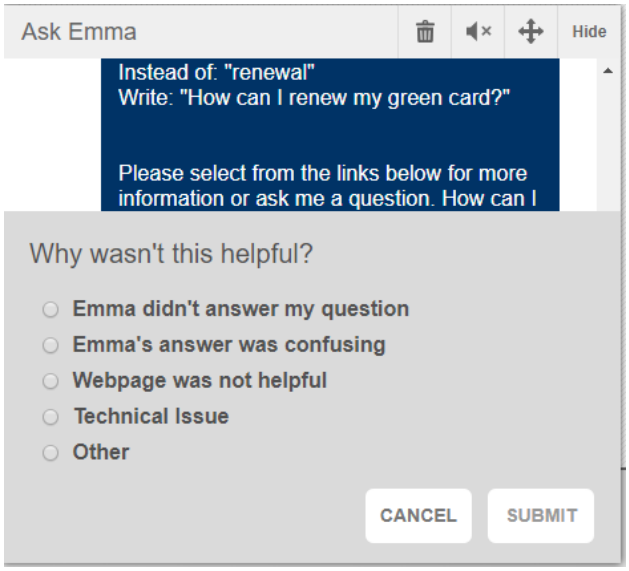


Figure 58: Chatbot Emma – providing feedback options.

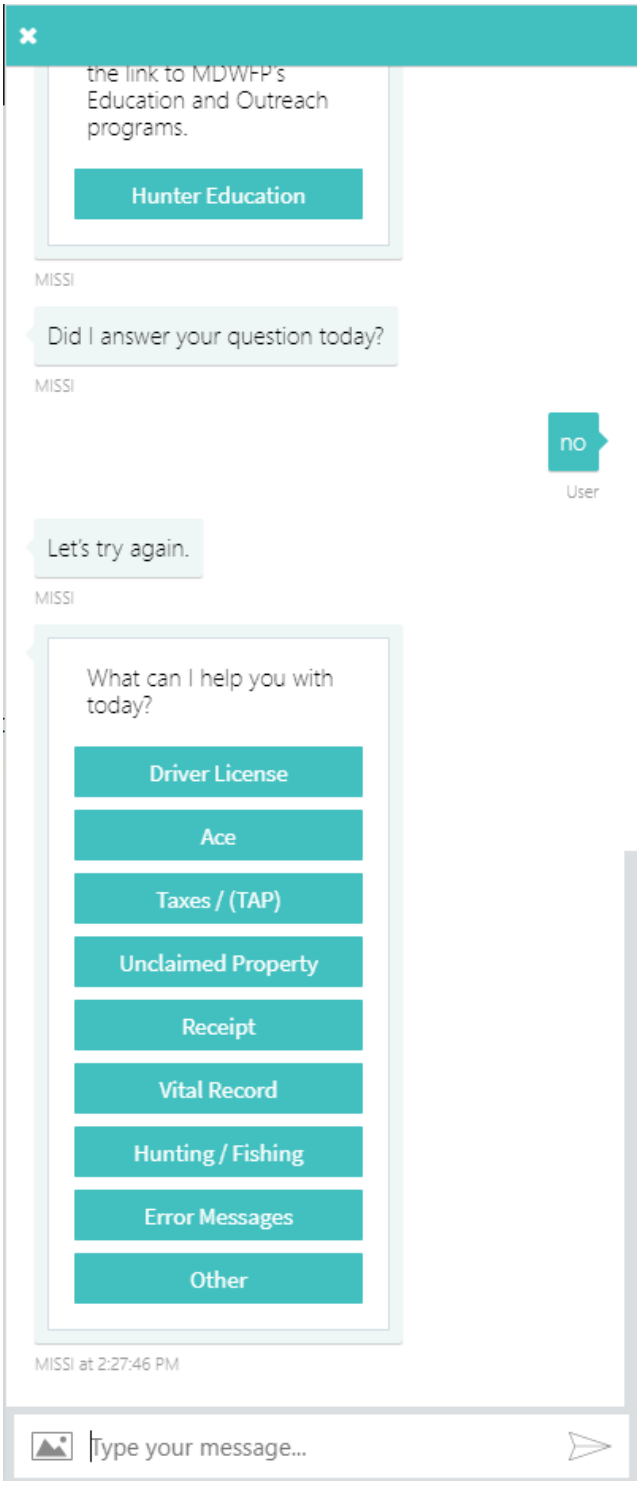


Figure 59: Chatbot MISSI – requesting feedback (yes/no).

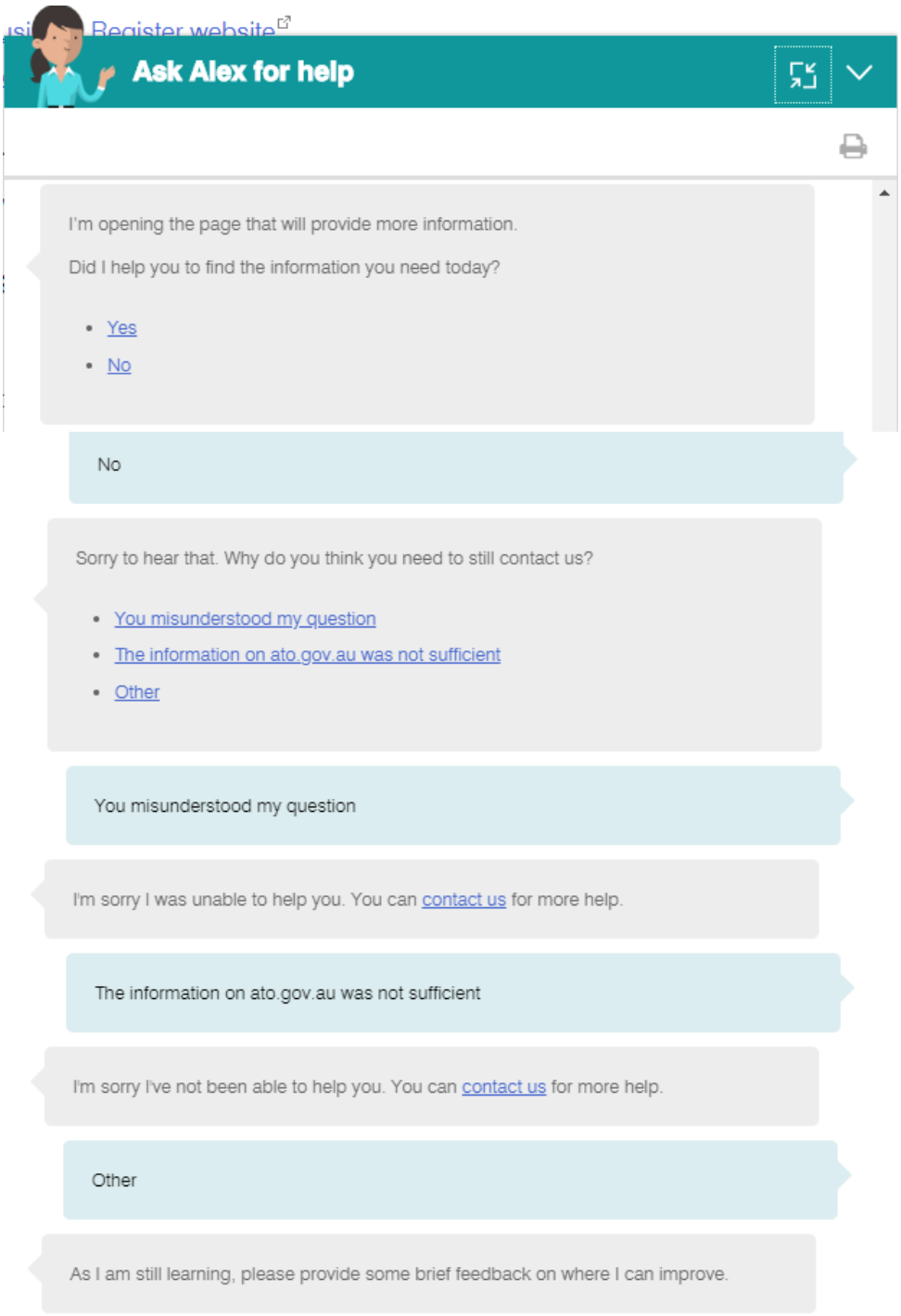


Figure 60: Chatbot Alex – Providing different ways of feedback.

