Fundamentals

*This document is intended to provide additional fundamentals to the proposed Cezanne-ai project and the research paper: “Cezanne-ai: a conversational AI open-framework for multi-domains, all the languages and limited data”.*

1. Objectives:

Give a fundamental perspective/overview on NLP (not only to conversational AI) & conversational AGI, present and future, using knowledge both from linguistics (spoken and literary) and Machine Learning/DL/Reinforcement Learning algorithms. Some of the topics addressed:

1. Why is it important for the AI industry to find efficient solutions and models that cover not only the most used 5-10 languages and the developed countries? Even if, Google and Facebook provided solutions for multi-language models (mBERT and XLM-RoBERTa) the fact that these models come with unbalanced datasets for different languages it will translate in poorer results and pretty much are classifying the countries in first degree and second-degree countries… and this will translate in greater economic disparities in the future.
2. Why is it a bad idea to generalize and not to particularize in terms of language, culture or even on a local level: the specificities of big and average size cities for example?
3. How to provide machines with the education and the right context in order to be involved actively in the conversation and to build them in terms of responsibility and ethics?

We will mainly target fundamentals in Conversational AI that had without a doubt a breakthrough in the last years and, even if it is far away from other AI/NLP domains that are using Machine/Deep Learning, we need to acknowledge the important steps in having a natural conversation with a robot. But multi-turn conversational AI (one-turn can hardly be considered a real conversation) is still an emerging research topic, to put it politely. The statement that is behind other domains like image recognition or machine translation is based on obvious evaluation that machines do a better job than humans in some areas of artificial intelligence and not on statistics that are not in the scope of this document[[1]](#footnote-1). Furthermore, we don’t have benchmarks like Glue, Squad or F1 as we have on other NLP tasks, so is hard to proper evaluate conversational AI models (we will not refer at chitchats bot that have some benchmarks, as our objective, stated in the research paper, is to go deep in the conversational AI). What is sure, is that expectations were very high for both conversational AI and conversational AGI in the past 5-10 years but is questionable if they are going in the right direction.

We are also characterizing the existing research on NLP/Conversational AI as either idea driven (somebody has a sexy idea that he wants to test and he is writing a research paper on this topic) or business driven (the businesses have needs and NLP might have solutions). We didn’t find a research that is driven by natural conversational principles and methods, a human ethology, or a semiotic/semiology paper correlated with machines capabilities as basis for NLP architectural models. If you read the most important NLP research papers: BERT[[2]](#footnote-2), ERNIE[[3]](#footnote-3) or Attention is all you need[[4]](#footnote-4), you will see that the authors didn’t even bother to provide fundamentals for their models and all they did was to focus on the results. Not to mention that ERNIE authors didn’t even answer questions for us to understand if the model or the datasets are responsible for their results (everybody was happy, however, that everything is open-source now). “Neural Approaches to Conversational AI” by Gao/Galley/Li is probably close to a fundamental paper on Conversational AI, but it is not enough, as they are more focusing on “what it is” and not on “what should be”. Furthermore, they are reducing a dialogue between a human and a bot to Q&As, tasks-oriented and chitchats, limiting the conversation substantially, as we have argued previously.

In terms of AGI, we will talk more about Baum research and the projects surveyed, but there are few initiatives in the conversational domain that are worth mentioning and only one in the top 4 projects is conversational linked (Open AI).

Before going to fundamentals, we will try to answer some important questions:

1. Why is it a good moment for a hybrid Conversational AI/AGI to become the focus of AI industry in a period when 5G’s, cloud services and blockchains are in front?

