Virtual Assistant - Research Report

Comparative Study 1 - Alexa - Skill development

1.Introduction: Alexa skills are voice-activated applications that allow users to interact with Alexa - Amazon's virtual assistant through voice commands. Alexa skills provide a new and innovative way for users to interact with technology, and offer developers a unique platform for creating voice-enabled applications. These skills are similar to mobile apps, but are designed to be used on voice-enabled devices such as Amazon Echo, Echo Dot, Echo Show, and other devices that have Alexa built-in.
Alexa skills can be developed by third-party developers or by Amazon itself. There are thousands of Alexa skills available, covering a wide range of categories such as news, weather, music, games, home automation, and more. Users can enable and disable Alexa skills through the Alexa app, and can also customize which skills Alexa uses by default for certain tasks. Developers can create Alexa skills using the Alexa Skills Kit (ASK), which provides a set of APIs, tools, and documentation for building voice-enabled applications.

2.Vendors chosen for Comparative Study: Voice Apps, XAPPmedia, and Bespoken, considered here are companies that specialize in creating Alexa skills and providing related services. However, there are some differences between them in terms of their offerings and target markets.

Voice Apps:Voice Apps is a company that specializes in creating Alexa skills for businesses and organizations. They offer a range of services, including skill design, development, and hosting. Voice Apps has created over 7,000 Alexa skills for clients across various industries. They have a focus on providing an easy-to-use platform for clients to create and launch their own skills

XAPPMedia: XAPPmedia, is a voice-activated solutions provider that creates Alexa skills for media companies, brands, and organizations. They offer a platform for designing, building, and launching custom Alexa skills. XAPPmedia has created over 1,500 Alexa skills for clients, including National Public Radio and Radio Disney. They have a focus on creating interactive and engaging voice experiences for their clients.

Bespoken is a company that provides testing and monitoring services for Alexa skills. They also offer a platform for building and deploying Alexa skills. Bespoken has created over 1,000 Alexa skills for clients, including Uber and Mastercard. They have a focus on providing high-quality testing and monitoring services to ensure that Alexa skills are reliable and perform well for end-users.

3.Comparative criteria: A comparative review is performed on the three vendors on the **type of skills offered** (Result 1) by the companies to identify the diversity of the services and the target audience of the vendors. The **alexa skill statistics** (Result 3) reflects the versatility and expertise of the vendors in designing and developing voice applications. The (Result 2) mentions the perks the vendors cater to their target audience. While (Result 4) **types of services** offered brief the high level services the vendors mention in their portfolios. Parallely (Results 5) **features criteria** explores the fine grained aspects that are included in the packages/plans which are considered during the skill building and pricing.

4. Results: The results of the 5 criterias are tabulated below,

Criteria	Voice Apps	XAPPmedia	Bespoken
1. Type of skills covered	There are a variety of skills built with voice apps from music, health, entertainment, education, games and flash briefings. One example being, 1. Guard dog - This app Guard Dog can help deter criminals by playing dog barking, growling, and snarling sounds while you're away. Burglars will likely avoid your home if they hear angry dogs inside. To have Guard Dog guard, "Alexa, tell Guard Dog to Guard" is the wake word.	XAPPMedia offers a range of Alexa skills from Voice commerce, podcast skills, Radio and news skills. Instance of an app, 1. Interactive Audio Ads: - This new amazon feature builds on the limited voice commands to buy something or ask for more information incorporated into Amazon Music ads. User's can hear the ad to buying the product and complete the transaction with "Alexa add this to cart" command	Bespoken covers a wide variety of Alexa skills from rentals, hotel bookings, Vehicle maintenance, smart home e-Commerce. 1. Mercedes-Benz: Bespoken helped Mercedes Benz R&D deliver exceptional quality for its connected car voice interaction experiences. 2. The Mars Agency : The Mars Agency cut down errors in the speech recognition of their voice app by more than 80% before launching thanks to our software and tech support.
2.Advantages	The upsides of using VoiceApps for Alexa skill development are, 1) Enhanced User Experience: Voice apps allow for a more natural and	The benefits of using VoiceApps for Alexa skill development are, 1) Provision to create personalized and	The advantages to using Bespoken for Alexa skill development, 1) Providing comprehensive tool for