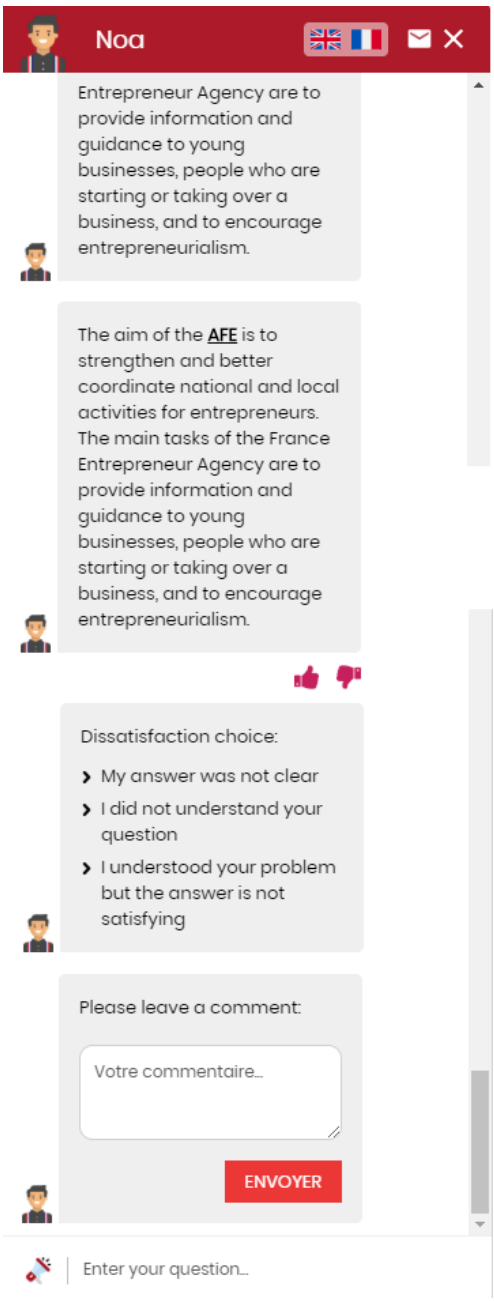


Figure 61: Chatbot Noa – Requesting feedback.

7.1.2. External sources

It is possible to receive chatbot feedback through mail or phone, therefore different channels should be put in place. Different public service channels are defined in the CPSV-AP[99].

[99] Detailed information on the channels can be found in the CPSV-AP specification on Joinup: https://joinup.ec.europa.eu/solution/core-public-service-vocabulary-application-profile/releases

7.2. Learning

Training a chatbot happens considerably faster and on a larger scale than via human training. While normal customer service representatives are given manual instructions, which they must be thorough with, a customer support chatbot is nourished with a large number of conversation logs, and from those logs, the chatbot can understand what type of question needs what kind of answer.

There are two main ways to train a goal-oriented chatbot: supervised learning with an encoder-decoder that directly maps user dialogue to responses. The second one is reinforcement learning, which trains a chatbot through trial-and-error conversations with either real users or a rule-based user simulator.

As discussed in section 4.1.3, generative chatbots are very difficult to build and operate. One way to build them is via recurrent neural networks. More information can be found on specialized sites.[100]

7.2.1. Manual improvements

Based on the feedback, discussed in the previous section, manual improvements can be made to the chatbots. For instance, an option would be to add new intents to increase the coverage on user queries. Another option would be to tweak the onboarding to set expectations around what the chatbot can do. Once these fixes are rolled out, the same feedback sources can be checked to validate the changes made.

7.2.2. Semi-automated improvements

The more generated conversational data, the more possibilities to retrain the underlying machine learning models and the better the accuracy and coverage. Semi-automated training can be done with deep reinforcement learning. More information on this topic can be found on specialized websites.[101]

Although it is technically feasible to fully automate the learning, it would risk losing control of the direction of the chatbot. In general, updates should be rolled out slowly and tested with a small subset of users first before going live to the entire user base.

7.3. Monitoring

Monitoring conversations between the user and the chatbot is crucial to get better understanding about the user needs and, for example, to notice network-level miss-predictions.

---

[100] https://medium.com/@BhashkarKunal/conversational-ai-chatbot-using-deep-learning-how-bi-directional-lstm-machine-reading-38dc5cf5a5a3

[101] https://towardsdatascience.com/training-a-goal-oriented-chatbot-with-deep-reinforcement-learning-part-i-introduction-and-dce3af21d383

When developing VeroBot (Finnish Tax Administration[102]), a very user-centered approach was used. The most important thing for an organisation in the development of a chatbot is to know the needs of its users.[103]

The coming subsections describe 15 key metrics for chatbot analytics (Table 10, Table 11 and Table 12).[104] Any mix of these metrics can be used to monitor a chatbot. Public administrations should make sure that the chosen metrics are tailored to the specific chatbot type and use case. For instance, if the chatbot is providing a specific service, the ideal conversation might be short but efficient, while a social chatbot might need longer conversations that are optimised for engagement.

It is also important to see these metrics in context, e.g. the length of a chatbot conversation could mean the chatbot had a high quality conversation that kept the user engaged, but could also mean that the chatbot wasn't able to provide the right answer while the user kept trying to get the right answer.

7.3.1. User metrics

Table 10: Chatbot analytics - User metrics.

User metrics	
Total users	This is the most basic metric. It captures the number of people using the chatbot. This matters because its trend shows the change in the number of users and therefore the amount of data the chatbot has been exposed to.
Active users	Active users can be defined as the people who read a message in the chatbot in a defined period.
Engaged users	The engaged users are the ones who communicate with the chatbot. They receive and send messages. This is important since the chatbot will be able to provide the conversation statistics based on this sub-sample. This metric provides one perspective on the effectiveness of the chatbot.
New users	This metric captures the overall success of the chatbot's visibility.

7.3.2. Message metrics

Table 11: Chatbot analytics – Message metrics.

Message metrics	
Conversation starter messages	This is the number of messages where a user starts the interaction by the bot.
Bot messages	Bot messages are the total number of messages sent by the chatbot in each interaction. This measures the length of a conversation between a customer and the chatbot. In case of misunderstanding or failure to comprehend the input by the user, the chatbot will say similar words repeatedly.