* For decades people have been using clicks, or GUI, or instructions or coding in order to use the computing capabilities. But in that way, we restrict human access to technology and create even more disparities. Why not let everybody have access by communicating with machines in their natural language and solve their needs by using a human-to-bot conversational method? The discussion over the conversational AI to become the new UI is more present than ever.
* Conversational AI is dominated by Google search, on one hand, and Alexa, Siri and Google Home on the other (together with AI/ML solutions/frameworks/infrastructure provided by Amazon, Google, Microsoft, Facebook and IBM, etc.). The former is mainly an information hub and not a solution-based environment. Every search has millions of possible answers and is becoming challenging to make selections for real people’s needs that are becoming more and more complex by days, not to mention the fact that Google ads are becoming kind-of disturbing (the fact that Google and Bing also have solutions for Q&A - multi-turn, is not so well-known information and these solutions are limited to some languages). The latter comes with solutions for your needs but are not conversational in a real sense. To have a virtual personal assistant, with whom you can have a task-oriented dialogue helps more with time management issues than in a personal/educational/development way. The same emerging market limitations apply also to Alexa/Amazon Lex, Google Home and Siri, not to mention the cultural and spoken language differences that are perceived by many users.
* For ages people are expecting to have a personal E.T. or a robot to interact with on many levels, especially on the emotional one. Even many of us that didn’t see the movie have this core expectation from AI: when will we be able to talk to a non-human? Of course, this implies some risks, but it is easier to control a Conversational bot than the whole internet, especially as the internet is becoming more and more dominated by commercial objectives. There are some countries (5) that have Microsoft XiaoIce, but we had expressed doubts that a robot created mainly for Chinese teenage girls can be well received globally, even if somebody will invest to make it available in other countries/languages.
* A conversational bot can replace operating systems (that are increasingly more difficult to use due to diversity, difficult integrations with other systems and constant deployment of new versions) and also can replace programmers. If you can talk directly to the machine in a conversational and structured way, why need intermediaries? There are two types of developers in IT. The creative ones that are capable of writing codes from the beginning to the end, and the ones that are GitHub/Hugging Face/E2E-models research oriented. In a market where we have a big demand for programmers, a Conversational AI bot can replace the latter category. If a bot can understand everything you are saying in a conversational environment (in natural language and without using a programming code) and at the same time will have access to GitHub database (for example), it will be able to replace some intermediates. It is paradoxical that the first employees that will be impacted by developments in Conversational AI will be the IT experts and not employees in the services industries. But of course, this is not true as it is demonstrated that AI is creating new jobs rather than repressing (2018 WEE forum: 58 million new jobs by 2022). Additionally, once a model becomes efficient, it will have an impact also in the way we operate PC’s/laptops/mobiles as clicks and hardware configurations might become obsolete.
* Everybody has a model or a need of expertise in their life and a possible conversation with this person will be very much appreciated/needed. This model/expert can be Steven Hawking, Einstein, Van Gogh, a CEO, a doctor, a lawyer… and because they are not (always) available through traditional channels of communication, opportunities are created for conversational bots to capture at the same time the personality (machine/deep education) and the experiences (machine/deep learning). Even if these possibilities can make many people uncomfortable, to say the least, we need to acknowledge that our children do not have books in their hands anymore, but smartphones. If these smartphones can’t educate as books do and are only able to provide information or task-oriented solutions, then the lack of conversational AI/AGI bots must concern us more. Books are also regarded as virtual friends in many ways and the selection of the book to read implies the same educational background as the one needed to select a conversational bot to interact.
* Even if in AGI the newest projects are small-to-medium size and governments are not aggressively pursuing AGI for its strategic advantage[[5]](#footnote-5), a hybrid conversational AI/AGI model can change the trends in favor of conversational bots if the presented fundamentals will prevail in front of the marketing/PR approach.

1. Fundamentals:

Some of the fundamentals have correspondence in cybernetics systems theory and we consider the work of Gregory Bateson to be essential for this paper. His holistic view is divergent with the math-based neural networks that is currently used in NLP/Conversational AI. Seeing everything only from one part of the relations can lead to disaster, as Bateson argued, even if math and logic are extremely important for any holistic approach in this field. Although AI will eventually find the path to Human-Level machines, it is impossible to separate humans from the environment in which they live and from the relations they are part of.

Many of the fundamentals are more linked to Conversational AI, but still, some of them apply also to machine translation or other NLP domains. And we will cover in the end also fundamentals on AGI.

*Please do not take into consideration the order (as it is random) and please take into consideration also the fundamentals provided by the research paper “Cezanne-ai: a conversational AI open-framework for multi-domains, all the languages and limited data”.*

* + 1. Accessibility and authenticity in AI

The first intention of this document is to go back to fundamentals. Why are we working/communicating with other people by using emails, messages, meetings, phone calls… and the custom of working with a computer is by clicks, GUI and coding? Is it more efficient this way or we haven’t found yet other ways of making use of the computing power of a CPU/GPU? Continuing on this path can be dangerous because we are creating parallel words that are becoming more and more confusing for people that get stuck in virtual lives and have difficulties to cope with realities. For example, in a war game simulation, the user as a general of his army is giving instructions to kill and destroy not by talking with his suborders and assessing the human implication but by clicks or by other prerequisites. Is this a simulation of a real war or just a sadistic way of making a war interesting and a way of relaxing? Is this an educational method that has positive effects in our current society?

