UX analytics for chatbots: metrics, method and case study

uxprobe

White Paper

Table of contents

"Where 10 years ago every company needed a website and five years ago every company needed an app, now every company needs to embrace messaging with AI and chatbots."

Murray Newlands[1]

Executive summary

If 2016 was the year of the rise of the chatbots, 2017 should answer the question "How can we make them better?" Because it's one thing to have a fun conversation with a piece of AI, and being amazed that the bot understands what you say and answers in a clever way; it's another thing to have a flawless chat with a bot to get something done: book a flight, order a pizza or find new tyres for your car.

That flawless chat won't exist without lots of efforts from the team building it. So what do they need? First, they need to set up clear KPIs based on the goals of the bot. Second, they need to measure in order to understand how the chatbot is used. Quantitative and qualitative metrics are the fundamentals for improving the quality of the chatbot.

Analytics for chatbots is going to be the new trend of 2017. Learn why that is so, and how Bridgestone has already benefited from the analytics solution provided by UXprobe.

"Chat bots are looking like the big trend of 2016, with Facebook M, Operator, Magic and more vying to help us out via instant message."[2]

Introduction

In 2016, we saw simultaneously Google, Facebook and Microsoft coming out with new products, the chat assistants, based on chatbots. Suddenly, in one year, a technological gap was filled in, and chatbots started to work well enough to be useful. At the same time, developing chatbots was made so much easier through different tools (Facebook Messenger, Microsoft, Smooch, Chatfuel, etc.)[3] that they came quickly within reach of any company.

Like Murray Newlands, expert in chatbot, said: "Where 10 years ago every company needed a website and five years ago every company needed an app, now every company needs to embrace messaging with AI and chatbots."[1]
If 2016 really saw the rise of chatbots, **2017 is where we start to ask: "How can we make them better?"**

Where it starts – a little bit of research on TechCrunch

A quick search on TechCrunch about "chatbot" gives **13 pages of results.** The very first reference of the word brings back to 2006, "When Verizon Chatbots Attack". B**etween 2006 and 2015, there are 7 articles** (still online of course, I don't know if they archive some of their articles). **For 2016, there are 10 pages with 9 articles** listed on each of them. It starts on February 16th, with an article entitled "On Chatbots". **For 2017,** there are already 2 pages of articles listed, which means **18 articles published** in 2 months and a half.

If you search for "chat bot" (in 2 words, between brackets) you get 6 pages of results; 1st article in 2007, but also starting really in 2016. If you don't put the brackets, you get 479 pages (probably because it includes everything with the word "chat").

Nice piece on the **history of chatbots**, starting in the year 2000 with SmarterChild, written by Matt Carbone, engineering leader at Talla.

"Chatbots will dislodge apps in terms of ubiquity, usage depth, and importance in our everyday life." Topbots magazine[4]

Definition and perspective

What is a chatbot?

A chatbot is a conversational interface where users can have a conversation with a piece of software to achieve a goal: find an information or a specific product, get a recommendation, answer a quizz, etc.
According to Shawn Hansen, Mixpanel's CMO, "the top five performing categories of bots are weather, fashion, humor, knowledge-based quizzes, and notification-driven news."[5]
One very famous chatbot is Poncho, a chatbot dedicated to personalized weather forecast "It's like being friends with your local weatherman, if he were a cool cat from Brooklyn."[6]

To learn about chatbots, right when it started, read that very first article about chatbots on TechCrunch in February 2016 ("On chatbots"). See also box p. 4 "Where it starts - a little bit of research on TechCrunch".

Why are chatbots so trendy?