[102] See section 4.1.2 for a more detailed analysis of this chatbot.
[103] Henri Harvima, Finnish Tax Administration.
[104] https://blog.aimultiple.com/chatbot-analytics/

In messages	This category shows the messages sent by the user and makes clear whether the user engages with the chatbot or not. If this category is significantly low, there might not be a need to use a chatbot.
Miss messages	Miss messages are messages the chatbot cannot process.
Total conversations	The total conversations are the number of conversations started and successfully completed on a given day.
New conversations	This is the amount of new conversations started. This captures both the inexperienced users and the conversations that are initiated by the returning users on a different matter.

7.3.3. Bot metrics

Table 12: Chatbot analytics – Bot metrics.

Bot metrics	
Retention rate	This is the percentage of users that return to using the chatbot on a given time frame.
Goal Completion Rate (GCR)	This captures the percentage of successful engagements through the chatbot. This shows the number of times the bot successfully processed the input and provided the asked information or service.
Goal completion time / taps /messages	Chatbots need to provide a seamless and efficient experience and have plenty of substitutes such as the public administration's web pages or apps. Minimizing the effort to complete a goal can improve user experience.
Fall back rate (FBR)	No robot is perfect, therefore chatbots are expected to fail sometimes, but what really matters is to know if those failures happen regularly or only in some extreme cases. The fall back rate captures that information. This is the percentage of times the chatbot failed or experienced a near-failure situation.
User satisfaction	Users could rate their experience to achieve further product excellence. More feedback possibilities can be found in section 7.1.

8. SAT VIEWS

The European Interoperability Reference Architecture (EIRA) is an architecture content metamodel defining the most salient architectural building blocks (ABBs) needed to build interoperable e-Government systems. The EIRA provides a common terminology that can be used by people working for public administrations in various architecture and system development tasks. The EIRA uses (and extends) the ArchiMate language as a modelling notation and uses service orientation as an architectural style.[105]

A Solution Architecture Template (SAT) is a specification extending the EIRA by providing support to solution architects in a specific solution domain. A SAT contains a motivation (principles, requirements), a goal and a description of the supported functionalities, a sub-set of the EIRA core Architecture Building Blocks (ABBs) covering the four views, a set of specific ABBs extending EIRA's views enabling specific functionalities to be provided by implementations derived from the SAT and the interoperability specifications of selected ABBs, and a narrative for each EIRA view.[106]

The purpose of the Catalogue of Services SAT is to provide guidance by defining a minimal, but holistic (legal, organisational, semantic and technical) interoperability architecture for the Public Catalogue of Services, and concretely in this case the public service chatbot architecture. The CoS SAT allows businesses and public administrations to have a common understanding of the most salient building blocks from the perspective of interoperability.[107]

As a data standard, the CPSV-AP fits in the EIRA architecture, divided in four views.[108]

8.1. The legal view

The CPSV-AP supports the Digital Single Market strategy for Europe, which is realized by several binding and non-binding instruments. In particular, in the context of a chatbot, the GDPR[109] regulation is a binding instrument that should be taken into account when a public service deals with sensitive data. Figure 62 shows the EIRA CoS SAT legal view.

[105] https://joinup.ec.europa.eu/solution/eira/about
[106] https://joinup.ec.europa.eu/release/sat-cos/v100
[107] https://joinup.ec.europa.eu/release/sat-cos/v100
[108] The EIRA CoS SAT ArchiMate file can be found on Confluence (limited access).
[109] See section 5.1.2.

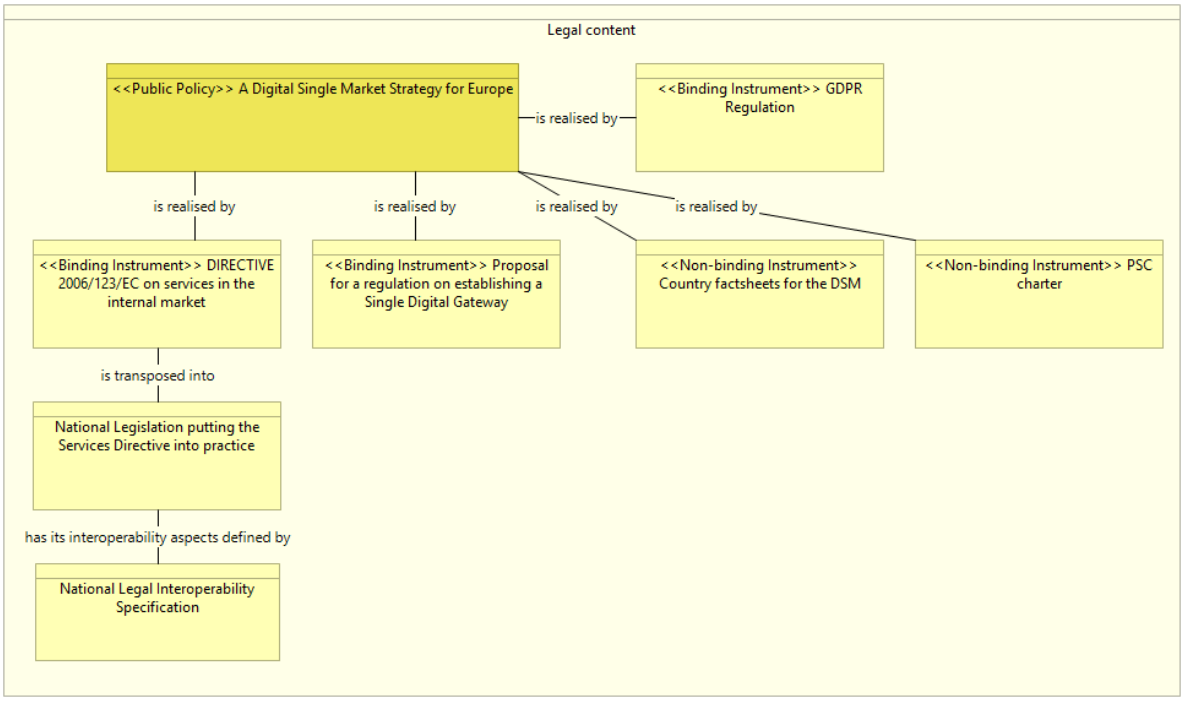


Figure 62: EIRA CoS SAT: Legal view.

8.2. The organisational view

As described in section 3.2, public organisations already provide catalogue of public services via online portals. Such portals can be integrated with chatbot services in order to enrich the service delivery. Figure 63 shows the EIRA CoS SAT organisational view.