	intuitive interaction between users and Alexa 2) Affordable Pricing 3) Improved Accessibility: Voice apps can make Alexa skills more accessible to users with disabilities or limited mobility, as they provide an alternative to traditional user interfaces.	interactive voice experiences 2) Comprehensive service to design, develop, and launch the skills 3) Options for clients to measure the performance of their alexa skill.	developing, testing and managing alexa skills 2) Bespoken offers a high degree of customization and flexibility, allowing businesses to create Alexa skills that are tailored to their unique requirements 3) Ongoing Support and Maintenance
3.Alexa skill statistics	Voice Apps has created over 7,000 Alexa skills for clients across various industries.	XAPPmedia has created over 1,500 Alexa skills for clients, including National Public Radio and Radio Disney.	Bespoken has created over 1,000 Alexa skills for clients, including Uber and Mastercard.
4.Types of services offered	Range of services, including integration of Alexa skills with API's , Testing, maintenance and support of Alexa skills and provide voice analytics with insights.	Platform for designing, building, and launching custom Alexa skills	Provides provision for testing and monitoring services for Alexa skills
5.Features	Voice apps offers plethora of features for the companies to choose from while developing the alexa skill, Skill Building: Features like getting data from third parties, using standard audio player, linking accounts, responding to users in multiple voices are few features under skill building that can be customized based on the chosen plan while creating Alexa skills. Support: Live chat support, self guided help feature, email support , customizing response time are some features offered	XAPPmedia offers a simple plug and play setup that requires no code and takes just minutes for the assistant to quickly start giving automated, highly relevant and accurate answers to questions. The features offered are not clearly mentioned explicitly on their website and can be made available upon request.	Bespoken has features such as, Development and Testing: command-line interface that enables developers to automate the testing and deployment of their Alexa skills , Bespoken Virtual Device to test their Alexa skills on any device with a web browser , Bespoken Debugger tool that enables developers to debug their Alexa skills in real-time.

5.Discussion: The study demonstrates the ability of vendors to understand their clients' needs and deliver high-quality voice solutions that cater to their target audience. It also reflects the potential for growth and innovation in the voice technology industry, as more businesses look to incorporate voice applications into their digital strategies.

Comparative Study 2 - Open-Source Voice Assistants

1.Introduction: Open-source voice assistants are software programs that allow people to communicate with devices through voice commands. These voice assistants are built using open-source technologies and aim to provide users with a more natural and intuitive way to interact with technology. Some examples of open-source voice assistants include Mycroft, Aimybox, and Rhasspy. These voice assistants come with a variety of features and capabilities and can be customized to meet the specific needs of users. One of the benefits of open-source voice assistants is that they are transparent, meaning users can inspect the source code to ensure their privacy and data security.

2.OpenSource Voice Assistants chosen for Comparative Study: Rhasspy, Jasper, and Aimybox are open-source software projects that enable natural language processing (NLP) and voice assistance.

Rhasspy:

Rhasspy is an open source, fully offline set of voice assistant services for about 16 languages. It enables users to communicate with their devices using natural language and is constructed using the Python programming language. Rhasspy integrates different open-source libraries and frameworks, including Kaldi, for speech recognition and natural language processing.

Rhasspy has several benefits, including the ability to operate on low-power devices such as Raspberry Pi, which makes it convenient to establish an affordable voice-controlled smart home or personal assistant. Moreover, it supports various languages, allowing users to communicate with it in their preferred language. Rhasspy also provides text-to-speech capabilities, enabling it to articulate responses to the user.

Jasper:

Jasper is an open source platform that permits developers to develop voice-controlled applications. It empowers developers to create personalized voice commands and automate tasks on their computers using natural language. Jasper is constructed on Python and incorporates several APIs and libraries, including the Google Speech Recognition API. Jasper can be integrated with various online services, such as Wolfram Alpha and Google Calendar.

One of the strengths of Jasper is its extensibility. It has a plugin architecture that allows developers to easily create new functionality for the assistant.

Aimybox:

Aimybox is a platform that enables developers to generate voice-controlled applications and virtual assistants. It relies on the idea of "conversational AI" and furnishes developers with a suite of tools and libraries to facilitate the integration of natural language understanding and generation features into their applications.

Aimybox furnishes an easy-to-utilize software development kit (SDK) that can be incorporated into several platforms, including mobile, web, and smart home devices. It can process multiple languages and employs sophisticated natural language processing approaches to comprehend and reply to user commands. Aimybox also permits developers to tailor the conversation flow, and it supports integration with various services like Google Assistant, Amazon Alexa, and Microsoft Cortana.

3.Comparative criteria: The (Result 1) **intended users** briefs about the extent of technical expertise required to adopt the open source. The **significant features** (Result 2) mentions the unique aspects of the open source voice assistant in terms of the extendability, ease of use, customizability as features.

While (Result 3) **Security** and (Result 4) **Privacy** factors are key trust factors for voice assistants. This section points out the potential threats in the respective open source platforms. Having a robust and accessible development platform is crucial for open source projects hence (Result 5) the **platform criteria** is used as qualitative analysis to measure the ease of adoption by users and enable collaboration.

