

**PUNTU  
EUS**

**OBSER  
VATORY**

***Analysis of the  
Basque Country's  
Internet Situation***

**2017**



***gure domeinua***

## LEGAL NOTICE

The PuntuEUS Fundazioa is the author of this report and the owner of the various content, references, texts, photos, images and logos, as well as of the design and layout of the report.

As a result of this ownership, the PuntuEUS Fundazioa offers the user the Creative Commons licence. This type of licence offered is as follows: *Attribution 4.0*. Licence covering all the various content referred to in the above paragraph.



# Contents

<b>1. FOREWORD.....</b>	<b>4</b>
1.1. Methodology .....	4
1.2. Information Sources .....	6
1.3. Collaboration .....	6
<b>2. SITUATION OF THE .EUS DOMAIN .....</b>	<b>8</b>
2.1. Geographical Distribution of the Domain .....	8
2.2. Typological Distribution of the Domain .....	10
2.3. Penetration of Basque-language Domain Parking.....	11
2.4. Distribution of the Dominant Languages.....	12
2.5. Density of the .EUS Domain in the Basque Country.....	13
2.6. Domain Quality.....	14
2.7. Distribution between men and women.....	16
<b>3. THE INTERNET IN THE BASQUE COUNTRY .....</b>	<b>17</b>
3.1. Distribution of the main TLDs.....	17
3.2. Number of active websites .....	18
3.3. 2016-2017 evolution .....	19
3.4. Internet Domain Density in the Basque Country.....	20
<b>4. THE BASQUE LANGUAGE ON THE INTERNET IN THE BASQUE.....</b>	<b>21</b>
4.1. The Presence of Basque in the main TLDs.....	21
4.2. Density of Websites with Basque in the Basque Country .....	23
4.3. Presence of Basque on Twitter .....	24
<b>5. PUNTUEUS FUNDAZIOA.....</b>	<b>25</b>
5.1. PuntuEUS 2020 .....	25

## 1. FOREWORD

The PuntuEUS Observatory is an initiative of the PuntuEUS Foundation and its aim is to measure the situation enjoyed by Basque on the Internet. The analysis conducted by this Observatory is made once a year, and the result is posted at <https://www.domeinuak.eus/en/observatory/>. Here, the results of last year's study together with this year's are shown.

The PuntuEUS Observatory analyses the presence of the Basque language on the Internet and also the presence of the Internet across the Basque Country, quantitatively and qualitatively, by measuring three spheres:

1. **Situation of the .EUS domain:** The main features of the websites that use the .EUS Domain are analysed.
2. **The Internet in the Basque Country:** Number and trends in the domain names registered in the Basque Country by the main TLDs.
3. **The situation of the Basque language on the Internet:** The presence of Basque on the Internet, in the main TLDs and, in particular, in the websites that use the .EUS domain.

The information published in this report is based on data for 2017, and reflects the situation of the Internet in the Basque Country in 2017.

### 1.1. Methodology

This study not only analyses how far the .EUS domain and the main Internet domains have penetrated the Basque Country, it also examines the presence of Basque and the other dominant languages.

To do this, the entire content of the websites corresponding to the domains was analysed and classified by language. That way, it will be possible to know how much content in each of the Basque Country's domains is in Basque, Spanish, English, French and other languages. Two strategies were used to conduct this domain analysis:

- **Domain level crawling:** First of all, the html website content corresponding to the domain is automatically downloaded using crawling techniques, taking those created in JavaScript into consideration as well. With this crawling process the redirects that may be within a domain are managed in an intelligent way. Parking pages

are also detected and blocked. After this crawling process, the text is extracted from the gathered html content and the languages in it are automatically identified by means of language model statistics. The language model used is capable of identifying all the texts that may be in a multilingual text. This strategy produces a lot of traffic when large websites are processed. That is why it is only used for processing low-content websites.