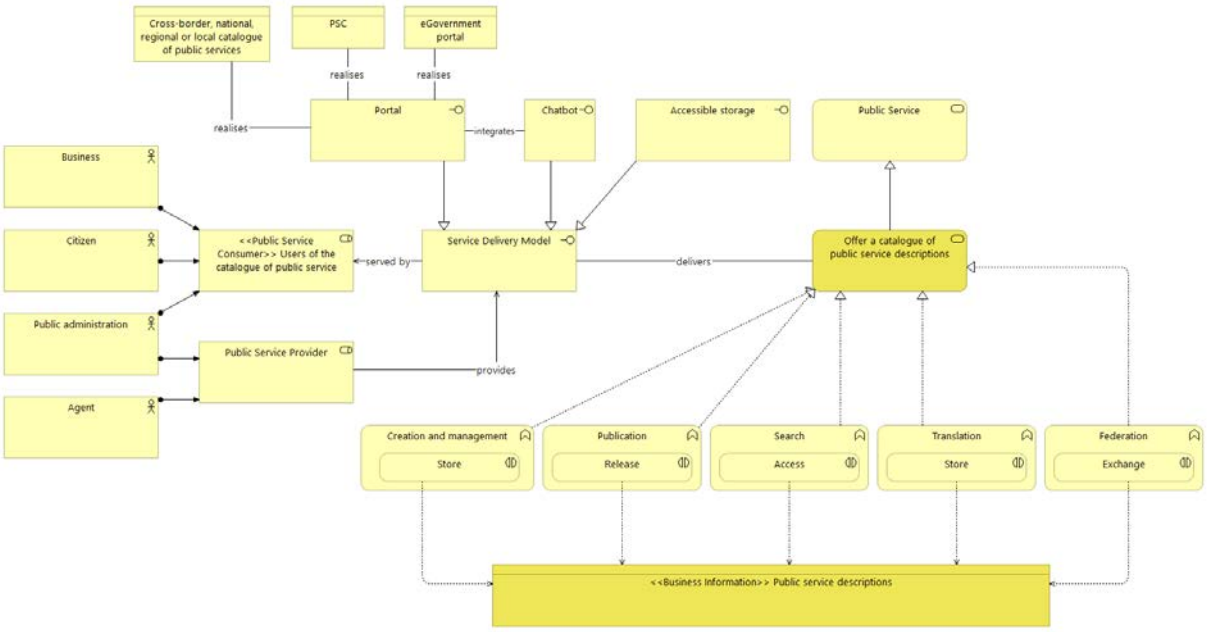


Figure 63: EIRA CoS SAT: Organisational view.

8.3. The semantic view

For the usage of a chatbot, a machine-readable format can come as a result of speech-to-text or text-to–speech operations. Figure 64 shows the EIRA CoS SAT semantic view.

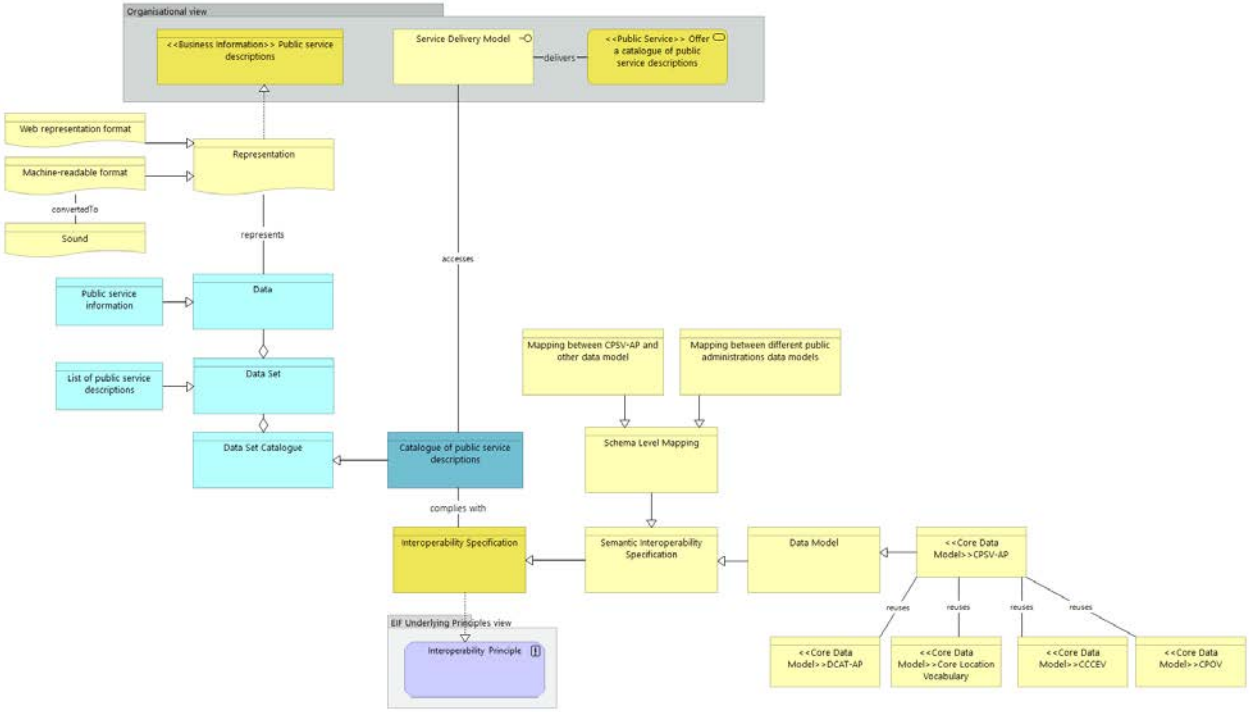


Figure 64: EIRA CoS SAT: Semantic view.

8.4. The technical view

In addition to the Metadata Management Service and the Data Exchange Service, a chatbot could use a Natural Language Processing Service based on services provided through different APIs, e.g. a translating service, a speech-to-text service, a text-to-speech service … Figure 65 shows the EIRA CoS SAT technical view on the infrastructure, while Figure 66 shows the EIRA CoS SAT technical view on the application.

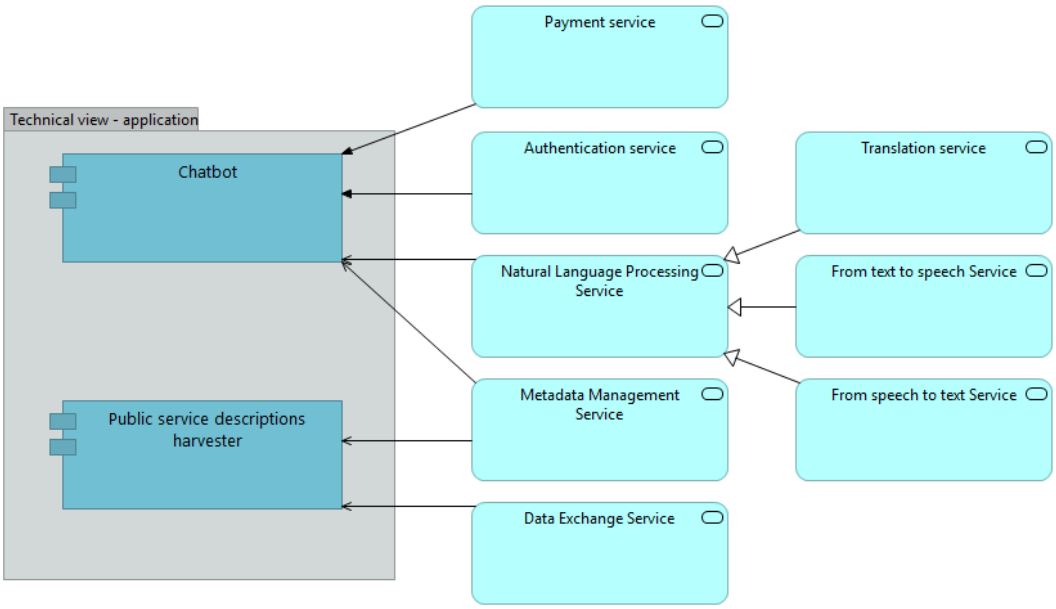


Figure 65: EIRA CoS SAT: Technical view – infrastructure.