4.Results:

The table below weighs up various criterias to perform a comparative study on the three open source voice assistant,

Criteria	Rhasspy	Jasper	Aimybox
1.Intended users	**Rhasspy** is intended for savvy amateurs or advanced users who intend to have a private voice interface to their chosen home automation software.	**Jasper** is claimed as an effective tool for developers, makers and DIY enthusiasts to create custom voice commands and automate tasks on their computers using natural language. The	**Aimybox** is specifically designed for implementation in Mobile applications, supporting both Android and iOs platforms adopted by Mobile Application Developers.
2.Significant features	1) Optimized to work with MQTT, HTTP and WebSockets 2) Rhasspy itself is licensed under the MIT license 3) Enhanced to handle uncommon words and pronunciations	1) Multiple language support 2) Presents itself with rich documentation 3) Developer-friendly API to build applications 4) Large community of developers contributing to the project	1) Fully customizable and extendable, you can connect any other speech-to-text, text-to-speech and NLU services 2) Voice skills logic and complexity is not limited by any restrictions 3) Can interact with any local device services and local networks
3.Security	Rhasspy the voice assistant unlike other voice assistants that need an internet connection to operate, Rhasspy can operate locally on a device, ensuring more security and privacy.	Open source software is prone to security issues due to lack of updates and maintenance, malware and virus attacks can result in unauthorized access to user's files/social media accounts. Jasper can be integrated	Aimybox, mainly designed for Mobile applications, may store sensitive user data in an insecure manner, such as storing user passwords in plain text. Similarly Mobile applications can be targeted by malware and phishing attacks, which

		with other third-party applications, which may have their own security vulnerabilities.	can result in the user's sensitive information being compromised.
4.Privacy	Rhasspy the voice assistant unlike other voice assistants that need an internet connection to operate, Rhasspy can operate locally on a device, ensuring more security and privacy.	Jasper emphasizes the "Control anything" and "Always listening" features which are identified as privacy concerns. Jasper has the capability to use voice commands from users to automate tasks on the computer, access email accounts, and update social media networks which have direct access to personal files / social media accounts.	The voice assistant on the mobile device may collect and process sensitive personal information, such as location data, contacts, and personal messages. As well as vulnerable to data breaches, which can result in sensitive user data being exposed.
5.Platform	Rhasspy is built on Python and uses Kaldi and Pocketsphinx, open source toolkits for Speech recognition.	Jasper is an open source Python-based platform used for creating voice-controlled applications.Jasper's design is specifically tailored for the Raspberry Pi and Linux.	Aimybox is an Open source under Apache 2.0, written in pure Kotlin.

5.Discussion:

The interesting results from the comparative study is that the choice of open-source platform has to be specific to the use case and requirements of the project. Jasper is a good choice for DIY voice assistant project whilst privacy focussed Rhasspy is an excellent offline voice assistant toolkit and Aimybox is highly customizable and modular.

Comparative Study 3b - Knowledge Graphs QA datasets

1.Introduction: KGQA comprises answering a Natural Language Question from using the knowledge facts present in a given RDF based Knowledge Graphs (KGs) such as Wikidata, Freebase, DBpedia, etc. KGQA systems use various techniques such as entity recognition, relation extraction, and semantic parsing to understand the meaning of natural language questions. They then use this understanding to identify the relevant nodes and relationships in the knowledge graph and retrieve the necessary information to generate answers.

2.KGQA Datasets chosen for Comparative Study: KGQA is an important area of research in natural language processing. The three datasets under comparative study are,

WebQuestionsSP: Created using Google Suggest API and Amazon Mechanical Turks the WebQuestions dataset is a collection of questions and answers. The dataset includes 6,642 question-answer pairs and utilizes Freebase as its knowledge base. The train and test split of the dataset is 3,778 data points used for training and 2,032 used for testing. Fig 1 - illustrates the structure of the questions in the dataset.


```
{
  "QuestionId": "WebQTest-290",
  "RawQuestion": "what do christians believe about heaven hell and purgatory?",
  "ProcessedQuestion": "what do christians believe about heaven hell and purgatory",
  "Parses": [
    {
      "ParseId": "WebQTest-290.P0",
      "AnnotatorId": 1,
      "AnnotatorComment": {
        "ParseQuality": "Complete",
        "QuestionQuality": "BetterAnsweredByDescription",
        "Confidence": "Normal",
        "FreeFormComment": "D "
      },
      "Sparql": "PREFIX ns: <http://rdf.freebase.com/ns/>\nSELECT DISTINCT ?x\nWHERE {\nFILTER
       '' OR langMatches(lang(?x), 'en'))\nns:m.01lp8 ns:religion.religion.beliefs ?x .\n}\n",
      "PotentialTopicEntityMention": "christians",
      "TopicEntityName": "Christianity",
      "TopicEntityMid": "m.01lp8",
      "InferentialChain": [
        "religion.religion.beliefs"
      ],
      "Constraints": [],
      "Time": null,
      "Order": null,
      "Answers": [
        {
          "AnswerType": "Entity",
          "AnswerArgument": "m.02wvcg4",
```


