

We tried hard, we got little. Things that didn't work, but won't stop us from trying

Mikel Bilbao (EUSTAT – Basque Statistics Office, Spain)

mikel@eustat.eus

Abstract

Dissemination technologies evolve at a very fast pace. We feel the urge to catch the wave, to stay relevant and up with the times. That's a difficult task and prone to errors, mistakes, misjudgments or, simply, excessive expectations.

Going out the beaten path, your beaten path, involves risk. In order to minimize it, we offer our experience in failure, so that others can learn from our mistakes. That's not to say, by any means, that other organizations shouldn't try themselves or there are not success stories involving the same technologies. Sure there are, we know some, actually.

Our exposition is intended for other to avoid our mistakes or take in account the limits of each experience, which are this four: Apps, Open Data, Facebook and Active web/social monitoring.

For each of them, we'll present, what we expected, what we did, the cost/effort involved, what we got (or didn't get), alternatives (that/if we know of) and what we plan for the future.

CONFERENCE OF EUROPEAN STATISTICIANS
Workshop on Statistical Data Dissemination and Communication
25-27 June, Cork, Ireland

We tried hard, we got little. Things didn't work, but that won't stop us from trying

Mikel Bilbao (EUSTAT – Basque Statistics Office)

mikel@eustat.eus

Paper

12/06/2018

Introduction

Users of data from official statistics offices are far from uniform. From casual citizens in need of a piece of data to settle a pub argument, to professional power users requesting microdata to perform their analysis, the range is extremely variable.

Dissemination technologies, furthermore, are also multiple and changing by the minute.

In an ideal world, we should try to match user segments to dissemination channels, but too often we just go with the trend and use them (all or part) indiscriminately.

At the same time, new technologies are always hyped as the “next big thing” in communication and the Swiss Army knife that’s going to get all of our problems solved and all of our users served and satisfied.

The truth is that the communication landscape is more often than not a minefield, with the biggest traps being, in order of importance: our own mistakes or misunderstanding of the technologies, the limits of such technologies, and an exceeding of expectations.

Over these next pages we will see how these three "mines" worked together or on their own to lead us to disappointment in four fields: Facebook, Active/web social monitoring, Apps, and Open Data.

An entire paper could certainly be written on each one, individually. Above all, Open Data is a hot topic in need of plenty of thought and action. So, think of this paper merely as an excuse to talk and dig deeper into the topics.

It is our hope for others to learn from our mistakes and, with luck, for others to point us in the right direction.

0. First things first: who are we?

One of the key variables involved in the success of our plans (and one that tends to be obviated) is context. Who are we, who are our users, how many are they, competitors. Sometimes context is the main –if not the only- reason for success or failure of a given action.

The Basque Statistics Office is a regional agency dedicated to producing and disseminating official statistics in the Basque Country.

The Basque Country, despite its misleading name, is an autonomous community (region) of Spain, with a population of two million people, and bilingual (Spanish/Basque).

Two million people are, then, our natural maximum number of users (and re-users).

We are not alone producing and disseminating official statistics. We face competition from the INE – Instituto Nacional de Estadística (Spanish National Statistics Office), both in relevance and the dissemination of data itself.

1. Facebook

1.1. In context

Facebook is –and already was when we approached it- the biggest social network on the Internet. Two thousand million active users flock to Facebook as of 2018Q1.

1.2. Our context

It seemed only natural to have a Facebook page when we approached social media. At a time when it wasn't too clear what a page, a personal profile or a group were and what was best for us, we struggled to set up two different pages, one for each official language in our community. Soon afterwards, the Basque Government dictated a social media presence policy for all of its departments and we had to start again from scratch, with only one bilingual account and a graphic design common to all government, thus, we lost the work already done, even our own logo.

1.3. What we expected

To be where the people are, where the debate is.

People are already discussing statistics and they are starving of data. Go there and provide. They will be grateful and your popularity will rise up to the clouds.

Or so we were told.

1.4. What we did

We went there, of course.

We set up our Facebook page (<https://www.facebook.com/eustat>) and began filling the wall with our scheduled press releases and events announcements. We hashtagged main keywords. We followed other institutions. We also added graphics and links where it was due (most of the time).

1.5. What we got

In a nutshell: very little.