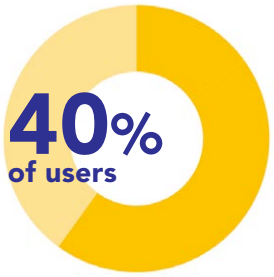
A lot has been written about how messaging apps are going to replace any other type of communication. For example, look how Slack is replacing emails for team communications, or read what Thomas Husson, VP and principal analyst at Forrester, says about it: "We expect messaging apps to play a key role throughout the customer life cycle but more specifically to enable brands to deepen conversations with their customers during the retention phases. Why? Because messaging apps combine the three keys to powerful relationships in any digital environment: **frequency of use, emotional connection, and convenience.**"[7]
You can see how chatbots would take over a large part of these online conversations: chatline for customer service, long forms, research/exploration tool.
See also box underneath.

How Google, Facebook and Microsoft threw themselves into chatbots in 2016

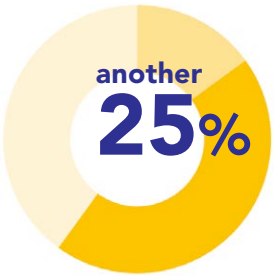
- **I/O Google,** May 2016 – watch the video.
  Sundar Pichai: "We want users to have an ongoing dialogue with Google."

- **Facebook F8,** April 2016 – Focus on Messenger, watch the video.
  Mark Zuckerberg: "No one wants to have to install a new app for every service or business that they want to interact with. […] We think you should be able to message a business the way you message a friend."

- **Microsoft Build,** March 2016 – How Microsoft personal assistant is integrated with Skype, watch the video.
  Lilian Rincon (Director of PM Bots and Conversation as a Platform at Skype): "We're moving into a world where you can plan trips, shop and even talk to an intelligent bot, all within your Skype chats."



stop talking to a bot
after the first message



are gone
after the second one

UX analytics for chatbots

Currently analytics tools rarely provide conversational analytics solutions and while the demand for building chatbot is rising, companies now want to know when and where their bots fail to conduct a successful conversation.

Who needs analytics for chatbot

"To me, a good chatbot analytics tool should monitor all conversations between the bot and its users so the bot owner knows the entire customer journey"[8]

What is good for marketers, is not enough when it comes to usability and improve your chatbot itself, then you need far more than that.
That's what we are busy with at UXprobe.

Why you should measure

To improve retention and engagement
According to Ilker Koksal: "nearly 40% of users stop talking to a bot after the first message and another 25% are gone after the second one. Sadly, daily and monthly retention numbers aren't stellar either, with daily rates hovering between 1-2% and monthly rates barely exceeding 7%."[4]

To improve quality
Measuring is the only way to improve the quality of a chatbot. If you don't know what's going wrong, you won't ever be able to fix it. Machine learning is a nice thing, but not all the bots are using artificial intelligence capable of learning by itself.

If you want to build a simple chatbot for a client, you need simple ways to measure in order to get your most basic KPIs.

KPIs to consider

The point of a chatbot is not to have nice little conversations with customers – no – it's to let your customers get something done – find out their flight schedule, order a concert ticket, find the nearest tyre dealer.
But "measuring success for companies is always a tricky topic: the right metrics for one company may not be the same for another. "[9]
Juliette Kopecky, VP of Marketing at Talla, gives two interesting examples[9]. At Hubspot, they were very focused on customer usage as a way to measure success since it was a good indicator whether or not a customer was going to churn. At Backupify, it was the opposite. Since their product was designed as a failsafe, if a customer didn't log in for weeks or months, that was fine as long as the product gave them what they needed when disaster struck.
In other words, each company has to decide for themselves what KPIs are the most relevant. Nevertheless, here are a few to consider:

- **Number of users**
- **Duration of a chat** (average, longest, shortest)
- **Users success rate**
- **Number of returning users** (how many times a user is using the chat)
- **Click Through Rate for links** (to evaluate how many leads / conversion have been made through your chatbot)
- **Satisfaction rate** (through a quick feedback survey in the end)

By measuring task success you can see very quickly how usable your chatbot is.