Fig 1: Sample Question structure from WebQuestionsSP

About the WebQuestionsSP dataset: Every datapoint in the dataset will have a unique questionid, Raw questions(the questions as they were originally submitted by users on the Google search engine) and Processed questions (questions that have been cleaned up and corrected by human annotators to remove errors and make them more understandable). The datapoint also has the SPARQL query to access the Freebase KB, annotation specific information and potential entity name for Entity resolution.

LC-QuaD 2.0: Linked Data-based Complex Question Answering Dataset is a Large Question Answering dataset with 30,000 pairs of questions and its corresponding SPARQL query. The target knowledge base is Wikidata and DBpedia, specifically the 2018 version. Fig 2: shows a sample data point from LC- QuaD dataset.


```
{
  "NNQT_question":"Did {Alexander_Hamilton} {occupation} {lawyer}?",
  "answer":[

  ],
  "paraphrased_question":"Did Alexander Hamilton practice law?",
  "question":"Is Alexander Hamilton a lawyer?",
  "sparql_dbpedia18":" ASK {
?statement1
 <http://www.w3.org/1999/02/22-rdf-syntax-ns#subject>
<http://wikidata.dbpedia.org/resource/Q178903> .
?statement1
 <http://www.w3.org/1999/02/22-rdf-syntax-ns#predicate>
 <http://www.wikidata.org/entity/P106> .
?statement1
<http://www.w3.org/1999/02/22-rdf-syntax-ns#object> <http://wikidata.dbpedia.org/resource/Q40348>. } ",
  "sparql_wikidata":"ASK WHERE { wd:Q178903 wdt:P106 wd:Q40348 }",
  "subgraph":"boolean one_hop right subgraph",
  "template":"Ask (ent-pred-obj)",
  "template_id":1,
  "template_index":0,
  "uid":0
```


Fig 2: Sample Question structure from LC QuaD 2.0

About the LC Quad dataset: LC Quad provides support for NNQT (Non-Native Question Type) questions, every data point in the dataset has a NNQT Question, a paraphrased question and the actual question. A corresponding dbpedia and wikidata SPARQL query is also included in the dataset. The question also has fields such as a subgraph is a connected subgraph of a knowledge graph that is constructed based on the entities and relationships relevant to a given question.

The subgraph information is used to represent the data that needs to be searched in order to answer the question. A template ID in LC-QuAD is a unique identifier that is assigned to each template. This ID is used to link a specific question to the appropriate template and to generate the corresponding subgraph for that question

QALD: Comprising multiple sub-tasks, such as multilingual question answering and temporal question answering, the QALD dataset serves as a standard benchmark for assessing the performance of QA systems operating on linked data. It contains natural language questions and their corresponding SPARQL queries that can be used to query linked data resources.. QALD has been utilized to compare and evaluate various question answering systems, including those that employ techniques based on information retrieval, semantic parsing, and machine learning. Fig 3- shows a sample datapoint from the QALD dataset.


```
"questions": [
    {
        "id": "99",
        "question": [
            {
                "language": "en",
                "string": "What is the time zone of Salt Lake City?"
            },
            {
                "language": "ba",
                "string": "Миндей вакыт поясы Солт-Лейк-Ситила"
            }
        ],
        "query": {
            "sparql": "PREFIX res: <http://dbpedia.org/resource/>
            PREFIX dbp: <http://dbpedia.org/property/>
            SELECT DISTINCT ?uri WHERE { res:Salt_Lake_City <http://dbpedia.org/ontology/timeZone> ?uri }"
        },
        "answers": [
            {
                "head": {
                    "link": [],
                    "vars": [
                        "uri"
                    ]
                },
                "results": {
                    "bindings": [
                        {
                            "uri": {
                                "type": "uri",
                                "value": "http://dbpedia.org/resource/Mountain_Time_Zone"
                            }
                        }
```


Fig 3: Sample Question structure from QLAD

About the QLAD dataset: Every datapoint in the dataset has a unique id for the question, the language it is presented in along with the question, corresponding SPARQL query for the intended Knowledge base. The answers are mapped to their corresponding resource entry in

the knowledge graph. Compared to the other two datasets this dataset looks more refined and exclusive.