At the same time, people’s needs for knowledge must not be limited and a solution is to make all communication with machines as genuine as possible. The same is with accessibility that is becoming more and more restricted as resources are limited.

* + 1. AI in emerging markets. Limitations & opportunities**[[6]](#footnote-6)**

One of the biggest differences between developed and emerging markets is in the way people communicate. In developed countries they are more comfortable with speaking with a lawyer, or a doctor, or any expert in general because they are familiar with the vocabulary and their countries are more services orientated (we really don’t need references to back this argument). The same applies for bots. For that reason, the communication is more straight-forward (also due to time constraints) and consequently Q&A and task-oriented bots (that are not ‘fundamentally’ conversational bots as we had argued in the research paper) are more useful in these kinds of countries. On the other hand, people in emerging countries search not only for answers but also for consultancy, not to mention that they have their own vocabulary (not adapted yet to expertise domains) and understanding is becoming very tricky.

For example, the first category could have a very direct request in terms of going to dinner: “Alexa, please select me the best sushi restaurant in Western New York and make me a reservation for 3 persons at 5PM”, or many times users would also give instruction related to the name/names of the restaurant. In that way Alexa would have the complete data in one turn conversation. In some instances, it would require two or three turns. The latter category would be for sure more advice-oriented: “Hey bot, can you help me with something? I want to go out with my friends to a special restaurant tomorrow. I think I want to eat some Asian food, but not too traditional …. and I would like to stay outside if it is sunny tomorrow…. what do you recommend?” (those are extracts from our live/current bot). And the conversation could go forever. In more complex domains like legal it can be a nightmare in terms of understanding the situation and the actual needs.

The first impression as developers is that a Conversational bot in an emerging country is impossible to implement. A way to do it is to deploy Alexa/Amazon Lex as it is and ask the user to adapt, maybe also to the language and pronunciation as there are not enough resources (interest) to adapt Alexa to the emerging market environment (at least for now).

But, after analyzing the situation, things might get in favor of conversational bots in emerging markets. Let’s explain why:

The cons for the developed market Conversational AI:

* Even if the request is straightforward, every utterance can be put in a million different ways. So, you will need a huge database to train in order to be sure you got everything the user requested.
* Expectations from the users are high. They will not accept the bot failure especially that the instructions were clear and next time they will not use its services.
* Even if the request/ utterance can be seen as complete/exhaustive, in real life it is not the case. Taking the above example: What kind of sushi, as there are many types and restaurant specializations? In western New York there are many sushi restaurants; which one? Do they have availability, parking spots ….? Many alternatives also mean a huge NLG database that needs to be processed…. Developed can be also seen as complex.

The pros in emerging markets Conversational AI:

* They are more open to conversations, even on a more personal level.
* They have lesser expectations, for now.
* Giving details as much as possible is very helpful even if the specific vocabulary is not accurate.
* They are opened to answer all kind of questions (additional/confirmations)
* The bot doesn’t need to understand everything from the beginning, but to find a way, through conversation and consultancy, to eventually arrive at a common ground with the user.
* Huge databases or complex embeddings are not as essential (this is in scope of the Cezanne-ai project), but the architecture of the conversational model is the key.
  + 1. Conversational doesn’t mean Q&A or tasks (only)**.**

Humans are expressing not only intents, but also having emotional reactions, chitchatting or just giving replies (solicited or not) to an intent or to an answer. Those replies can be ironic, metaphorical or implicit. From our experience with bots, they treat everything as an intent, either they are designed only for replying to an intent, or they don’t have capabilities to understand ironies (for example).

* + 1. All languages are unique (not only in wording)**[[7]](#footnote-7).**

Language isn’t just a modality to express yourself. It incorporates the culture and civilization of a nation, and, most of all, it is a thinking enabler. Additional hypothesis:

* + In today’s globalized civilization, many people speak in one language and think in another, most commonly in their native language.
  + In some languages, the verb is at the end of the sentence facilitating a more elaborate thinking.
  + In other languages you say the important things at the end of the sentence/phrase
  + In others, you have specific motion verbs that are treated differently from other verbs.
  + Some languages have been influenced by many other languages with opposite structure.
  + Other languages are more musical or more metaphorical.
  + There are people that study other languages just to read a book in the native language of the writer and there are voices saying that you cannot fully understand Heidegger or Dostoyevsky in other languages even if you have a perfect translation.
  + Semiotics[[8]](#footnote-8) is arguing that languages are unique not only in wording but also in the structure.