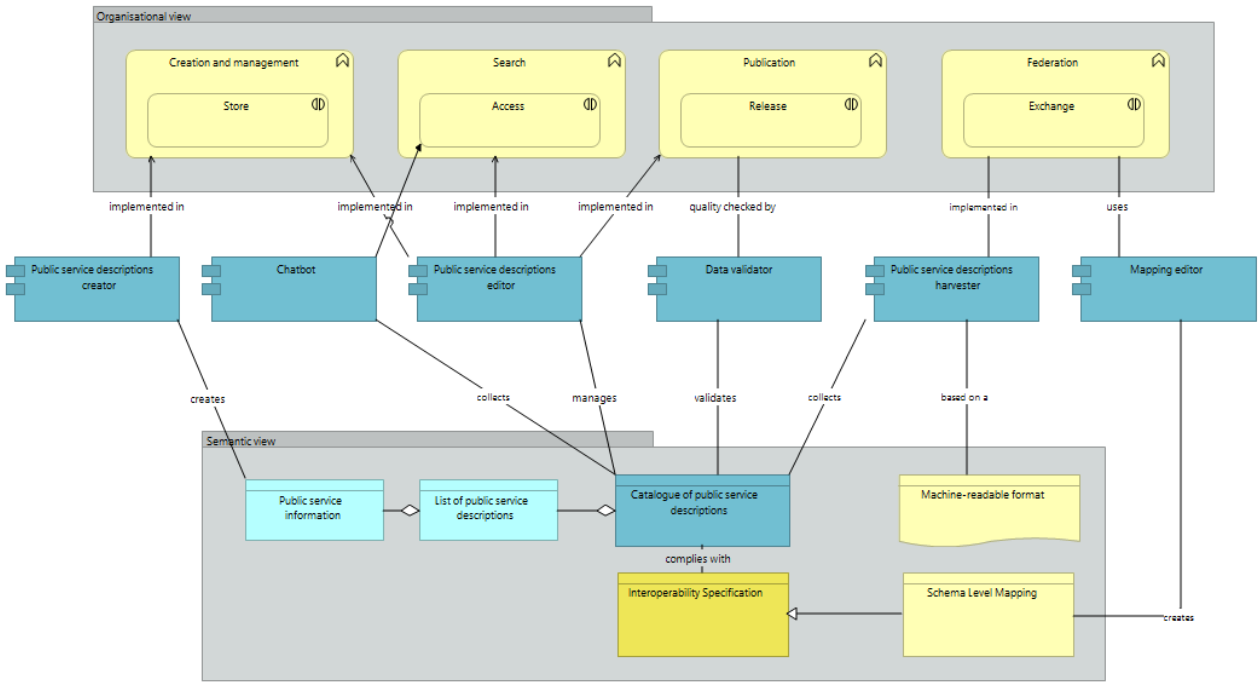


Figure 66: EIRA CoS SAT: Technical view – application.

9. RECOMMENDATIONS

This section provides an overview of recommendations when building public service chatbots.

9.1. Type and platform

This document recommends the use of a federated network of goal-oriented chatbots. This way, different information services as well as the provision of public services can be integrated in chatbots with relatively limited resources. As only the general design principles have to be standardised and afterwards only a limited cooperation needs to take place to keep aligning the chatbots, the risk of cooperation-related issues is much lower than when making a generative chatbot.

Moreover, public administration departments can act independently: they would have the possibility to join the chatbot network, but not the obligation. Another department that would like to make multiple more advanced chatbots on the other hand, would also have this opportunity. The same principle works on a European level.

A recommended platform, for public administrations to integrate one or more chatbots in, is through the public administrations webpage as some platforms might be used less or disappear in the future, but also for security reasons (see section 5.4). Public administrations could additionally implement the chatbot in social network platforms or messaging apps, but this would have an additional cost because of the security concerns that should be addressed.

A chatbot is only one search functionality or one way to provide a public service to support the user. Keeping it visible on the webpage as an icon or a bar seems the most appropriate as the user can still make use of the webpage while using the chatbot. Public administrations should keep in mind to optimise the chatbot experience for mobile as well, as there is a huge amount of people who look for information through their phones. It would be recommended to integrate the chatbots in the voice assistants as well, since voice queries are on the rise[110]. Figure 67 shows that voice assistants are (mainly) used for looking up information (Ask a question).

[110] According to Google, 20% of all searches are voice. According to Comscore, 50% of all online searches will be voice-based by 2020.

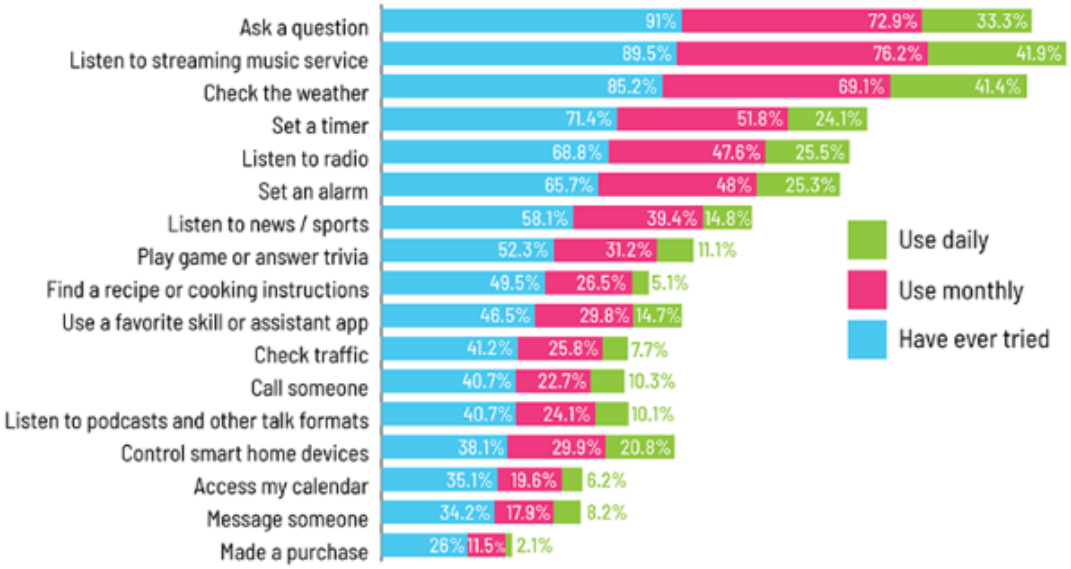


Figure 67: Smart Speaker Use Case Frequency January 2018.[111]

9.2. Building the chatbot

As public administrations would like to make the search for information and services easier, this document recommends using a broad view when developing chatbots. This means that public administration chatbots should include as many **relevant functionalities** as possible (e.g. payment functionality).

For the provision of public services, **authentication** is often necessary. This document advises to implement the authentication option in the relevant chatbots (e.g. through eID or mobile app) and to base it on the eIDAS principles[112] as to make it possible for users to use eID across European borders.

This document advises to use **multilingual** chatbots and more specifically a chatbot that combines the native language(s) with English, as to be useful as well for people who are not proficient in the native language of the country.

As (some) public administrations receive heavy traffic (see e.g. Figure 68), a **cloud-based** technological solution is recommended, which is also in line with the European Cloud Strategy[113].