3.Comparative Criteria: The foremost criteria compared are the (Result 1) **dataset size**, **languages** it is available in and the knowledge base (KB) it supports. Language and dataset size are important factors for datasets as they directly impact the quality of the models it is tested upon. The **Question type** (Result 2) is compared as a criteria as the complexity of datasets determines the generalizability of the KGQA models so the models do not overfit to simple or limited knowledge graphs and can generalize well to unseen or more complex data. The **Questions created using criteria** (Result 3) is crucial to determine the naturalness of the question which increases the hits to the Knowledge base. The **Created by and purpose of creation** (Result 4) helps us to determine the intent and the use case the dataset is created for.

While (Result 5) gives a comprehensive justification of the **level of generalizability** of the datasets.

4. Results:

Criteria	WebQuestionsSP (2016)	LC-QUAD 2.0 (2019)	QALD 9+
1(a) Knowledge base	Freebase	Wikidata	Dbpedia, Wikidata
1(b) DatasetSize	5810	30,000	500
1(c) Languages	1	1	9
2. Question Type	Simple	Complex Complexity Level : 1) Multi-fact questions 2) Temporal questions 3) Questions that utilize qualifier information	Complex
3. Questions created using	Google Suggest API The Freebase database	The dataset is created using a multi-stage process involving	The authors used QUANT, a Question Answering curation

	is queried using Lambda Dependency-Based Compositional Semantics framework. The authors use an almost fully automated approach that maps natural language phrases to logical predicates. Finally, logical predicates that match the neighboring predicates are captured using bridging mechanisms.	template generation and paraphrasing. First, a set of entities are obtained based on Wikipedia Vital Articles. A set of SPARQL queries are manually created that satisfy a large number of intentions from the human perspective. This is followed by fetching a set of predicate entities that satisfy them. A subgraph of subject, predicate, objects and qualifiers is created with the collected entities that fit to the SPARQL query. They are then used to generate natural language template based queries. Once this standard form is obtained, they are crowd sourced through **Amazon Mechanical Turks** for labellers to verbalized the NL template based queries, followed by the second step of generating paraphrases for them (again achieved by the labellers)	interface that is used to speed-up the process of updating new QA pairs from the DBPedia RDF dataset. The questions are then manually annotated (SPARQL queries) to generate both the train and test sets. 17 volunteer participants and 290 crowd workers from Amazon Mechanical Turk4 and Yandex Toloka5 translated questions into 8 languages, such as: German, Russian, French, Armenian, Belarusian, Lithuanian, Bashkir, and Ukrainian, to create the QALD 9+ dataset (+ indicating an iterative improvement over QALD by multilingual translation by crowd workers)
4.Created by and purpose of	The datasets are created by Researchers at	The dataset was created by a group of	The dataset was created by researchers from

creation	Stanford University Jonathan Berant Andrew et. al.	researchers from the University of Bonn Germany and their work was supported by the Fraunhofer-Cluster of Excellence “Cognitive Internet Technologies” (CCIT).	France and Germany (University of Hamburg, University of Applied Sciences Kothen and Universite de Lyon)
5(a) I.I.D Generalizability	Yes	Yes	Yes
5(b) Compositional Generalizability	No	Yes	Yes
5(c) Zero-shot Generalizability	No	Yes	**Yes***

*According to the given reference paper, the datasets can also be compared to each other based on a “generalizability factor”. Every knowledge base has a set of Relations, Entities, Blank Nodes, Classes and non-unique Literals. Using these components, and a set of constraints on them it is possible to divide the datasets into several categories based on their generalization capability. They are categorized into the following types based on the test set:

1. **I.I.D generalization (Independent and Identically Distributed):** In this case all the questions in the test set follow the same distribution of the train set in terms of the schema terms (classes, relations and join operations), however may differ in terms of actual entities and literals.
2. **Compositional generalization:** This form of generalization adds further complexity to the previous form. Here, the structure of the query in the test set differs from the train set (compositional diversity).
3. **Zero-shot generalization:** This form states that for all questions in the test dataset there exists at least one schema item that is not seen in the train set (which could be a relation, entity or operation). This is the highest level of complexity for the testset, with respect to which the model's generalization capability can be tested.

The authors also provide a mechanism for identifying if a dataset can be splitted into train and test splits that allow any form of generalization as listed above. The algorithm work as follows:

Given a question q from the test set, a unique set of schema terms formed on the train set, the aim is to determine the level of generalizability of the question q w.r.t to points 1,2 and 3 discussed above. First the schema terms from the question are extracted. If they correspond to an entry in the dictionary of the set of unique scheme terms in the train datasets (each row corresponds to schema terms in a question where rows are unique), then it is considered an I.I.D generalization. In the other case, if all schema terms lie within the train set scheme terms they are considered to be Compositionally Generalizable. Finally if there exists at least one schema term in the test question that is not there in the unique set of schema terms in the entire train set, it's considered to have the highest level of complexity (zero-shot) generalization.