What is important to measure

Success and failure
In the user experience (UX) domain, a task is what a user is supposed to do when he wants to go from A to B (beginning to end). In chat lingo, a task is called the intent. Your task will be defined by the intent of the chat, what are the users supposed to do with your chat.
For example, a typical intent for a chatbot used in customer support would be to fix an issue or have an answer to a question related to an issue. The task will end when the customer has a solution to her problem. A verification question can be added at this moment: "Was this helpful?" or "Are you satisfied with the solution?"

With a tool like UXprobe, you can **define your own tasks and determine what is success and what is failure.** By measuring this, you can see very quickly how usable your chatbot is. You can determine how often a chat ends in a success for a user.
Moreover, you can see how long it takes users to complete and what path they use to get there.

Chasing bugs and errors
When developing a chatbot, it especially difficult to generate all the flows and have all the proper entities and synonyms in the chat dictionary. It is then easy for chatbots to get stuck and that's what you have to fight against.

By measuring every time a chatbot reaches a dead end, you can filter and group all the error cases. Looking at only those sessions with errors, you will easily find the bugs in the chatbot logic and missing synonyms from the dictionary.

Collecting feedback
By the means of a micro-survey, like thumbs up/thumbs down, five-star rating or NPS, you can quickly determine the satisfaction rate of your chatbot.

Quick overview: talks & meetups about chatbots in Belgium

- **Pieter Rahier,** UX designer at Pàu (Brussels) "The Near Future for Articial Intelligence and Conversation Bots" at UX Antwerp Meetup, April 2016
- **Alexis Safarikas,** digital strategist at Springbok Agency (Mechelen) "A step further in UX: Conversational Interfaces" at UX Antwerp Meetup, November 2016
- **Sjoera Roggeman,** "Chatbots: building intelligent systems" at UX Antwerp Meetup, January 2017 - video
- #BeBot Meetup founded in November 2016, first meetup January 2017
- Brussels Chatbot Meetup founded in January 2017, first meetup March 2017
- 2 other meetups around: Utrecht & Amsterdam

Building a chatbot, yes.
But let's make sure it does a good job.