[111] https://www.smartsheet.com/voice-assistants-artificial-intelligence
[112] https://ec.europa.eu/digital-single-market/en/trust-services-and-eid
[113] https://ec.europa.eu/info/sites/info/files/ec_cloud_strategy.pdf

Figure 68: Users that had at least one session in a week on gov.uk.[114]

It is hard to recommend a specific cloud-based bot-building platform or ecosystem for public administrations, as there are many providers. Figure 69 provides a non-exhaustive overview of existing bot building platforms. It is also possible to create a public administration cloud platform to avoid vendor lock-in (e.g. the UK[115]).

This document does advise to use a cloud-based solution, but to recommend a specific architecture, a provider or a custom solution, it is needed to (1) evaluate the feasibility of various vendor solutions, (2) define the scope of capabilities and data requirements, and (3) determine a strategy for integration with existing back-office systems.

Figure 69: Overview of bot building platforms.[116]

[114] https://www.gov.uk/performance/site-activity
[115] https://www.cloud.service.gov.uk/
[116] https://blog.aimultiple.com/chatbot-ecosystem/

On the architectural components, discussed in section 5, this document recommends to use the pattern matching, and an NLU – NLP – NLG combination. All of those elements are needed to create a great chatbot. NLU takes up the understanding of the data based on grammar, the context in which it was said and decide on intent and entities. NLP will convert the text into structured data. NLG generates text based on structured data.[117] The big five bot building platform providers (see Table 13) use a similar architecture. The advice on specific specifications to use depends on which kind of chatbot / chatbot network the public administration would like to use. A vendor/custom-built framework evaluation would be necessary as well on this topic.

Table 13: The big five bot building platforms.

Framework	Link
Amazon Lex	https://aws.amazon.com/lex/
Dialogflow (Google)	https://dialogflow.com/
IBM Watson	https://www.ibm.com/watson/how-to-build-a-chatbot
Luis (Microsoft)	https://www.luis.ai/home
Wit.ai (Facebook)	https://wit.ai/

Lastly, this document recommends using APIs, as they are a key building block of chatbots. Some public administrations already publish APIs on a government portal (as mentioned in section 3.1.1), but this document advises to step up this effort as these APIs do not only support public administration chatbots, they could also be used in private sector chatbots. To be more effective and to better leverage the APIs in the private sector, public administration APIs could be published on private platforms as well (e.g. The Programmable Web[118], Rapidapi[119], …).

[117] https://towardsdatascience.com/nlp-vs-nlu-vs-nlg-know-what-you-are-trying-to-achieve-nlp-engine-part-1-1487a2c8b696

[118] https://www.programmableweb.com/category/government/api

[119] https://rapidapi.com/search/government

10. NEXT STEPS

This document described a public service chatbot architecture, from needs to governance. Using a chatbot has multiple benefits, but there is of course a cost for implementing this technology. Nonetheless, the more complex the structure of information, the more useful a chatbot would be as the main benefit of a chatbot is the decrease of user search time. As there is a huge amount of public services in all countries, chatbots are definitely useful here. Especially on a European level, where another level of complexity is added, a (network of) chatbots(s) would be extremely useful. In this context, this document provides a few possible next steps to make this a reality.

10.1. Supporting the Digital Single Market

Chatbot technology fits in the Digital Single Market Strategy[120]. The DSM[121] Strategy is built on three pillars:

- Access: better access for consumers and businesses to digital goods and services across Europe;
- Environment: creating the right conditions and a level playing field for digital networks and innovative services to flourish;
- Economy & Society: maximising the growth potential of the digital economy.

Implementing chatbot technology on Member State and European level would support all three pillars:

- Chatbot technology provides easier and faster access to public services and information than traditional channels (see section 3);
- Creating common design principles and a platform to exchange ideas, information and best practices on chatbots would start an interesting environment (see section 3.3.3);
- Being a pioneer in implementing this technology will leverage it to the economy as a whole and to society, as chatbot (development) information would be more visible, centralised and thus lowering the cost of building a new one. On top of that, a chatbot is great marketing for a public administration to be seen as a technology savvy organisation.

More concretely, chatbot technology could support the SDG greatly (see 3.3), a DG GROW[122] action. To achieve this goal,

- A regular communication channel could be set up with DG GROW to discuss the development and governance of a chatbot (network) on the Your Europe portal;
- An evaluation could be made of the feasibility of various vendor or custom solutions based on providing the 21 procedures through a chatbot;

[120] https://ec.europa.eu/digital-single-market/en/policies/shaping-digital-single-market

[121] Digital Single Market

[122] Directorate-General Internal Market, Industry, Entrepreneurship and SMEs

- The capabilities and data requirements and determining a strategy for integration with existing back-office systems, based on the 21 procedures, could be scoped;
- A chatbot network PoC could be developed to measure the feasibility of the solution; and
- A chatbot cooperation platform could be set up. Next to exchanging ideas, information and best practices, this platform could be used to build a community, to create public administration design principles and to create chatbot governance.

10.2. Supporting Member States

Some member states already make use of chatbots (see section 11.211.2), but are not developing them with the CPSV-AP in mind. Considering this, a possible chatbot network, as mentioned before, would transfer the end-user to Member State chatbots. A Member State chatbot PoC that is based on the CPSV-AP could be considered as a next step to align with these use cases.

11.SOURCES

11.1. Used resources

Table 14: Used resources.

Description	Region	Link
API Catalogue	Global	https://www.programmableweb.com/category/government/api
API Catalogue	Global	https://rapidapi.com/search/government
API portal	FR	https://api.gouv.fr/
Authentication	Global	https://www.itsme.be/en/lang-splash
Chatbot analytics	Global	https://blog.aimultiple.com/chatbot-analytics/
Chatbot architecture	Global	https://thinkgrowth.org/how-to-nail-a-great-chatbot-experience-7bbfd98026d6
Chatbot architecture	Global	https://dzone.com/articles/understanding-architecture-models-of-chatbot-and-r
Chatbot architecture	Global	https://medium.com/analytics-vidhya/building-a-simple-chatbot-in-python-using-nltk-7c8c8215ac6e
Chatbot architecture	Global	https://towardsdatascience.com/architecture-overview-of-a-conversational-ai-chat-bot-4ef3dfefd52e
Chatbot architecture	Global	https://bigdata-madesimple.com/how-do-chatbots-work-an-overview-of-the-architecture-of-a-chatbot/
Chatbot architecture	Global	https://www.marutitech.com/chatbots-work-guide-chatbot-architecture/
Chatbot architecture	Global	https://medium.com/@surmenok/chatbot-architecture-496f5bf820ed
Chatbot architecture	Global	https://www.newgenapps.com/blog/complete-guide-on-chatbot-architecture
Chatbot architecture	Global	https://blog.aimultiple.com/chatbot-ecosystem/
Chatbot architecture	Global	https://towardsdatascience.com/nlp-vs-nlu-vs-nlg-know-what-you-are-trying-to-achieve-nlp-engine-part-1-1487a2c8b696
Chatbot databases	Global	https://www.researchgate.net/publication/315364724_Chatbot_Using_a_Knowledge_in_Database_Human-to-Machine_Conversation_Modeling
Chatbot development platform	Global	https://dev.botframework.com/
Chatbot development platform	Global	https://wit.ai/
Chatbot development platform	Global	https://dialogflow.com/
Chatbot development platform	Global	https://www.ibm.com/watson/how-to-build-a-chatbot