Using this idea the authors further classify the above three datasets based on whether it is possible to achieve these forms of generalizability on their test sets. As it can be observed for the **WebQuestionsSP** dataset, it is I.I.D generalizable but not compositional or zero-shot generalizable. On the other hand both the **LC-QUAD 2.0** dataset and the **QALD-9** dataset is I.I.D, Compositional and Zero-shot generalizable, making them **more suitable options** for **evaluating KGQA systems**.

5. Discussion:From our study it can be observed that LCQuad 2.0 is a well-rounded dataset for KGQA model assessment, that has significantly more data points from various complexity levels, in addition to ensuring I.I.D, zero-shot and compositional generalizability. However, the dataset does not support multilingual KBs. In such scenarios a good option would be to choose a QALD 9+ dataset, that supports over 8 languages while meeting all the three complexity levels.

Comparative Study 4 - Chatbot platforms

1.Introduction: Designed to emanate human conversations as text or audio, chatbots provide a great way for businesses to provide quick and effective solutions to common customer queries and problems. With the increasing popularity of chatbots in various industries, including customer service, healthcare, and finance, there is a growing demand for more intelligent and sophisticated chatbot systems. Chatbot platforms are software tools or frameworks that provide developers and businesses with the necessary tools to create, deploy, and manage chatbots.

2.Chatbot platforms chosen for comparative study: The chatbots chosen for comparative study are,

Chat fuel: This simple easy-to-use chatbot builder platform is specifically designed for Instagram and Facebook messenger. Chatfuel linked to the desired page of the business account creates automated responses to common keywords from users. The bot uses NLP to understand and recognize phrases, and send users predefined answers. The automated responses can be simple text responses, or can be configured as custom automated flows.

The intended users for Chatfuel includes entities such as e-commerce enterprises, customer support and marketing teams, and any other organization seeking to enhance their customer engagement by automating their customer interactions.

Mobile monkey: Mobile Monkey is designed to cater to both novice and technical users, and provides advanced features such as broadcast messages (messages sent to multiple users, whether all users or a specific segment), drip campaigns, chat blasts to boost a business's marketing strategies. The platform provides a drag-and-drop interface for building chatbots, making it easy for users with no coding experience to create conversational experiences.

Aivo: Aivo Suite's automated conversational agents bots offer the possibility to design a service experience in a simple and no-code way, understand the behavior of your customers, solve queries instantly and increase the NPS using conversational artificial intelligence. The target audience of Aivo chatbot are e-commerce, banking, healthcare, and telecommunication domain.

3. Comparative Criteria: The **significant features** (Result 1) of the chatbot platforms elucidates the special features the platforms offer in terms of the services, performance factors, customer support and likewise. With (Result 2) listing the **reputable clients** as the users of the chatbot platform, the trustworthiness and experience of the provider in the market can be compared.

The **predefined template criteria** (Result 3) explains the versatility the platforms offer for its customers. The templates are designed to help businesses quickly create chatbots that can engage with customers and automate certain tasks, such as answering common questions, scheduling appointments, or processing orders.

Pricing (Result 4) is included as a criteria for qualitative analysis. The list of messaging **channel supported** (Result 5) allows us to compare the seamless experience across different channels supported by the chatbot platforms

4.Results:

Criteria	Chat Fuel	Mobile Monkey	Aivo

1.Significant features	1. Features like broadcast messaging 2. Audience segmentation through CRM, and analytics 3. Chatfuel also offers automation of customer support, and improves sales from messenger, Facebook pages, comments and ads 4. Chatfuel website Claims 98% of the customer inquiries are handled by chatbot	1. Automating sales outreach can be beneficial for businesses in optimizing their workflow and enhancing customer interaction through the provision of useful information on the behavior and interests of potential customers. 2. MobileMonkey uses multi-channel chatbots. That means the web chat you create works in more than one messaging app. 3.Ideal for mobile messenger marketing	1. Video Conversational AI: Aivo offers a distinctive customer service experience by providing a hyper-realistic avatar generated through AI technology. This avatar can interact with customers in a nearly-human manner and help resolve their queries instantly. By utilizing a virtual video assistant, customers can chat naturally with the avatar and receive personalized assistance. 2. Aivo offers a stellar 90% accuracy in conversations handling multiple languages, naturalness and open dialogues.
2.Renowned clients	Lego Adidas Netflix British Airways Visa Nissan Levi's Source	Toyota Ford Holiday Inn	IBM Subway Virgin mobile Visa Movistar picPay Banco Azteca Source