Some figures to get an idea of the extent of the failure:

- As of today, we have just **262 followers**.
- **6 people** have rated our page (1 of them a Eustat employee).
- Our most successful publication this year reached a mere **190 people**, and it is a clear outlier; the usual reach ranges from 0 to 20.
- **Reactions** to publications (likes, shares,...) **are usually non-existent** for everyday posts (those announcing the availability of new statistics) and slightly better but still very poor when the post is about an event (a course, for example).
- **465 visits to our website** originated from Facebook this year so far. The whole of Facebook, not just our page.

1.6.- What went wrong

We believe that those 465 users visiting our website coming from Facebook shed light on the main problem here: Facebook is not the right channel for us.

Apparently, people are not discussing statistics on Facebook now and, as far as we know, they don't seem to be eagerly waiting for official statistics.

Besides, our own major flaw has been not "joining the conversation". We've been afraid of the never-ending nonsensical arguments so common on social media and believed that it was not the place for us to be. Therefore, we only made announcement posts and engaged in conversation only when directly addressed.

Finally, earlier this year, Facebook changed its behaviour, prioritizing people over businesses, and personal profiles over pages, making it even more difficult for us to reach an audience.

1.7.- Effort/achievement ratio

The effort involved in this action is very low, meaning that even when we consider this channel a failure every single "friend" is a success, even though they are few. So, all things considered, we can say the effort/achievement ratio is high. High enough.

1.8.- Plan of action

As mentioned, the effort is so low we can afford to maintain our presence on Facebook even when the number of users is also low.

And, we'll stick to Twitter, which is working very well for us.

2.- Active web and social media monitoring

2.1.- In context

Active monitoring of internet and social media is a tool for evaluating the presence of a person or organisation, both quantitative and qualitative.

There are a number of solutions offering this service; dozens, in fact. They usually scrap the public web/social media in search of mentions of predefined chains of characters (trademarks, names of products, even competitors' names and products) and also provide a tentative, far from perfect, contextualisation to find out whether a mention has been positive, neutral or negative.

2.2.- Our context

In Eustat we've been conducting user satisfaction surveys, newspaper scrapping, and web analytics for years. When we discovered the existence of these services we believed they could be very valuable within the same course of action.

2.3.- What we expected

It was our intention to complement the data we already gathered with web analytics.

Analytic tools give a lot of valuable information but they relay in actual traffic reaching our website. We know how many users visited our web and where they came from. That's a lot, but what's missing here are plain text mentions and links that nobody clicked on.

2.4.- What we did

We hired a provider for a three month pilot, which gave very good results. We thus extended the action for a whole year of brief weekly and deeper monthly reports.

2.5.- What we got

The first month was a success. We found an important complaint about one of our surveys in a blog post and as it was fair and reasonable but never reached us through an official channel, we believed the monitoring tool had served a very important purpose.

It just so happens that this was the first and last real achievement. For the rest of the year we only found neutral mentions with little real value, most which were in newspapers that had already been spotted by our journalist during his daily routine.

Neutral mentions are good, as we aspire to be neutral, but they bear little value for knowing our users and improving our services. Negative remarks are actually more useful.

2.6.- What went wrong

First of all, there is an inherent limit to this technology: it cannot reach private areas, and on social media, specifically on Facebook and LinkedIn, private areas are huge, even bigger than public ones.

There is a second limitation important to us: so far we don't know of any service of this kind which is able to monitor in the Basque language.

On top of this, we found out that the services we used were far -very far- from perfect at web scraping. Way too often we faced the fact that mentions we knew about hadn't been reported.

2.7.- Achievement/Effort ratio

The effort achievement/ratio is low.

The price of the contract was not high, but it was certainly a waste of money considering the results.

2.8.- Plan of action

We stopped the action but didn't abandon the idea completely.

We set up a free account on one of these services (Mention) and we are keeping an eye on this technology, waiting for it to improve results.

In the meantime, if we manage to hook something it will be welcome.

3.- Apps

3.1.- In context

The world is going mobile, there's no denying the fact. Every graph we are presented with is more than eloquent. It seems only natural, then, to enter the ecosystem of the mobile world.

3.2.- Our context

Eustat, as an organisation, has been always vigilant of new technologies for every stage of the statistical task. When the "age of the apps" came to a world where "there is an app for that" we felt it necessary to have an "app for statistics".

3.3.- What we expected

We thought an app should be more comfortable to use on a mobile device and would help us reach people on the move or out of the office more easily.

Besides, as we were to be among the first (if not the first) statistics offices in Spain to have an app, we expected a bit of a buzz about it.