It is highly debatable if you can use the same algorithms/models for all languages. Developing NLP models or language interfaces for a specific language is not only a necessity for preserving that language but it is also a fundamental requirement for building a Conversational bot if you are thinking systematically and on the long term.

* + 1. In order to understand you need to listen**[[9]](#footnote-9)**.

The most important condition for an efficient conversation is to listen. Both agents need to listen to each other and not have preconceived ideas. If the bot uses Markov Decision Processes (MDPs), it will have an unnatural impact on the overall dialogue fluency, even if individual machine answers will be “very” natural in terms of language. Not everything an agent is inputting is an action or an intent (we can argue also that most of the utterances are not exhaustive/complete), and for that reason reinforcement learning framework can translate into a forced topic conversation.

* + 1. Systemic vs Analytic**[[10]](#footnote-10)**

To better understand the “language”, you need to have systemic thinking and, even better, involve people that have this capacity. Mathematicians, programmers, engineers are well known for their analytical background. In math you can break everything into sequences and then build it again using some attention layers (as today NLP models are implemented), but in language (spoken, written; industry based or literary) it is not the case in many instances, and we will give some examples. Furthermore, in IT integration you can use an IT architect to fill this gap of systemic approach and it is the right way to do it. But the question is: can you use an IT architect to build a house or an NLP model with AGI enhancements? Even if the question is formulated in a misleading way and implies a negative answer, NLP means also processing and that makes the answer very tricky.

But let’s make a fundamental point on the systemic matter.

* Many times, the names of a movie/ book/ play/article are not translated accurately, not because the translation is bad, but because the translated name would not be symbolic for the movie itself when confronted with other languages and their cultural implications.
* In many languages you have expressions that cannot be captured, not even by the right semantics of the words. The expression can only be captured as a whole. The same applies to ironies, sentences with “but”, the implications of punctuation over semantic or a simple “no” in a phrase that can completely change a meaning. Btw, the simplest way to test a conversational bot is to make use of noes.
* A book, an article, a phrase … cannot be summarized by a uniform splitting into sequences, even if you keep interferences/attentions (from neural networks). In many cases the important aspects are concentrated in certain parts and you risk missing the point. Division in parts is arbitrary if it is not done by scientists that can find relations (Gregory Bateson).

Yes, a math professor will interrupt you if you start a demonstration erroneously and will correct you. This is the concept of the attention model in NLP. But a language professor would probably let you finish to better understand what you are saying, and this is not the case in many advanced NLP models.

* + 1. Verbal/written communication is less than 10% of the whole communication**[[11]](#footnote-11)**.

As needs, thoughts, emotions are becoming more and more complex, language and human communication needed to find ways to adapt. For this reason, communication doesn’t mean only wording. Punctuations, emoticons, gifs, expressions, jargon, abbreviations are becoming more and more important to capture what someone is saying. At the same time, there are cases when words are implicit in a sentence and capturing those words (that are missing) is important too.

The fact that NLP is using all kinds of processes/clean-ups algorithms to remove from the beginning many parts of what is transmitted is fundamentally wrong. Everything needs to be kept one way or another, because nothing is for granted. The uppercase letters are of importance, also, not to mention that together with the stemming process you can miss the actual intention.

Everything needs to be entered in the encoder- decoder in a structured way, as humans do. This is part of an education process, not a learning process.

* + 1. It is not conversation if you need to repeat everything**[[12]](#footnote-12)**.

A conversation in which you need to repeat yourself all over again is contra-productive. This is the main challenge in AI, and only few solutions were implemented inside NER, LSTM, reasoning and Google Home, summarization-augmentation… but is for sure not enough. MRC models (Machine Reading Comprehension) are the perfect example for making your life impossible. A solution that encodes all interactions and makes reasoning at every step is not only expensive in terms of processing, but the probability of the machine messing up the answers is very high in a real conversation (not a role play). Without a memory update on every interaction based on a fundamental selection, conversational AI could be just a dream. Deep Learning models give hope that this memory update is even possible in the future but if you process without proper understanding, you pretty much cancel the possibility to have future interactions on the same subject/intent.

