The voice of the customer for Digital Telcos

V. Richard Benjamins
David Cadenas
Pedro Alonso
Telefonica, Spain

Antonio Valderrabanos
Josu Gomez
Bitext, Spain

Abstract
In the midst of the digital revolution, the telecommunications industry is undergoing major changes. One of the changes affecting telcos is the increase in data sources from which to get customer feedback – big data, social media. Where this used to be fully controlled by companies through their call centres, websites and shops, today much feedback is expressed in social media, blogs, news sites, app stores and forums. Telefonica has taken up this challenge and opportunity, and is now systematically listening to the voice of its customers online.

The problem

Only a few years ago (around 2010), telcos received a hard wake-up call when Whatsapp started to significantly decrease SMS revenues. Since then, the choice for over-the-top (OTT) products and services has multiplied by orders of magnitude, and telcos run the risk to loose the end-customer contact, and to be forced into a connectivity-only offering. The answer of the telecoms industry to this threat has many aspects (beyond the scope of this paper), including the launch of digital services such as financial services, security, video, etc; leaner working methodologies (lean start-up); and much more customer-driven development and in-live management.

Listening to, understanding, and acting on customer insights are key for launching, growing or (rapid) killing of customer propositions. In this work, we present our approach to systematically and automatically listen to customers on the Internet as soon as a product has gone live[1].

Our approach is built around three main concepts: (i) crawling the Internet, (ii) concept identification & sentiment analysis, and (iii) visualisation, and is set up in such as way that any person with "advanced excel skills" is able to self serve the needed dashboards in a matter of hours.

Crawler

We use a commercial crawler of Sysomos (www.sysomos.com), which crawls social media, blogs, news, media and forums on a continuous basis. The input is any Boolean combination of keywords (and, or, not). Selecting the right search terms is important to avoid inclusion of noise. The output is a set of posts, tweets, and articles containing the specific search terms. Where possible,

---

[1] Notice that customer insights are also key for the conception of new propositions, but that is outside the scope of this paper.

geolocation is provided. We also incorporate reviews from app stores (e.g. Google Play and iTunes) if appropriate.

Concept detection and sentiment analysis

We use Bitext (www.bitext.com), which provides an API that receives as input the set of retrieved "items" and as output provides:

- The concepts mentioned in the items
- A set of possibly multiple opinions that constitute the items. E.g. a tweet or news item may contain several opinions about different concepts
- The neutrality or degree of tonality of the opinions (how positive or negative)
- The concepts the opinions are about
- The phrases used to express the tonality of opinions (sentiment)

Bitext applies semantic and linguistic technology to perform those tasks. We currently use it for Spanish, English and Portuguese. Sentiment is assigned to opinions based on dictionaries annotated with tonality, which we have tuned for "digital products" (e.g. in the world of digital products, "cheap" is usually something positive, whereas in general, it can be both positive and negative. Out of the box, Bitext's technology is about 70% accurate, and after tuning to the digital domain this is increased to 80%-90%. Most of the time, irony is not interpreted correctly.

Figure 1 Concept cloud representing how sentiment about objects is expressed. Size represents number of opinions; colour represents tonality as in next figure.

Visualisation

For visualisation, we use Tableau (www.tableau.com), which is an easy-to-use (both for development and for viewing) tool for building interactive dashboards. Interaction allows users to filter for viewing only negative or positive comments, or for different languages, to drill down into more detail, and to always review the original content. For each product or service we monitor, we use the following dashboards: Mentions (*by date, source, location, language, most active users & influence factor, most shared content*), locations (*geography of mentions – about 60% of the mentions*), concept clouds, sentiment (*concepts triggering opinions, phrases expressing sentiment, degree of tonality, date*), app store review analysis. Figure **1** shows a concept cloud of how tonality is expressed, and Figure **2** shows the breakdown of tonality of the opinions.
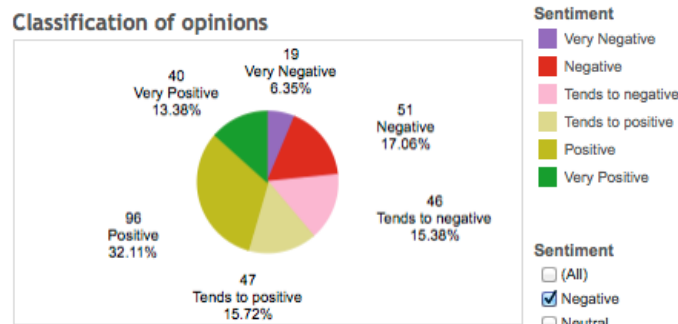
Figure 2 Distribution of tonality of opinions detected in the mentions.

Conclusions & learnings

We have learned that –apart from the typical reputation tracking that social media analytics is used for- it is a valuable tool for getting quick and economic customer feedback and insights for products. The tool is able to detect specific issues people complain about, such as for example customer care quality, price and pricing issues, registration process, and specific product features like crashes, unclear interfaces, battery drain, etc.

One thing business users see as very positive is the fact that the tool is available from day1 of launch, which enables quick responses to typical overlooked product issues, and complements the internally available product KPIs such as downloads, registrations, active users, etc. In general, we see that in the early days after commercial launch, comments are mostly positive reflecting the fact that most are announcements and promises of the great features of the product. Over time, more and more feedback comes in based on actual usage of the product.

Regarding the 70%-90% accuracy of the sentiment analysis software of Bitext, we have learned that when there are thousands of mentions per month, this does not cause a major problem. The dashboards mostly show aggregated information and the main trends, concerns, issues, etc. come out clearly. However, when there are less than 100 mentions a month, especially false positives (e.g. something seen as negative which in fact is not or vice versa), harm the credibility of the tool towards business users.

A final learning is that for some business owners, it is not easy to deal with a lot of negative feedback. And the fact that it is so easy to get feedback and that it is based on publically available knowledge makes it harder to "hide" the insights. This is however above all a cultural issue. In the lean, digital world, negative feedback should be embraced and taken as an opportunity to quickly improve products based on real customer insights.

In this brief paper, we discussed the opportunity for telcos (and in general, large enterprises) to take advantage of social media analytics as a valuable and economic tool for obtaining customer insights. This is however just one step in the journey to become a full digital telco, which eagerly listens to any relevant external and internal data source about customer and markets, including internal product data, call centre data, open data, paid-for-data, analyst data, screen-scraped data, and APIs.