3.4.- What we did

We built an App for Android, iOS and Windows Phone.

The app was an abridged version of the web, giving access only to the main statistics: population, economic trends and a tool for CPI calculations.

What set the app apart from the mobile version of our web that we had at the time was that it can be used without an internet connection and it allowed the user to set up alarms for statistics release dates.

3.5.- What we got

Since we published the apps in November 2013, it has been installed **536** times on Android and **61 remain active**; on iOS, only **4** remain active from **340 total installations**.

3.6.- What went wrong

We're not sure, we can only guess. Some ideas:

- .- The world goes mobile but statistics users not so much. Our product is still mainly consumed in offices and schools/universities, meaning PCs/Laptops, and not so much on mobile devices.
- .- Barely anyone needs a non-connected tool. That is not a real advantage nowadays.
- .- Being connected, a restriction in content is a clear drawback. Users want everything.
- .- When connected and wanting access to all content, a well-designed responsive website is a lot more useful than a crippled app.

3.7.- Achievement/Effort ratio

The achievement to effort ratio is very low. Developing apps is an expensive affair and from the results **offered we don't believe it's worth it**.

3.8.- Plan of action

We got rid of the windows phone app, as the platform is a zombie nowadays, and maintain the Android and iOS app without investing in updates.

Eustat -the data gathering area, not the dissemination area- has developed an app for respondents to the Time Expenditure Survey. Maybe this is a field of action more suitable for an app. We'll see.

4.- Open Data

4.1.- In context

Ever since the Obama administration gave protagonism to Transparency and Open Data, the majority of public administrations have taken it on board as their own.

The main purpose of open data is not direct dissemination, but reusing, republishing, with the main target not being citizen or organisation end-users, but developers, re-users.

4.2.- Our context

The Basque Government caught the wave and made Open Data compulsory for all offices, setting up a general repository for this purpose.

Statistics, for the most part, have been always open. Free access and reusability have been a custom in statistics offices for years. Public service has always been paramount in this respect. This is why converting statistics into "open data" was almost trivial.

Almost trivial, because our production and dissemination systems were independent from those of the Government and an effort was needed to link both databases.

Besides, we lacked some metadata that had to be added.

4.3.- What we expected

We expected exactly what we were promised: re-utilisation of our statistics. Apps, web apps, web services built upon our datasets.

4.4.- What we did

We linked our dissemination system to that of the Government, added the necessary metadata to our datasets, and published most of our statistics in no proprietary formats (CSV).

At the same time, we made an API to our databank available so that re-users could take advantage of another way to put our data to good use.

4.5.- What we got

There are 58 apps/re-utilisation services registered on the open data portal of the Basque Government, not just for statistics, but for every kind of data..

This alone is a very poor outcome, but when it comes to statistics there are only two, both being in-house-developments, neither of which is directly related to the Open Data initiative. Both existed before it.

We can confidently say, then, that there is no one single app, service or development of any kind set up in the Basque Country with open statistics.

4.6.- What went wrong

This is a hot topic, most likely in need of a far greater scope than this paper: is Open Data working as intended so far?

To begin with, it's a very difficult question to ask. The most accurate answer is probably "we don't know". When discussing the open data portals we normally see the number of open datasets as a measure of success. Secondly, because there are usually few of them, the number of registered apps. But we never get to see final user figures. How many people are benefitting from the aperture of data? Statistics, specifically.

We barely know. There's no easy way to find out.

Coming back to our own case, we believe we lack the critical mass to produce re-users. There are two million of us, scarcely 10 percent of whom need statistics on their everyday life. Disseminators are even more scarce. Among these, very few, if any, are developers. And among the developers, how many have the incentive and time to use our data? That is, our data instead of data from other sources (e.g. INE). The answer, we think, must be very close to zero. At least that is what is shown from the number of registered apps on the Government data portal.

Besides, it's a proven fact that most developers prefer geolocated and/or real-time data to work with (e.g. traffic, weather,...). They are probably more profitable, and statistics lack these two characteristics.

4.7.- Effort/achievement ratio

We made a big effort, in the beginning, to be able to link databases with those of the Government. Once established it works almost automatically.

The cost/effort is, therefore, diminishing with time.

4.8.- Plan of action

We'll continue publishing all of our datasets on the open data portal and we'll be keeping an eye on the development of the technologies implied; everything relating to "linked open data" is particularly interesting.