Argument: if the DL is 70% efficient this also translates into a more pessimistic side of the story: you lose 30% at every turn and in that way at the third turn you are below the results of a simple logistic regression.

* + 1. The importance of the syntactic functions over POSs (part of speech)**[[13]](#footnote-13)**

Many times, humans are not aware of POSs when they are speaking, instead they carefully formulate sentences by making the right relations between the words they are using. Machines must focus not on nouns, for example, but on subject, predicate, complement and attribute and the syntactic relations (that is different from parsing in our view). We didn’t find in any NLP research papers this debate/topic and it is very odd (aka NER is not quite a subject, the complements are not always nouns and in many languages adjectives and nouns can be part of the Predicate).

* + 1. Reinforcement learning (RL) vs Supervised learning (SL) vs cybernetics systems theory vs creativity process**[[14]](#footnote-14)**.

To use Reinforcement Learning to achieve the best answers/reactions in conversational AI is, of course, not a fundamental decision as a conversation is not a game of chess in which the scope of the machine is to win. Conversation is not about rewards and about exceeding human-level capabilities even in the debate states. For that reason, RL is not researched in this paper, but a more interesting discussion would be in the policy choosing of a conversational infrastructure. The E2E -deep reinforcement learning-model is using rewards or attentions to bypass systems theory policies but at the same time it is impossible for the machine to be trained on every possible policy and to be able to simulate, especially in deep conversational topics or in the more creative art of conversation.

A neural network (RL or SL), especially in the back-propagation layer, needs to be very carefully adjusted to allow feedback and to take into account the deep conversational topics. Feedback is specific to conversation and was described by Gregory Bateson in the systems theory. The main difference between feedback and reward is the fact that feedback is impacting both agents’ behavior/policies (users’ and machines’), on the other hand, reward is not taking into account that also the user utterances are impacted, considering only the machine actions in a ceteris paribus environment. In a conversation, both agents have dynamic approaches– “You never step in the same river twice” – Heraclitus.

* + 1. Common interest**[[15]](#footnote-15)**

For a conversation to be both efficient and interesting the two parts need to find a common ground. Second most important thing is to know how to respond to emotional reactions. If not, it is highly probable to have a short dialogue without a conclusion/result. Taking this into account, the bot needs to show interest in the client’s needs, not only to understand or to give the right answer. How can it do this?

* Ask additional questions, which should be specific, not general.
* Confirm in some cases the user’s intentions.
* Involve in conversational or deep conversational topics but find ways to redirect always to the main topic.
* React to emoji and understand emotional reactions.

In practice, Conversational KB-QA as part of Reinforcement Learning covers many of these requirements, but the way they are doing it is not conversational by definition. Firstly, conversational is not the same as role plays. Secondly, using specific trained or programmed databases for queries is very much a sales approach (that is no longer accepted by users as far as we know) instead of a consultancy approach that is more conversational. Thirdly, there are big chances to ask a wrong question that will show the user that you don’t even listen. Finally, you risk the machine taking over the initiative, leaving the user with no answer for his initial need.

In reality, the fact that the user is building his question not in one-turn conversation but in a multi-turn conversation is due to the common way of structuring a question (introduction-context- and then the question, in one or more steps) not because he doesn’t know his needs or how to actually ask/synthesizes the questions (we do not include here task-oriented multi-turn dialogues).

“A man walking is never in balance, but always correcting for imbalance” – Gregory Bateson

* + 1. Grounding**[[16]](#footnote-16)**

The French believe that the most important thing in wine production is “terroir”. If you want to translate that in Conversational AI it will mean something like this: more important than semantics or the actual conversation is grounding. The environment in which the user and the machine is activating needs to be the same or understandable by both, if not the conversation will not feel real. For that reason, Conversational Bots grounded will be more effective than a generalized bot, even if the latter knows the correct answers. At the same time, there are domains targeted by Conversational bots that cannot work without grounding no matter what state-of-the art models you are developing. For example, you cannot ask for legal consultancy to a lawyer that practices in a different country than yours, or for restaurant recommendation to an evaluator that is working in another city than yours.

* + 1. Exceptions are exceptions**[[17]](#footnote-17)**.

All languages have exceptions in terms of having words/wording that do not follow the same pattern as the rest of the words in a phrase/sentence. For example: expressions, ironies, metaphors, quotes... Treating everything in a general way can have an important effect.