- **Domain level web searches:** The idea behind this strategy is to make use of Web browsers (Google, Bing, etc.) in measuring how much presence a language has in a website. By running a search comprising the most significant words in a language (language filter words) in a specific domain in the web browsers, we can calculate the number of content items in the language. That way we avoid having to download website content. That is why we use this strategy to process large-sized websites. We do not apply this to small-sized websites because many websites with little content are not fully indexed in the browsers. To confirm that the language filter words have functioned correctly, we classify -by means of statistical models- the first results returned by the browsers according to language; the purpose is to check the number of pages returned by the web browsers, in accordance with the language filters.

It goes without saying that this measuring process that needs to be completed is very complex, and that is why even if the precision of the two strategies is very high, there is a margin of error. At the end of the day, the measuring process comprises a number of steps, and each of these steps has a small error rate, which accumulates throughout the whole chain. According to our calculations, the precision of the results of the measurement is between 70% and 80%.

We are incorporating improvements into the system to reduce this error rate in the yearly measurements, and these improvements are making their presence felt in the results. In the 2017 analysis, for example, more redirects were taken into consideration in the crawling process, and many domain parking pages were automatically blocked. We have, however, used a new statistical language model to be able to identify pieces of texts in Basque in multilingual texts, and we have also taken websites with very short texts into consideration.

In addition to these automatic strategies, the presence of Basque in the case of the .EUS domains was measured manually, thus achieving a higher level of precision.

With respect to domain distribution, the following domains were analysed:

- gTLDs or generic Top Level Domains: .EUS, .COM, .NET, .INFO, .ORG and .BIZ
- ccTLD or country code Top Level Domains: The .ES and .FR domains were analysed. In the case of the .EUS domain, various pieces of data needed to be able to make language analyses are not public, and this is indicated throughout the corresponding analysis.

## 1.2. Information Sources

References and information sources used for the project:

- <https://ntldstats.com/country/ES>
- <http://www.statdns.com/>
- <https://webhosting.info/domain-name-statistics>
- <http://www.eurid.eu/>
- <https://stats.centri.org/gtlds>
- <https://opendata.afnic.fr/>
- <http://www.umap.eus>
- <http://fundacio.cat/es/observatorio/domini-cat>
- <http://www.dominios.es/dominios/es/todo-lo-que-necesitas-saber/estadisticas>

## 1.3. Collaboration

The PuntuEUS Observatory project is an initiative of the PuntuEUS Foundation. Nevertheless, it is clear that to be able to run such a project, the collaboration of other players is essential. We were assisted by the following players when conducting the analysis and presenting the data:

- Fundació PuntCAT
- AFNIC
- EURID
- Public entity Red.es
- The Elhuyar Foundation contributed towards the technological side needed to analyse the Basque language content.
- The cartography and spatial data of the Basque Country: Gaindegia & Euskalgeo.

- **Analysis of Basque presence on Twitter: Umap.eus / CodeSyntax**

**This project has received funding from the following institutions:**

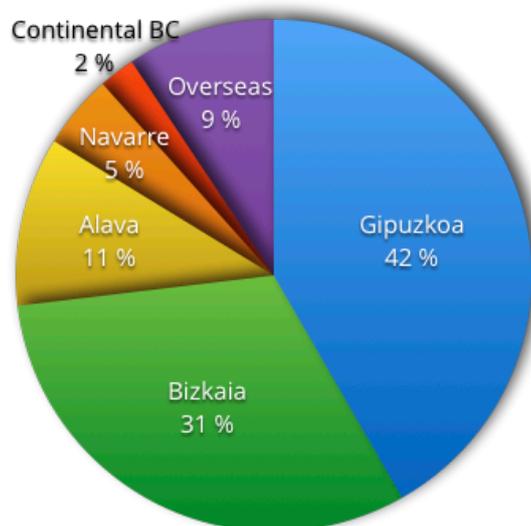
- **The Directorate for Language Equality of the Chartered Provincial Council of Gipuzkoa**

## 2. SITUATION OF THE .EUS DOMAIN

### 2.1. Geographical Distribution of the Domain

The .EUS Internet domain does not make any reference to a specific geographical area, but to a cultural and language community. So the domain is spread across the world with 9% of the registrations having been made outside the Basque Country so far. That points to the importance of the domain's widespread acceptance.