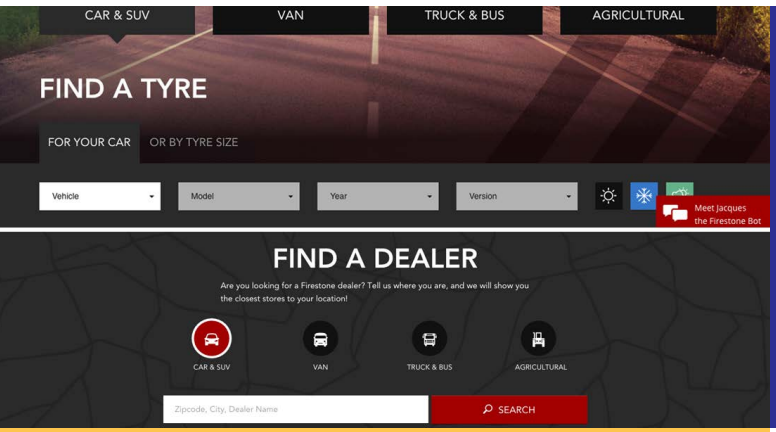
Case study: a chatbot for Bridgestone

The company & project

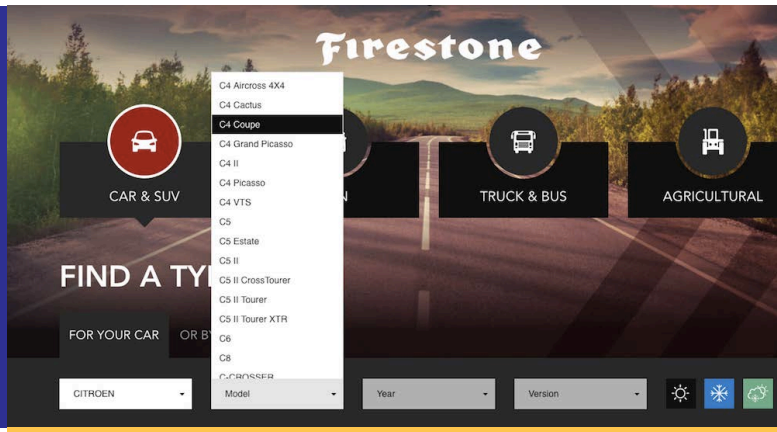
Springbok is a Belgian based digital agency that provides digital services for companies in need of a digital strategy, marketing consulting, or specific technical support. Springbok aims to be at the forefront, bringing in disruptive ideas and solving technical challenges.
That's why, when **Bridgestone** asked Springbok to improve the website of their brand **"Firestone"**, they didn't go the easy way. They went a step further by offering to build a chatbot. And that made especially sense for that project as you're going to discover it now.

BRIDGESTONE

SPRINGBOK



Pic 1. Firestone UK website has 2 very clear goals: find a tyre & find a dealer



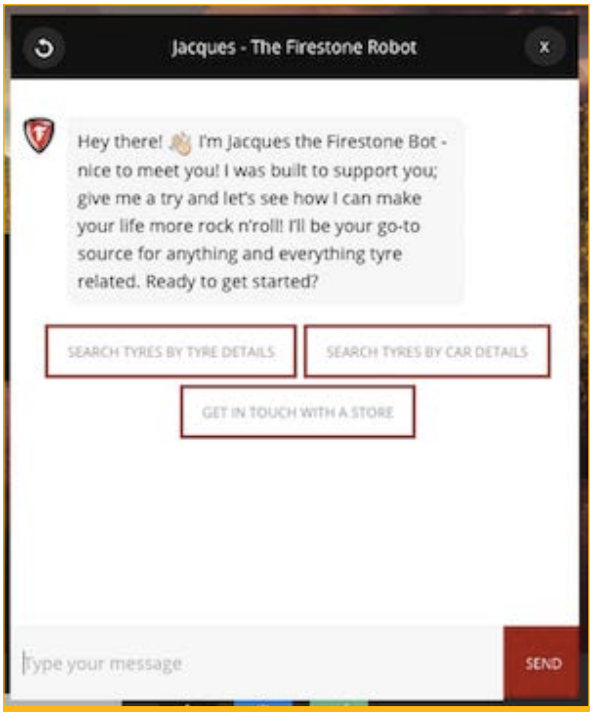
Pic 2. Drop-down menus on Firestone UK website

The challenge of the chatbot

Firestone's website has two clear goals (SEE PIC 1):

- entice the visitor to choose new tyres for his car (section "Find a tyre")
- or find a vendor to buy them (section "Find a dealer")

This is more complicated than it seems.
Finding the right tyres for a car is difficult because a lot of parameters come into account. The current process requires visitors to go through a long series of drop-down menus (SEE PIC 2) to introduce all the required information about their vehicle or tyre's specifications.



Pic 3. Firestone chatbot - choosing one of the intents

In that case, a chatbot is a very appropriate solution to ease the pain of a long, time-wasting process. **Potentially it can guide visitors through the whole experience in a helpful and easy way.** Giving information – even specific ones – during a conversation should make it a much more enjoyable experience than filling in a form. Should. If it works well. Alexis Safarikas, the digital strategist at Springbok, is the person who brought the idea of a chatbot replacing the form on the website. Once the project was accepted, he quickly realised that he had to ensure that the chatbot he was building for Bridgestone was indeed working well. To do so, Alexis had in mind a few important KPIs he wanted to measure.