* + 1. Fundamentals on AGI**[[18]](#footnote-18)**

We have two main and different views on AGI. There are people that don’t think in the possibility of an Artificial General Intelligence or an Artificial Narrow Intelligence and other people that are concerned that we don’t have enough safety measures in place to tackle the negative impact of AGI, if implemented. Both of these views have a negative impact on developing at least a human-level AGI that are also known as weak AI or soft AGI.

In some AI branches we understand the risk associated with AGI implementation, but fundamentally, we cannot have a conversational AI bot without Human-level AI/AGI. Why? Conversation presumes the existence of two agents. If you have on one side a human and on the other side an alien, the probability for the conversation to be incomprehensible is very high, especially in a society where informal communication and empathy prevails. Now, using neural networks can be seen as an AGI concept if you have the objective to simulate human thinking, but the intuitions behind this model, that is math based, is not quite an AGI material. This is the reason why we don’t see an interesting dynamic in NLP/DL as we see in other AI domains and therefore it is mandatory to have a hybrid model for the conversational bot. It is, however, a big question which type of AGI is proper to use inside the Human-level AI that anyway is considered to be soft.

* + 1. Neural Network vs Labyrinth**[[19]](#footnote-19)**

Many times, it is advisable to go back to fundamentals also in areas where everybody has a clarified mindset. To use neural networks (it is custom in AI to use models from others domains/branches) to make translations or to find in a database the right answer to a question can be a good decision for cases when you want to substitute human brain with machine capacities. The big question of this research paper is if a neural network is suited for a conversational bot (that is more than a chitchat or a Q&A platform) as human knowledge is sometimes more intuitive and based on past experiences (like reading a book). Knowledge is not a mapping or labeling process and using a book instead of a database for training purposes (extracting the right output) will get you to this conclusion.

Human knowledge, that is essential in communication, is more resembling to a labyrinth that has two important differences from a neural network:

1. Intuition (a subjective decision to choose the desired path to move forward);
2. Finding different sources of inspirations (like reading a book and helping the machine to make queries instead of training databases).

A database is giving the bot (or to humans) solutions, but knowledge, to help him in a conversation, can only be obtained from books (not from knowledge graphs as ERNIE is proposing). This is the reason why it is important to make changes in terms of NLU and also implement new enhancements and models in NLG (intuitions and book queries).

* + 1. Machine Education**[[20]](#footnote-20)**.

For a human to be able to have an intelligent conversation, or to be able to translate, or to speak, or to summarize an article/book … he needs two things:

1. education (meaning personality)
2. knowledge (meaning learning, experience).

When you hire a person to do jobs related to the abilities in discussion (related to NLP) you will necessarily look for both: education and experience, and for sure the education in the linguistic domain will be an important criterion. If everybody agrees with that and we are building models for the machines to duplicate natural language, then why are we using only Machine/Deep Learning and not also Machine/Deep Education that can be regarded also as a AGI concept?

1. Why can a hybrid NLP/AGI model be a fundamental choice?

Let’s take as an example the medicine evolution in time, and more specifically the doctor’s advice over time.

The inception phase. Sophism. If a drug (medicine) makes you well, then 1000 drugs will make you 1000 better. This sophism was also experimented in NLP. Is the objective to train 10x more data familiar?

The 2nd phase. Gurus. If a medicine works for one disease, it should be used for many diseases, as it is a providential medicine. This is also familiar with different revolutionary methods in AI that were extended for most of the AI fields.

The 3rd phase. Quality. If a medicine has a positive effect, a qualitative shift will improve the results. This is the main focus today in data science.

The current phase in medicine. Hybrid. Changing your lifestyle/habits and doing sport, together with drug prescriptions, will bring benefits to your body and mind. This can translate into a conclusion that maybe conversational AI is some steps behind.

1. Considerations on BERT, ERNIE, Transformers, Open-domain Q&A’s, Intents/Entities/Forms models and Open-domain dialogues frameworks with summarization-augmentation.