Chatbot development platform	Global	https://home.pandorabots.com/home.html
Chatbot development platform	Global	https://botpress.io/
Chatbot development platform	Global	https://botkit.ai/
Chatbot development platform	Global	https://rasa.com/
Chatbot development platform	Global	https://github.com/gunthercox/ChatterBot
Chatbot development platforms	Global	https://geekflare.com/chatbot-development-frameworks/
Chatbot feedback	Global	https://chatbotsmagazine.com/amplifying-user-intelligence-with-chatbot-feedback-loops-b8e6ded391ec
Chatbot general report	Global	https://www.drift.com/blog/chatbots-report/
Chatbot learning	Global	https://medium.com/@BhashkarKunal/conversational-ai-chatbot-using-deep-learning-how-bi-directional-lstm-machine-reading-38dc5cf5a5a3
Chatbot learning	Global	https://towardsdatascience.com/training-a-goal-oriented-chatbot-with-deep-reinforcement-learning-part-i-introduction-and-dce3af21d383
Chatbot preprocessing	Global	https://medium.com/@surmenok/natural-language-pipeline-for-chatbots-897bda41482
Chatbot search term evolution	Global	https://trends.google.com/trends/explore?date=2009-06-10%202019-06-10&q=chatbot
Chatbot security	Global	https://dzone.com/articles/4-chatbots-security-measures-you-absolutely-need-t
Cloud platform	UK	https://www.cloud.service.gov.uk/
Cloud Strategy	EU	https://ec.europa.eu/info/sites/info/files/ec_cloud_strategy.pdf
Core Vocabularies	EU	https://joinup.ec.europa.eu/solution/e-government-core-vocabularies
CoS Workshop	EU	https://joinup.ec.europa.eu/collection/semantic-interoperability-community-semic/event/technical-workshop-catalogue-public-services-vis-vis-sdg-implementation-16-may-2019-930-1730-cet
CPSV-AP	EU	https://joinup.ec.europa.eu/solution/core-public-service-vocabulary-application-profile/about
CPSV-AP Specification	EU	https://joinup.ec.europa.eu/release/core-public-service-vocabulary-application-profile/221

Digital public services	FI	http://www.finlex.fi/fi/laki/ajantasa/2016/20160571
Digital Single Market Strategy	EU	https://ec.europa.eu/digital-single-market/en/policies/shaping-digital-single-market
eGovernment Action Plan	EU	https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52016DC0179&from=EN
eIDAS	EU	https://ec.europa.eu/digital-single-market/en/trust-services-and-eid
EIRA	EU	https://joinup.ec.europa.eu/solution/eira/about
European Data Portal	EU	https://www.europeandataportal.eu/en/homepage
GDPR	EU	https://eur-lex.europa.eu/eli/reg/2016/679/oj
Government AI strategies	Global	https://emerj.com/ai-sector-overviews/ai-government-current-ai-projects-public-sector/
Government AI strategies	Global	https://medium.com/politics-ai/an-overview-of-national-ai-strategies-2a70ec6edfd
Government AI strategies	UK	https://www.gov.uk/government/publications/government-transformation-strategy-2017-to-2020/government-transformation-strategy
Government AI strategies	US	https://www.cio.gov/fed-it-topics/modernization/digital-strategy/
Government AI strategies	Global	http://www.oecd.org/gov/digital-government/Digital-Government-Strategies-Welfare-Service.pdf
IDPR	EU	https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32018R1725
itsme	BE	https://www.itsme.be/en/lang-splash
Lexical database	Global	https://wordnet.princeton.edu/
MySQL	Global	https://www.mysql.com/
NLP Library	Global	https://www.nltk.org/
NLP Library	Global	https://nlp.stanford.edu/software/
NLP Library	Global	https://spacy.io/
NLU Toolkit	Global	https://opensource.google.com/projects/syntaxnet
Once-Only Principle	EU	http://toop.eu/once-only
Open data catalogue	FR	https://www.data.gouv.fr/fr/
Private sector chatbots sectors + platforms	Global	https://thinkgrowth.org/100-best-bots-for-brands-and-businesses-c281c5279229
PSI directive	EU	https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32013L0037&from=FR

Public service catalogue	FI	https://www.suomi.fi/frontpage
Public service catalogue	NO	https://fellesdatakatalog.brreg.no/
Public service catalogue	UK	https://standards.esd.org.uk/?uri=service%2F1296&tab=details
Real time data API provider	Global	https://sendbird.com/
Real time data API provider	Global	https://www.pubnub.com/
Relational database	US	https://commons.wikimedia.org/w/index.php?curid=17875170
SAT CoS	EU	https://joinup.ec.europa.eu/release/sat-cos/v100
Single Digital Gateway Regulation	EU	http://data.consilium.europa.eu/doc/document/ST-10069-2018-INIT/en/pdf
Voice Assistants	Global	https://www.smartsheet.com/voice-assistants-artificial-intelligence
Web traffic	UK	https://www.gov.uk/performance/site-activity
Your Europe portal	EU	https://europa.eu/youreurope/citizens/index_en.htm

11.2. Overview of researched public administration chatbots

Table 15: Chatbots researched for this document.

Chatbots, EU bots in bold	Organisation	Link
Alex	Australian Taxation Office	https://www.ato.gov.au/
BonnBot	Bonn city	https://govbot.bonn.de/
Chip	Los Angeles City	https://www.labavn.org/#
Citibot	North Charleston City	https://www.northcharleston.org/connect/
Dolores Landingham Bot	U.S General Services Administration	https://github.com/18F/dolores-landingham-slack-bot
EMMA	U.S. Citizenship and Immigration Services	https://www.uscis.gov/emma
Gov.sg	Singapore Government	https://www.facebook.com/pg/gov.sg/about/?ref=page_internal
HELP	Pune Municipal Corporation	https://pmc.gov.in/en
Jamie	Inland Revenue Authority of Singapore	https://www.iras.gov.sg/irashome/default.aspx
Kamu	Finnish Immigration Service	http://startingupsmoothly.fi/
Missi	Mississippi government	https://www.ms.gov/Technology
Noa	prefecture of Île-de-France	https://french-tech-central.com/

PatRek	Finnish Patent and Registration Office	http://startingupsmoothly.fi/
Pierre	Municipality Woluwe-Saint-Pierre	Under development
Rammas	Dubai Electricity & Water Authority	https://www.dewa.gov.ae/en/rammas
Sam	Australian Government, Department of Health	https://headtohealth.gov.au/sam-the-chatbot
Sigma	Agency for Administrative Modernisation	https://eportugal.gov.pt/
TravelBot	Transport for London	https://www.facebook.com/tfltravelbot/
VeroBot	Finnish Tax Administration	http://startingupsmoothly.fi/
Virtueller Bürger-Service-Assistent	City of Berlin	https://service.berlin.de/virtueller-assistent/virtueller-assistent-606279.php

ANNEX I. SINGLE DIGITAL GATEWAY REGULATION.

Table 16: Annex II, Single Digital Gateway Regulation.