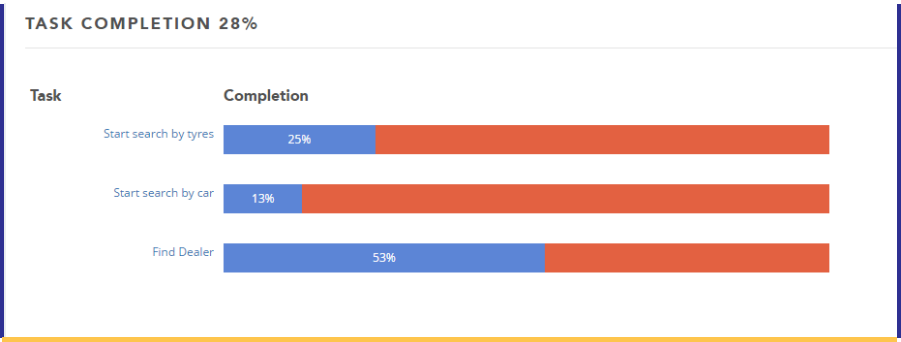
What Springbok wanted

A chatbot is a conversational interface. The users of a chatbot have a conversation with a robot to achieve something: this is called the intent. It is what your users intend to do with your chatbot, what they want to achieve. Chatbots can have closed or open intents.

In Bridgestone's case, of course, the chatbot has closed intents. Bridgestone knows exactly what they want their users to achieve with the chatbot.
3 simple intents were identified (SEE PIC 3):

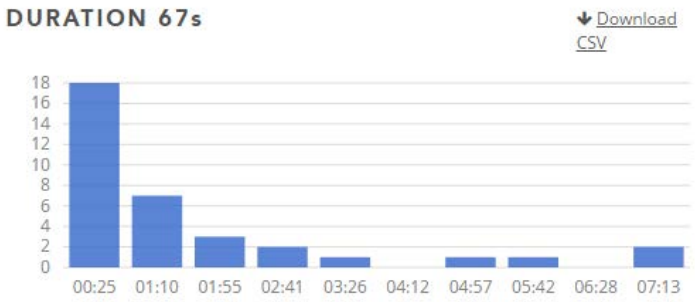
- Find tyres through car information
- Find tyres through tyre characteristic
- Find and contact a dealer

The most important KPI Alexis wanted to get was **the success rate for each of the intents.** That, according to him, would give a good indication on how well the chatbot itself



Pic 4. Task completion graph in UXprobe dashboard

was working. He was also eager to know the duration of each chat, the click-through rate of the links to vendors… But how to get these numbers? There was no easy answer because, unfortunately, there aren't too many analytics around to help with that.
Until Springbok came in contact with UXprobe.



Pic 5. Task duration graph in UXprobe dashboard

UXprobe's contribution

UXprobe has developed an online software to analyse User Experience by measuring success and satisfaction in apps, websites, but also in chatbots.

Therefore UXprobe brought a solution perfectly suited to Springbok's needs and requirements.

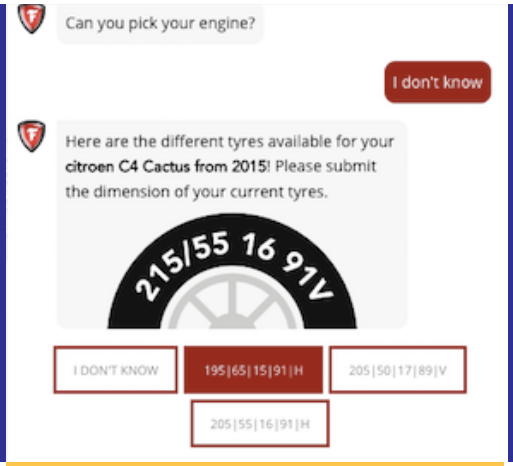
- **Definition of goals and measurements:** UXprobe starts as a service with an evaluation of the intents to make sure you get the appropriate metrics
- **Easy to connect to REST service:** UXprobe works with all the existing tools and NLP service that are already implemented.
- **Automatic collection of data** both quantitative and qualitative
- **Analysis in UXprobe's dashboard:** online dashboard displaying a UX metrics overview with relevant graphs and charts.

