

# Robustness Challenge with Dialog Systems, Myths and Solutions

---

**Eric Charton, Ph.D.**

Senior AI Director, Scientist



Let's say : we want to build a bot strategy from scratch ...



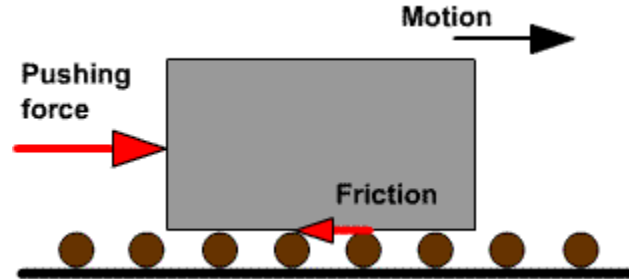
Instantly, dozens of vendors and SME appears

## Common point between all the solutions

---

In 5 months, over 20 vendors audited, from startups to majors. Common point :

- Dozens of famous competitors use the solution ... but we can't see it
- If you want to test, you need to order a POC
- We can't test the system
- Most often, there is no demo from the vendor website (eat your own food ?)





If you do not want to let us see your product, is it really as robust as you claim ?

13:50

Démarrer



Hello Eric. I'm Remi BetaBot, your Money Transfer Advisor! I can assist you with your banking needs by setting up rate alerts and transfers.

This chat will be recorded and stored by Facebook and will be accessible by [redacted] I'm happy to answer all of your questions, but please do not disclose any personal or account information, such as account/card numbers or passwords. [redacted] foreign exchange rates quoted are subject to change at any time without any notice.



How can I help?

Check rates



Sure thing! Which rate would you like to check? Here are some suggestions!

what is the rate for mexican peso



Sure thing! Which rate would you like to check? Here are some suggestions!

Check rate to EUR

Check rate to GBP

Check rate to INR

Check rate to PHP

Check rate to USD

Other

[? FAQs](#)


Need some inspiration on what questions to ask? Try typing "How do I view my balance?" and hit enter.

You can also ask about changing your username and password, transferring money, reporting a lost card, and much more. If I can't answer you, just say "Talk to an associate"! 🙄

[i lost my pin](#)

No prob! Did you forget your PIN, or would you just like to change it?

[Forgot my PIN](#)

If you've forgotten your PIN, just call at 1-800-263-2263 or visit a  branch to get it sorted out.

[how can i save money](#)

We have a few ways to move your money online, including e-Transfers to send money within Canada, and Western Union or Global Transfer for international transfers.

Where is your money headed?

[Within Canada](#)[Internationally](#)

Ok

What is the goal?

Z

I don't know the word "ok".

What's my goal in this game?

I don't know the word "what's".

Z

I don't know the word "goal?".

Do you know anything?

Z

I don't know the word "do".