3.Predefined templates	Chatfuel offers an array of predefined templates for eCommerce , Real estate, Restaurant Beauty, Sales , Healthcare Insurance, Travel, Hotel	Facebook messenger templates	Aivo provides customized conversational bots for the domains such as Retail and E-commerce Fintech ,Banking ,Telecom
4.Pricing	The pricing for chatfuel mentioned in their website is, For facebook and instagram, Entrepreneur - 14.99 $ with 500 conversations / Month Startup - 24.99 $ with 1000 conversations / Month Small Business - 59.99 $ per month with 3000 conversations / month Growing business - 119.99 $ per month with 10000 conversation / month For Whatsapp business version priced at 49 $ and enterprise edition at 300 $	Mobile monkey presents itself with 3 products namely, **InstaChamp** (free-$20/month) that is aimed at social media creators looking to grow their audience **FunnelPro** (from $120/month) is aimed at marketers looking to "fill their sales pipeline" **RoboBDR** (from $120/month)	The pricing for Aivo is based upon the starter, business, advanced package. Starter - 99 $ per month with support to all messaging channels, KB management, analytics, chat support Advanced - 399 $ per month with SMS, whatsapp support, Zendesk integration, satisfaction management Custom - with multi bot option, multi language, live chat and phone support, salesforce integration
5.Supported Platforms	Facebook Messenger Instagram Website Whatsapp	Facebook Messenger SMS Native web chat Twitter LinkedIn	Instagram WhatsApp

5.Discussion:

The interesting results from the study are the advanced features the platforms cater to its audience for instance, user segmentation features by Chatfuel allows users to segment their audience based on their interactions and behavior, and personalize responses accordingly. Similarly MobileMonkey's Chat blasting feature allows users to send messages to a group of subscribers on multiple channels simultaneously. This feature can be used for promotional messages or to notify customers about new products or services.

On the other hand, Aivo's video conversational AI has the capability to offer a personalized and captivating customer experience. Through video chat, customers can observe and listen to the chatbot, which can help create a sense of trust and connection. This feature is particularly significant for businesses that aim to provide a customized customer experience, such as those in the healthcare or financial industries.

Comparative Study 5 - Data-to-Text Datasets and Evaluations

1.Introduction:Data-to-Text Generation (D2T NLG) can be described as Natural Language Generation from structured input sources such as tables, graphs, databases, and other types of structured data formats. This technology is beneficial for various applications, including automatic report generation, data journalism, and business intelligence. Data-to-text datasets are critical for advancing the field of NLG by providing a means for training and evaluating NLG models that can automatically generate coherent and fluent text descriptions based on

structured data inputs.Datasets for data-to-text generation typically focus either on multi-domain, single-sentence generation or on single-domain, long-form generation.

2.Data-to-Text datasets chosen for Comparative study: The dataset chosen for comparative study are MultiWOZ (Multi-domain Wizard-of-Oz), Robocup and WikitableT.

MultiWOZ (Multi-domain Wizard-of-Oz):Multi-Domain Wizard-of-Oz dataset (MultiWOZ) contains fully-labeled written conversations between humans on various topics and domains. This dataset is introduced to aid the dialogue research community with the limited availability of data has hindered significant progress. This dataset is at least ten times larger than any previously annotated task-oriented corpora, consisting of k dialogues. This dataset presents itself as an open-source dataset with labeled dialogue belief states and actions. The dataset has the data collection process and its analysis, which is entirely crowd-sourced without hiring professional annotators. It also has benchmark results for belief tracking, dialogue act, and response generation, demonstrating the dataset's usability and setting a baseline for future studies. An instance of datapoint from the hotel dataset of MultiWOZ as shown in Fig:4 presents with hotel name, address, area, location id, price, star rating of a hotel.


```
{
    "address": "124 tenison road",
    "area": "east",
    "internet": "yes",
    "parking": "no",
    "id": "0",
    "location": [
        52.1963733,
        0.1987426
    ],
    "name": "a and b guest house",
    "phone": "01223315702",
    "postcode": "cb12dp",
    "price": {
        "double": "70",
        "family": "90",
        "single": "50"
    },
    "pricerange": "moderate",
    "stars": "4",
    "takesbookings": "yes",
    "type": "guesthouse"
},
```


Fig 4: Sample datapoint from Hotel MultiWOZ dataset

Robocup: RoboCup is an initiative in which research groups compete by enabling their robots to play football matches. Playing football requires solving several challenging tasks, such as

vision, motion, and team coordination. Framing the research efforts onto football attracts public interest (and potential research funding) in robotics, which may otherwise be less entertaining to non-experts.