After analysing the data on the registrations made in the Basque Country, it became clear to us that the domain is spread across the Basque Country, but there are significant imbalances between one province and another. These data display greater similarity with the conclusions of the Sociolinguistic Survey and Map produced by the Government of the Basque Autonomous Community and the Euskararen Erakunde Publikoa (Public Institution for the Basque Language) than with the demographic distribution of the provinces, certain exceptions notwithstanding: Navarre, in particular, and the Continental Basque Country, to a certain extent, should have a greater weight.



*Figure 1: Geographical distribution of the .EUS Domain*

	Gipuzkoa	Bizkaia	Araba	Navarre	Continental Basque Country	Overseas
% of domains	43%	30%	11%	5%	2%	9%



## 2.2. Typological Distribution of the Domain

The typology of organisations that register the .EUS domain points to the general and diverse value of the domain.

As is usual on the Internet, the main users of the .EUS domain are companies. In addition to them, it is possible to see that the remaining segments of our society are also widely represented in the .EUS domain: public institutions, education organisations, organisations in the field of the Basque language and culture, private individuals, mass media, sports organisations and others, too (trade unions, political parties, etc.).

What is noteworthy in this section is the fact that the growth trend in the .EUS domain is taking place in companies, in particular, while the segment corresponding to the public institutions is decreasing in the same proportion; this points to the normal penetration of the domain into society.

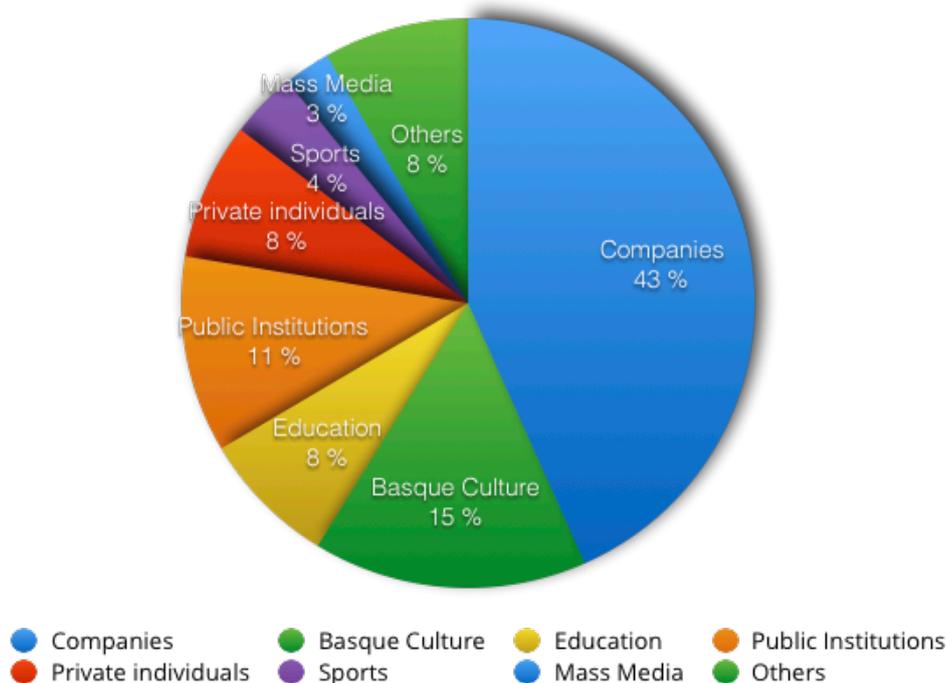


Figure 2: Typological Distribution of the .EUS domain

	Companies	Basque Culture	Public institutions	Education	Private individuals	Others	Mass Media	Sports
% of domains	43%	15%	11%	8%	8%	8%	4%	3%

### 2.3. Penetration of Basque-language Domain Parking

When one looks at the whole registered domain parking by the main TLDs\* in the Basque Country (.com, .org, .net, .es, etc.) it becomes apparent that the number of .EUS domains registered today is not very high.