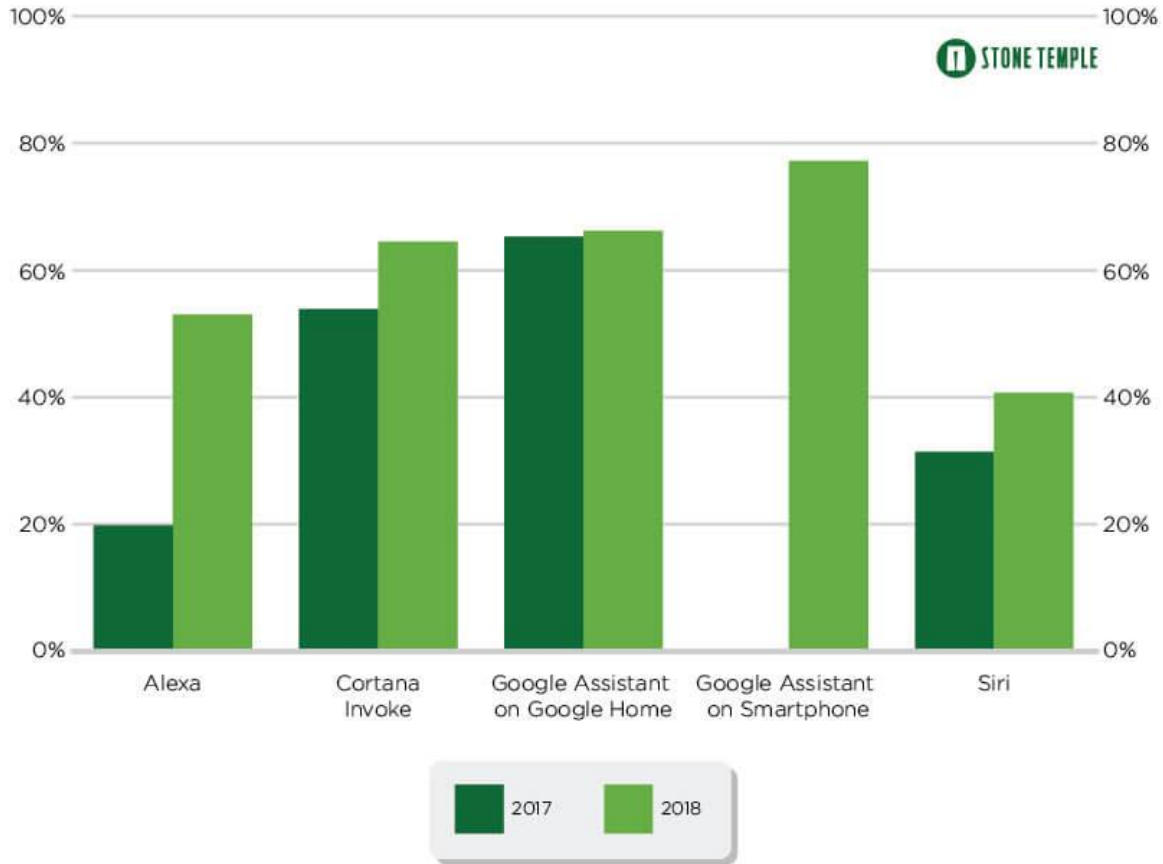
Z



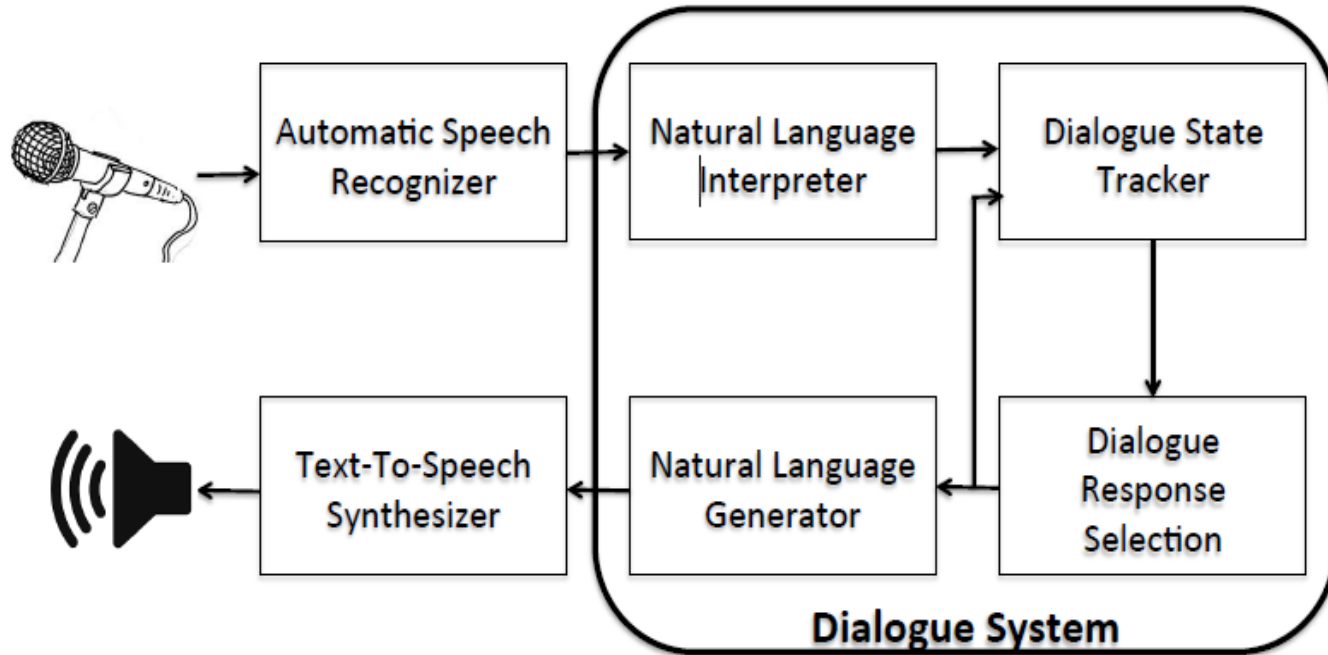
*Facebook has scaled back its ambitions and refocused its application of "artificial intelligence" after its AI bots hit a 70 per cent failure rate. The Information, tests, Feb 2017*