WikitableT: Wikitablet presents generating Wikipedia sections as a data-to-text generation task and creates a large-scale dataset, WikiTableT, that pairs Wikipedia sections with their corresponding tabular data and various metadata. WikiTableT contains millions of instances, covering a broad range of topics, as well as a variety of flavors of generation tasks with different levels of flexibility. This dataset has benchmarked training and decoding strategies on WikiTableT. A sample datapoint from the dataset is shown in Fig 5 which has document title from where the data is derived, tokenized text, document title about the topic from wikipedia articles.

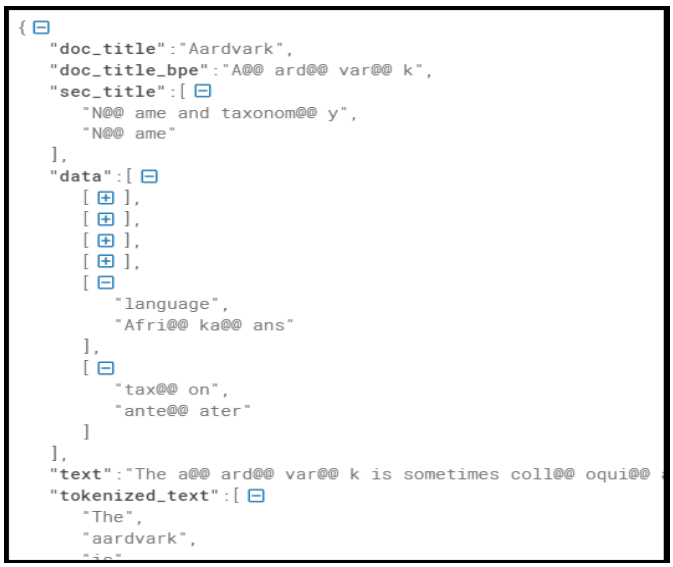

Fig5: Sample datapoint from WikitableT

3. Comparative Criteria: The criterias considered for comparison are,

Datatype (Result 1) is crucial for comparing datasets effectively and selecting the appropriate techniques and tools for processing and analyzing the data. While **Task type and intended audience** (Result 2) is an important factor when comparing datasets because different tasks require different types of data. Multiwoz focuses on task-oriented dialogue systems, Robocup on robotics research and Wikitable on general topics so the research community can benefit

from the task type details. **Domain focus** (Result 3) of a dataset is important for selecting the appropriate dataset for a given research question or problem, ensuring the quality of the data, improving interpretation and analysis, and addressing ethical considerations. **Evaluation Metric** (Result 4) is important for comparison in selecting the appropriate dataset and analytical methods, and for interpreting the results of the analysis accurately. The **data complexity** (Result 5) , complex datasets may require more sophisticated visualization techniques determining the resource allocation, method / model parameter selection , performance evaluation.

4. Results: The results of the criterias compared for the data-to-text datasets are,

Criteria	MultiWOZ	Robocup	WikitableT
Data type	MultiWOZ uses natural language text as their input and output. MultiWOZ uses human-human written conversations as input.Each dialogue consists of a goal, multiple user and system utterances as well as a belief state	RoboCup uses structured game state data as its input.	WikiTableT uses natural language text as their input and output. WikiTableT pairs Wikipedia sections with their corresponding tabular data and metadata
Task type and intended audience	MultiWOZ is designed for task-oriented dialogue systems. MultiWOZ is intended for developers of task-oriented dialogue systems	RoboCup is designed for robotics research and development of intelligent agents for game playing. RoboCup is intended for researchers working on robotics and artificial intelligence	WikiTableT is designed for data-to-text generation.WikiTableT is intended for researchers working on data-to-text generation

Domain focus	MultiWOZ covers multiple domains such as hotels, restaurants, and transportation	RoboCup focuses specifically on sports domain exclusively on football (soccer games)	WikiTableT covers a broad range of topics since it is based on Wikipedia articles.
Evaluation Metrics	MultiWOZ are typically evaluated using metrics such as BLEU and ROUGE	RoboCup is evaluated based on the performance of the robots in the football games	WikiTableT are evaluated using metrics such as BLEU and ROUGE
Data complexity	MultiWOZ require generating complex natural language text based on structured data	RoboCup requires making real-time decisions and actions in a dynamic environment, based on structured game state data.	WikiTableT require generating complex natural language text based on structured data

5.Discussions:The interesting similarity between RoboCup, Wikidata, MultiWOZ, and data-to-text datasets during comparative study is that they all involve the use of data to generate natural language text. In all cases, the challenge is to transform structured data into coherent and understandable natural language text. This requires the use of natural language generation techniques, which involve selecting appropriate language patterns and generating text that is both informative and grammatically correct.