But the unique features of the .EUS domain need to be taken into consideration, and on the basis of this, when analysing the market, we focussed on the websites that have some of their content in Basque.

That being the case, when comparing the number of .EUSs with the total number of websites among the main TLDs that have Basque, the domain still has room for growth.

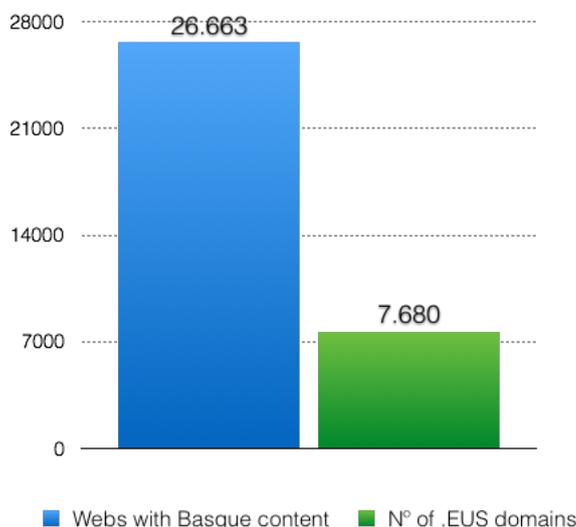


Figure 3: Number of .EUS domains with respect to the Basque language Internet

	Euskarazko edukia duten webguneak	.EUS domeinuzen kopurua
Domeinu-kopurua	26.663	7.680

\* The .EU domains are not included in this analysis

## 2.4. Distribution of the Dominant Languages

Four languages predominate among the websites using the .EUS domain: Basque, Spanish, French and English. The distribution of these languages can be seen in the figure below, which highlights the fact that Basque (90%) and Spanish (80%) are mostly used in the websites that use the .EUS domain. They are followed by English (55%) and French (21%) with a very small presence, and other languages (6%).

Basque and Spanish are on a par in this distribution and, to a certain extent, that is understandable, because most of the websites are multilingual ones and because many of them are located in the Peninsular Basque Country.