# Answers Attempted YOY Comparison



## Chatbot are dialog systems



# Two main families of Dialog Systems

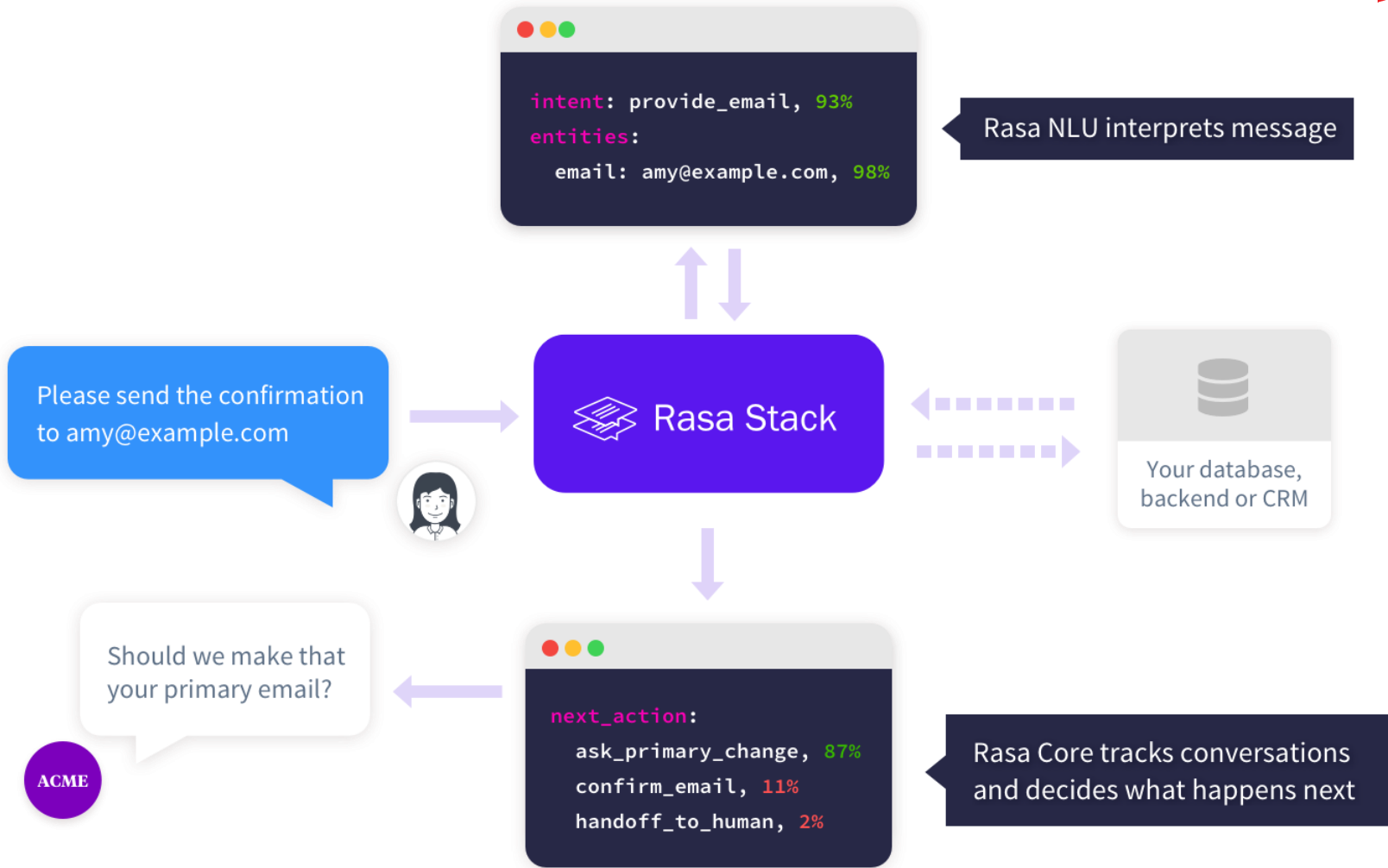
---

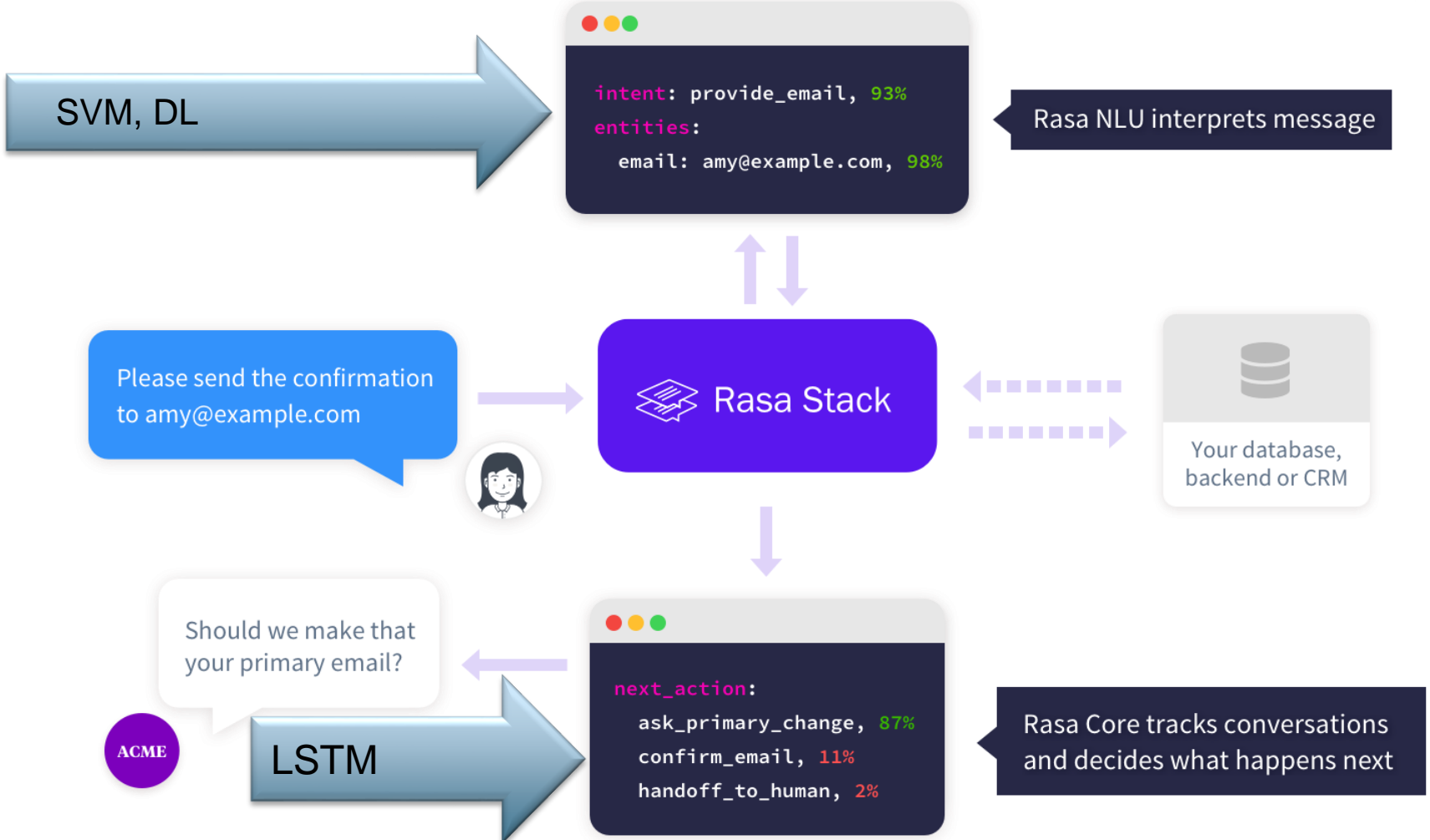
## Rule based systems

- **A team of engineers, linguists craft manually a dialog tree.**
  - ❑ Pro : easy to deploy, simple technology, good precision level on simple use cases
  - ❑ Cons: limited intelligence, high cost to progress, limited generalisation