Numbers are in

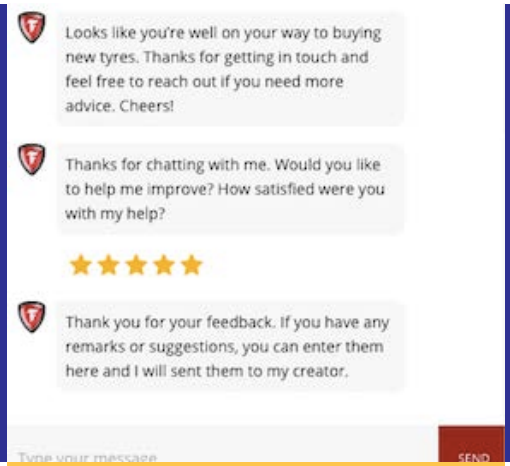
The base unit to measure the user experience (UX) is a task; it is the equivalent of the intent. It is then trivial to use UX metrics based on tasks to measure the quality of chatbots by using the intent as the base unit.

It was therefore very easy to get the numbers that Alexis was so eager to collect, simple but critical KPIs to measure the quality of the chatbot:

- **Success rate on 3 different intents:** users reaching the end of the chat with the right type of tyres, that's a success! (SEE PIC 4)
- **Chat duration:** how long does it take to successfully complete a chat (SEE PIC 5)
- **Click-Through-Rate:** how many users clicked on the button "Contact the dealer"
- **Satisfaction rate:** quick feedback panel at the end to rate the experience on a scale from 1 to 5
- **Specific bugs identification:** tagged as error when a session is too long when the chat cannot understand the users' answer/language



Pic 6. Firestone chatbot, task "Search by car"



Pic 7. Firestone chatbot, micro-survey in the end

With these KPIs in hand and the possibility to compare them with the existing website, Alexis has control over the chatbot he's building for Bridgestone.

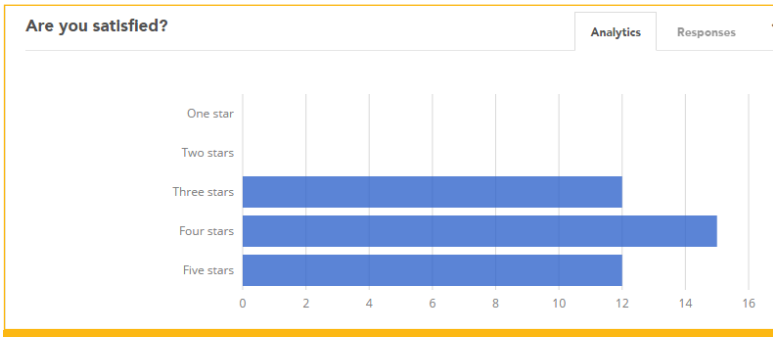
Results and learnings

The success rates show that the chatbot is more successful when the search starts by tyres than by type of car: 25% success against 13%. (SEE PIC 6) The number of visits shows also that more users chose to search by tyres (262 users against 166 for search by car). The success rate for finding a dealer is much higher: 53%.
The drop-off rates range from 40.45% (search by tyre), 31.92% (search by car) to 21.95% (dealer locator).

At the end of the chat, 24 users clicked to request contact information from dealers.
The qualitative survey got 15 answers: 3x "3stars" + 7x "4 stars" + 5x "5 stars" (SEE PIC 7 & 8), and Firestone received 2 emails feedback: "cool" + "I like this bot!".

The comparison with the existing website showed that visitors were preferably using the chatbot for searching tyres, and the website for searching for a dealer. The conversion rates for finding a dealer were almost the same (19.5% for the chatbot, 21.51% for the website).