Life events	Procedures	Expected output subject to an assessment of the application by the competent authority in accordance with national law, where relevant
Birth	Requesting proof of registration of birth	Proof of registration of birth or birth certificate
Residence	Requesting proof of residence	Confirmation of registration at the current address
Studying	Applying for a tertiary education study financing, such as study grants and loans from a public body or institution	Decision on the application for financing or acknowledgement of receipt
	Submitting an initial application for admission to public tertiary education institution	Confirmation of the receipt of application
	Requesting academic recognition of diplomas, certificates or other proof of studies or courses	Decision on the request for recognition
Working	Request for determination of applicable legislation in accordance with Title II of Regulation (EC) No 883/2004 (1)	Decision on applicable legislation
	Notifying changes in the personal or professional circumstances of the person receiving social security benefits, relevant for such benefits	Confirmation of receipt of notification of such changes
	Application for a European Health Insurance Card (EHIC)	European Health Insurance Card (EHIC)
	Submitting an income tax declaration	Confirmation of the receipt of the declaration
Moving	Registering a change of address	Confirmation of deregistration at the previous address and of

		the registration of the new address
	Registering a motor vehicle originating from or already registered in a Member State, in standard procedures (2)	Proof of registration of a motor vehicle
	Obtaining stickers for the use of the national road infrastructure: time-based charges (vignette), distance-based charges (toll), issued by a public body or institution	Receipt of toll sticker or vignette or other proof of payment
	Obtaining emission stickers issued by a public body or institution	Receipt of emission sticker or other proof of payment
Retiring	Claiming pension and pre-retirement benefits from compulsory schemes	Confirmation of the receipt of the claim or decision regarding the claim for a pension or pre-retirement benefits
	Requesting information on the data related to pension from compulsory schemes	Statement of personal pension data
Starting, running and closing a business	Notification of business activity, permission for exercising a business activity, changes of business activity and the termination of a business activity not involving insolvency or liquidation procedures, excluding the initial registration of a business activity with the business register and excluding procedures concerning the constitution of or any subsequent filing by companies or firms within the meaning of the second paragraph of Article 54 TFEU	Confirmation of the receipt of notification or change, or of the request for permission for business activity
	Registration of an employer (a natural person) with compulsory pension and insurance schemes	Confirmation of registration or social security registration number
	Registration of employees with compulsory pension and insurance schemes	Confirmation of registration or social security registration number

	Submitting a corporate tax declaration	Confirmation of the receipt of the declaration
	Notification to the social security schemes of the end of contract with an employee, excluding procedures for the collective termination of employee contracts	Confirmation of the receipt of the notification
	Payment of social contributions for employees	Receipt or other form of confirmation of payment of social contributions for employees

ANNEX II. LIST OF ACRONYMS

Name	Meaning
ABB	Architectural Building Blocks
API	Application Programming Interface
CoS	Catalogue of Services
CPSV-AP	Core Public Service Application Profile
DSM	Digital Single Market
EIRA	European Interoperability Reference Architecture
EU	European Union
FBR	Fall Back Rate
GCR	Goal Completion Rate
GDPR	General Data Protection Regulation
NER	Named Entity Recognition
NLG	Natural Language Generation
NLP	Natural Language Processing
NLU	Natural Language Understanding
OECD	Organisation for Economic Co-operation and Development
OOP	Once Only Principle
PSI Directive	Public Sector Information Directive
SAT	Solution Architecture Template
SDG	Single Digital Gateway
SDGR	Single Digital Gateway Regulation
UK	United Kingdom
US	United States

ANNEX III. SECURITY TESTING GUIDE

Table 17: Security testing guide.

Security control	Test Case	Recommendation
Use of public chat platforms	Does the platform support end to end encryption	If not, use public platforms to host bots meant for informational purposes only. For any sensitive functionalities, the user may be redirected to a secure page
Encryption	Is end to end encryption implemented	The channel must be encrypted at all points between the sending and receiving systems, if sensitive information is being handled
	Is Transport Layer Security enforced securely	Implement TLS 1.2 or above with AES 256 bit encryption using only secure ciphers
	Does the chatbot handle sensitive data such as financial data or Personally Identifiable Information (PII)	Implement message encryption in addition to channel encryption using a secure symmetric cipher such as AES 256 bit
	Is Perfect forward secrecy implemented	Ensure that different encryption key s are generated per session
Authentication	Authentication tokens	Ensure that authentication tokens are generated and used in a secure manner, as per best practice session management guidelines
	Authentication timeout	Authentication tokens must timeout after a fixed time period
	Multi-factor authentication	Re-authenticate the user through means such as username and password, eID, OTP, biometric authentication before handling sensitive information
Integrity Check	Is an integrity check mechanism in place to prevent tampering / corruption of data	A secure hash of the requested pay load can be sent along with the message and compared at server side. Implement on chatbots handling user functionalities and not just serving informational functions
Social engineering attacks	Has the organization implemented any mechanisms to create awareness against phishing attacks	For example, banks may give a warning to customers intimating them to not communicate any card details or other sensitive information over chat
	Chatbot identity verification	Implement a mechanism wherein a customer can verify that he is communicating with the intended receiver. E.g.: Facebooks “verified” icon
Infrastructure Security	Secure Network Architecture	Ensure that the application architecture complies with Secure Network Architecture principles

	Use of secure components	Ensure that the underlying technology has no known vulnerabilities
	Patch management	Implement a robust and timely patch management cycle
	Secure Configuration	Implement Server and Infra hardening through secure configuration. E.g.: CIS benchmarks
Data Storage and Handling	Only necessary data is collected by the application	Decrease the organization's liability by limiting the data collected
	Data Encryption	Ensure that database table encryption is implemented for PII and financial data
	Data Access Control	Ensure that only necessary personnel have access to sensitive data. Review access control measures implemented.
	Data Masking	Ensure that critical information such as card number, PAN number and account number are masked when displayed in the chat window
	Data Retention	Ensure that data retention period is defined and is being adhered to. All copies of data should be disposed of in a secure manner when the retention period elapses
	Analysis of Data	If data is analyzed and processed by the organization or sent to a third party, ensure that the data is anonymised / pseudonymised to protect identities of customers
Self-Destruction of Messages	Do messages in the chat window disappear automatically after a fixed period of time	Ensure that chat messages self-destruct
Application Security	Input Validation	Ensure that only desirable input is entered and data is displayed in a sanitised format to protect from injection and scripting attacks
	Database injection (SQL injection, XPath injection)	Ensure that backend database queries are secure from injection attacks.
	Code Review	Analyse application source code to check for vulnerabilities in data collection and handling
Pollution in communications channel	Are there checks in place to detect anomalous activity from users?	
Use of external NLP Services	Does the application send chatbot data to third party NLP services (e.g.: IBM Watson)	Ensure that data sent to third party services does not contain PII or sensitive information
Logging and Access Rights	Are logs of chatbot sessions collected	Collect detailed logs of all chatbot sessions
	Are these logs encrypted	Store the logs in an encrypted format

	How are the logs stored and secured	Is access control to the logs monitored and restricted

An action supported by ISA²

ISA² is a EUR 131 million programme of the European Commission which develops digital solutions that enable interoperable cross-border and cross-sector public services, for the benefit of public administrations, businesses and citizens across the EU. ISA² supports a wide range of activities and solutions, among which is the Semantic Interoperability Community (SEMIC) action.

ISA² solutions can be used free of charge and are open source when related to IT.

More on the programme

ec.europa.eu/isa2

Contact ISA²

isa2@ec.europa.eu

Follow us

@EU_ISA2
@Joinup_eu

isa2 programme