## Corpus based systems:

- **A set of question and answers sets and dialog sample is collected and submitted to an algorithm for training (Deep Learning, SVM, POMDP)**
  - ❑ Pro: great capacity of generalisation, high potential of intelligence, quick update by new training
  - ❑ Con: the corpus collection is sometimes difficult. Technology can be difficult to master





## Solutions to increase robustness



- Increase question formulation coverage (NLU Model)
  - Paraphrasing
    - Using Generation models
    - Using rules based system
  - Word embedding
- Increase dialog capacities (Dialog model)
  - Programmatic discourse handling
  - Dialog training set generation

---

## Paraphrasing : increasing the intent identification

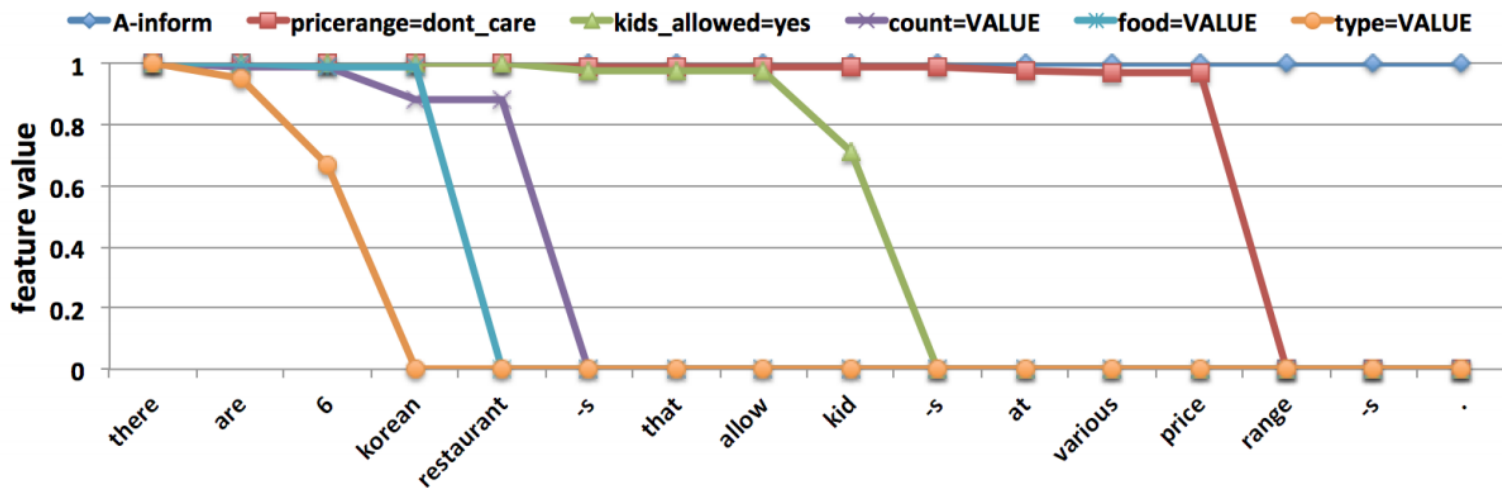
---



# Paraphrasing & Natural Language Generation



- **Semantically Conditioned LSTM-based Natural Language Generation for Spoken Dialogue Systems** ([Tsung-Hsien Wen](#), [Milica Gasic](#), [Nikola Mrksic](#), [Pei-Hao Su](#), [David Vandyke](#), [Steve Young](#))