The drop-off rates and some other errors that were identified thanks to UXprobe suggested that the quality of the chatbot was not optimal. After investigation, Springbok saw that the problems were related to the reliability of the API service they were using. Detected soon enough, they could find a solution in a short period of time.
Even if there is still room for improvement at this point, Alexis is now confident to tell his client that he's driving the new chatbot towards success.



Pic 8. Satisfaction graph in UXprobe dashboard

You now can see how quality impacts user success or how user success impacts user satisfaction.

Conclusion

In 2017, if you want to answer the question "How can I make my chatbot better?", you will need at least two things. First, set up clear KPIs based on the goals of the bot. Second, measure the user experience of your chatbot with UX analytics service, like UXprobe. This will allow you to very easily measure:

- **User success** – important for your stakeholders to gauge the usefulness of the chatbot
- **User satisfaction** – critical to understand if the chatbot is delighting or frustrating your customers
- **Chatbot quality** – so it can be improved

Not only will you gather three important metrics but also, these three metrics will be in relation to each other. In other words, you now can see how quality impacts user success or how user success impacts user satisfaction.
This will give you actionable data to make your chatbot the best it can be.

Do you also want to improve the quality of your chatbot?

Call us +32 (0)485 69 78 35 **Send us an email** happy@uxpro.be
Check our Website www.uxpro.be **Watch our video** http://youtu.be/xRYuRvhlWVQ
Follow us on Twitter @uxprobe **Facebook** Look for UXprobe **Google+** Look for UXprobe

uxprobe

References

1. **Entrepreneur,** "Top 10 Best Chatbot Platform Tools to Build Chatbots for Your Business", by John Rampton https://www.entrepreneur.com/article/289788
2. **TechCrunch,** "Assist Has No App. It's A Free Travel/Shopping Chat Bot For SMS And Facebook Messenger", by Josh Constine https://techcrunch.com/2016/01/06/assist-bot/
3. **Arpatech,** "Top 7 Chatbot Platform Tools To Build Chatbots For Your Business", by Aiesha Hasan http://www.arpatech.com/blog/chatbot-platform-tools-for-your-business/
4. **Top Bots,** "When Bots Go Bad: Common UX Mistakes in Chatbot Design" http://www.topbots.com/common-chatbot-ux-mistakes-in-bot-design/
5. **Top Bots,** "4 Critical Steps to Maximize Chatbot Retention & engagement", by Mariya Yao http://www.topbots.com/4-critical-steps-to-maximize-chatbot-retention-engagement/
6. **Market Wired,** "Poncho wins the chatbot popularity contest; keeps the yarn ball rolling with Kik launch and new suite of features" http://www.marketwired.com/press-release/poncho-wins-chatbot-popularity-contest-keeps-yarn-ball-rolling-with-kik-launch-new-suite-2167998.htm
7. **Forbes,** "The Future Of Messaging Apps", by Thomas Husson https://www.forbes.com/sites/forrester/2016/09/22/the-future-of-messaging-apps/
8. **Chatbots magazine,** "Conversational analytics", by Ilker Köksal https://chatbotsmagazine.com/conversational-analytics-87b4e82778d3#.z29hdijtu
9. **Talla Company,** "Measuring Success in the Land of Chatbots", by Juliette Kopecky http://blog.talla.com/2016/08/measuring-success-in-the-land-of-chatbots/

For more technical content about analytics & chatbots

- **"Chatbot Analytics - Getting started",** by Paul Davies https://www.uxpro.be/chatbot-analytics-getting-started/
- **"Chatbot Analytics: Hacking ChatFuel",** by Paul Davies https://www.uxpro.be/chatbot-analytics-hacking-chatfuel/

Call us +32 (0)485 69 78 35 **Send us an email** happy@uxpro.be
Check our Website www.uxpro.be **Watch our video** http://youtu.be/xRYuRvhlWVQ
Follow us on Twitter @uxprobe **Facebook** Look for UXprobe **Google+** Look for UXprobe

uxprobe