As solutions already exist for building conversational AI bots, it’s only natural to take them into account, especially that some of them have shown important results for some languages, domains or tasks. A framework that is not using the current instruments cannot be considered serious. But also models that are only considering fine tuning (by using additional datasets or different optimizations of parameters or pipelines) in the conversational AI area can be regarded as un-sufficient in an environment characterized by many as emerging (even in English language)

1. Ethics and principles in Conversational bots

One of the big pros for using fundamental approaches in developing bots and not processing huge databases is the increased control that we have in dealing with regulatory requirements in terms of gender, sexual orientation, religion…. At the same time, using databases for learning algorithms without copyrights can be considered stealing even if at this moment those methods are allowed (an uncaught thief is an honest merchant).

A model needs to be also aligned with General Data Protection Regulations for each country, not to use the client’s personal data. The same is with conversational/ deep conversational answers (inside NOG/NLG) that need to be given under literary copyrights.

Please find Bibliography referrals in the research paper *“Cezanne-ai: a conversational AI open-framework for multi-domains, all the languages and limited data”.*

1. *The etiquette ‘emerging’ is used also by Gao/Galley/Li in the 2019 paper: Neural approaches to Conversational AI.*  [↑](#footnote-ref-1)
2. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding - Jacob Devlin, Ming-Wei Chang, Kenton Lee, Kristina Toutanova- 2018, https://arxiv.org/abs/1810.04805v2 [↑](#footnote-ref-2)
3. ERNIE: Enhanced Language Representation with Informative Entities - Zhengyan Zhang, Xu Han, Zhiyuan Liu, Xin Jiang, Maosong Sun, Qun Liu – 2019, https://arxiv.org/abs/1905.07129v3 [↑](#footnote-ref-3)
4. Attention Is All You Need - Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Lukasz Kaiser, Illia Polosukhin – 2017 - https://arxiv.org/abs/1706.03762 [↑](#footnote-ref-4)
5. „Survey of Artificial General Intelligence Projects for Ethics, Risk, and Policy” by Baum/Fitzgerald/Boddy – 2020 [↑](#footnote-ref-5)
6. See also Harvard Business School article for the big picture related to this topic – Strategies that fit Emerging Markets – June 2005, https://hbr.org/2005/06/strategies-that-fit-emerging-markets [↑](#footnote-ref-6)
7. See also Peter Dizikes, Unique languages, universal patterns <https://news.mit.edu/2012/unique-universal-languages-0223>. Even if the article can be seen as divergent from our view, the fundamental conclusions are similar. [↑](#footnote-ref-7)
8. Semiotics on Wikipedia - https://en.wikipedia.org/wiki/Semiotics [↑](#footnote-ref-8)
9. for more on this we will refer to philosopher John Locke entire work, https://ro.wikipedia.org/wiki/John\_Locke [↑](#footnote-ref-9)
10. See also J. de Rosnay,Analytic vs. Systemic Approaches**,** <http://pespmc1.vub.ac.be/ANALSYST.html> [↑](#footnote-ref-10)
11. See also Dustin Smith, Nonverbal Communication: How Body Language & Nonverbal Cues Are Key

    <https://www.lifesize.com/en/blog/speaking-without-words/> [↑](#footnote-ref-11)
12. See also Brandon Lin, Investigating the Machine Reading Comprehension Problem with Deep Learning <https://towardsdatascience.com/investigating-the-machine-reading-comprehension-problem-with-deep-learning-af850dbec4c0> [↑](#footnote-ref-12)
13. „Is POS Tagging Necessary or Even Helpful for Neural Dependency Parsing?”- Zhou, Zhang, Zhang, Li - 2020 [↑](#footnote-ref-13)
14. See also „Gregory Bateson's cybernetic methodology: The ecosystem approach in empirical research” - [Pasikowski](https://www.researchgate.net/profile/Slawomir-Pasikowski) - 2017 [↑](#footnote-ref-14)
15. See also „Knowledge Base Question Answering” - Dhingra – 2017, Wu - 2015 [↑](#footnote-ref-15)
16. „You are thinking your culture thoughts” – Gregory Bateson [↑](#footnote-ref-16)
17. See also Kaufer: Irony, Interpretive Form, and the Theory of Meaning: <https://www.jstor.org/stable/1772026> [↑](#footnote-ref-17)
18. See also: Ragnar Fjelland: Why artificial intelligence will not be realized: <https://www.nature.com/articles/s41599-020-0494-4> [↑](#footnote-ref-18)
19. See also a debate over the epistemology theory by Britannica: <https://www.britannica.com/topic/epistemology> [↑](#footnote-ref-19)
20. This fundament has similarities with Inria- Flowers project approach: developmental robotics. [↑](#footnote-ref-20)