(a) An example realisation from SF restaurant domain

## Paraphrasing



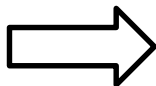
#	Example Dialogue Acts and Realizations from SF Restaurant Domain
1	<p>inform(name="red door cafe", goodformeal="breakfast", area="cathedral hill", kidsallowed="no") <i>red door cafe is a good restaurant for breakfast in the area of cathedral hill and does not allow children .</i> <i>red door cafe is a good restaurant for breakfast in the cathedral hill area and does not allow children .</i> <i>red door cafe is a good restaurant for breakfast in the cathedral hill area and does not allow kids .</i> <i>red door cafe is good for breakfast and is in the area of cathedral hill and does not allow children .</i> <i>red door cafe does not allow kids and is in the cathedral hill area and is good for breakfast .</i></p>
2	<p>informonly(name="dosa on fillmore and kiss seafood", pricerange="expensive", near="lower pacific heights") <i>there is no place other than dosa on fillmore and kiss seafood that are expensive near to lower pacific heights .</i> <i>dosa on fillmore and kiss seafood is the only expensive restaurant near lower pacific heights .</i> <i>the only listed restaurant near lower pacific heights in the expensive price range is dosa on fillmore and kiss seafood .</i> <i>i apologize , dosa on fillmore and kiss seafood is the only expensive restaurant near lower pacific heights .</i> <i>i apologize , dosa on fillmore and kiss seafood are the only expensive restaurants near lower pacific heights .</i></p>
#	Example Dialogue Acts and Realizations from SF Hotel Domain
3	<p>inform(type="hotel", count="182", dogsallowed="dontcare") <i>there are 182 hotels if you do not care whether dogs are allowed .</i> <i>there are 182 hotels if you do not care whether they allow dogs .</i> <i>182 hotels are available if dogs allowed or not is not an issue .</i> <i>there are 182 hotels if allowing dogs or not is not an issue .</i> <i>there are 182 hotels if whether dogs are allowed does not matter .</i></p>
4	<p>informonly(name="red victorian bed breakfast" accentscreditcards="yes" near="haight" hasinternet="yes")</p>

Discourse generation : using  
human behavior to handle the  
dialog utterances more friendly

## Example of conversation mechanisms that can be implemented



**User:** Est-ce que la banque vend des bananes?



+ Interpreter: predict class #6 with confidence **0.02**  
+ Policy: Confidence too low: fallback

**Bot:** Désolé, mais je ne trouve pas avoir la réponse a votre question.

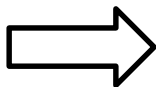
Essayer de reformuler, ou naviguer avec les boutons pour restreindre la recherche.

Voir les sujets principaux

## Using keywords and Named entities to handle dialog programmatically



**User:** Qu'est-ce que je fais avec une mise en demeure?



+ Interpreter: Class predicted #6 with confidence **0.5**  
+ Interpreter: recognize «*mise en demeure*»  
+ Policy: Medium confidence level. Return questions with keyword

**Bot:** Je ne trouve pas la réponse à votre question.

Cependant, j'ai trouvé ces questions qui pourraient vous aider:

**À qui dois-je m'adresser si je reçois une plainte d'un client ou une mise en demeure?**

# Use case and KPIs



# Example :

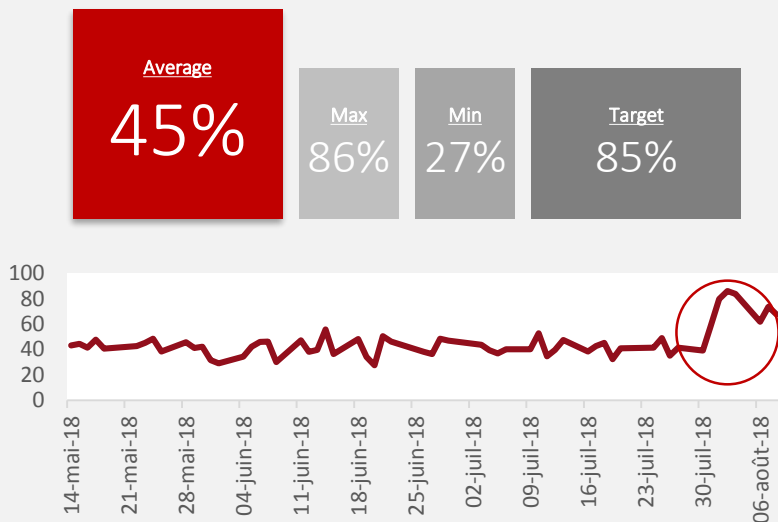
The screenshot displays a banking portal with a chat window overlaid. The chat window is titled "Clavarder avec Expert" and shows a conversation with a virtual assistant. The customer asks: "Cliente a une transaction douteuse. Que dois-je faire ?". The assistant responds: "Suivez les indications dans la procédure Refus de transaction." and "Puis-je faire autre chose pour vous?". The customer replies: "Oui, j'ai une autre question." The assistant then provides instructions: "Pour clavarder avec un expert, entrez 'Expert' à tout moment." and "Suivez les indications dans la procédure Transfert / Correction de points." The chat window also shows a list of accounts and a search bar at the bottom.

Nom du compte	Numéro de
<b>Comptes bancaires</b>	
PROG.	01711 - 292
PROG.	10421 - 251
E.PROJET	10421 - 765
MCP	10421 - 766
<b>Cartes de crédit</b>	
MC OR AG	525897200
<b>Financements</b>	
TAUX VAR	10421 - 000
+ HY	000020600
<b>Investissements</b>	
FMUT RER	10421 - 238
FMUT RER	10421 - 239
FMUT RER	10421 - 238
RE B FIX	01711 - 000
RE B FIX	10421 - 000
RE B EVO	10421 - 0000579230686313
RE B EVO	10421 - 0000579230686763
RE B CBG	10421 - 0000579232812324
Clavarder avec EMC	10421 - 0000579234122945

# Impact of paraphrasing on the system



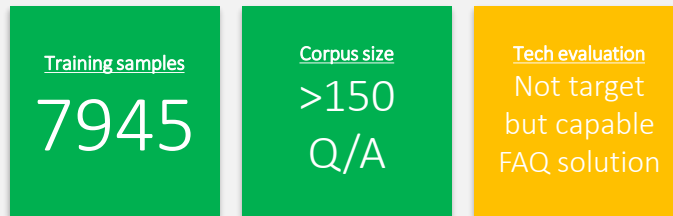
Evolution of chatbot confidence level



**Insights:**

- Adding of new training data (automatically generated) on July 31st lead to a jump in the confidence level : from 41% to 73%.

Training corpus structure

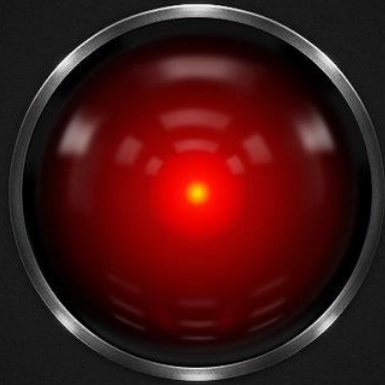


**Insights:**

- The pilot started with a classic FAQ -> 22 questions and 22 answers
- Within the first 4 months we have increased to 7945 training elements (questions) and 150 answers
- Pilot developed a technical solution to accelerate bot learning
- Used solution (Icontek) is a viable option for FAQ style bot architecture; chat functionality needs to be expanded



HAL 9000



Thank you !

Thank you ! Questions?

▪ [Eric.charton@bnc.ca](mailto:Eric.charton@bnc.ca)