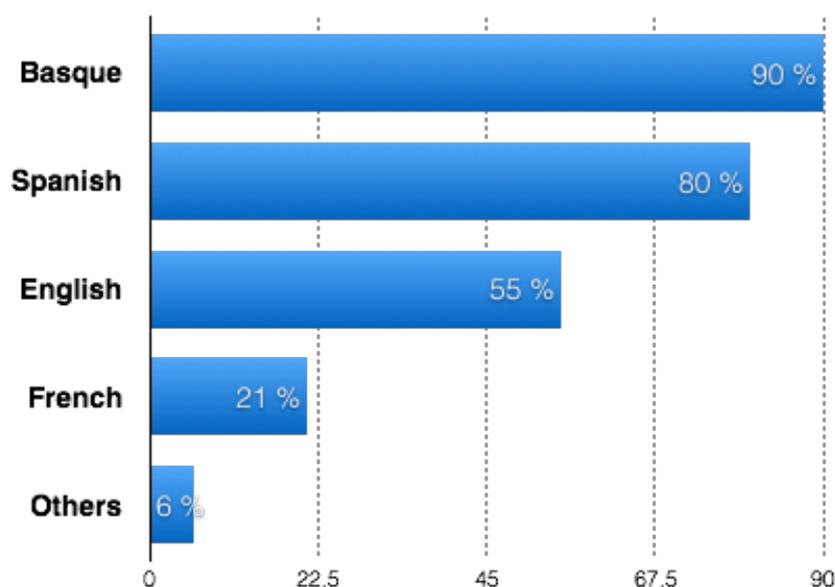


Figure 4: Language distribution in the .EUS domain websites

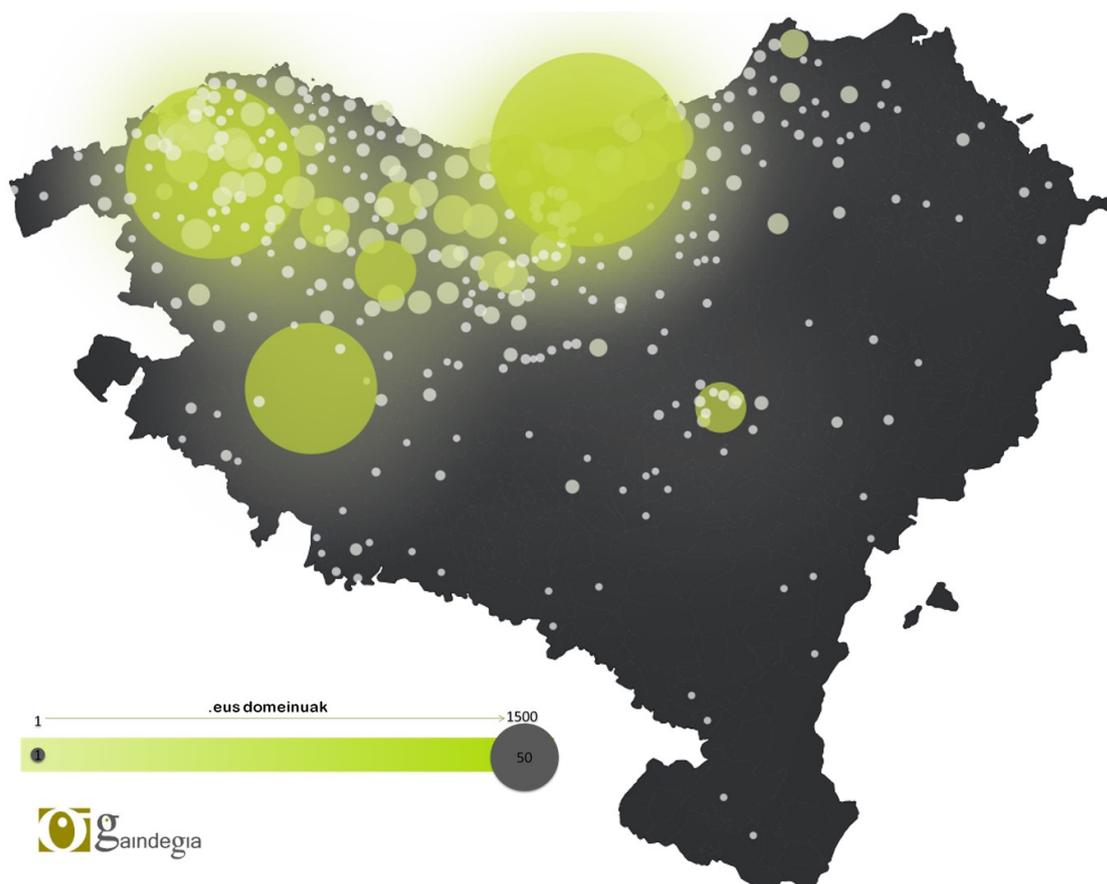
	Basque	Spanish	English	French	Others
% of Websites	90%	83%	47%	12%	11%

## 2.5. Density of the .EUS Domain in the Basque Country

The figure below shows how .EUS domain density is distributed across the Basque Country according to the geographical location of registrations. The circle size indicates the degree of density with respect to the number of domain names registered in each town/city.

There are no surprises in this result because the future of the Basque language can in fact be said to lie in the cities, in other words, in the areas with a greater number of inhabitants. Nevertheless, and in line with what has been seen in the previous sections, the low number of registrations in Navarre and the so-called BAM (area comprising Baiona/Bayonne-Angelu/Anglet-Miarritze/Biarritz) should be taken into consideration

An updated version of this map can be found at <http://mapa.domeinuak.eus>.

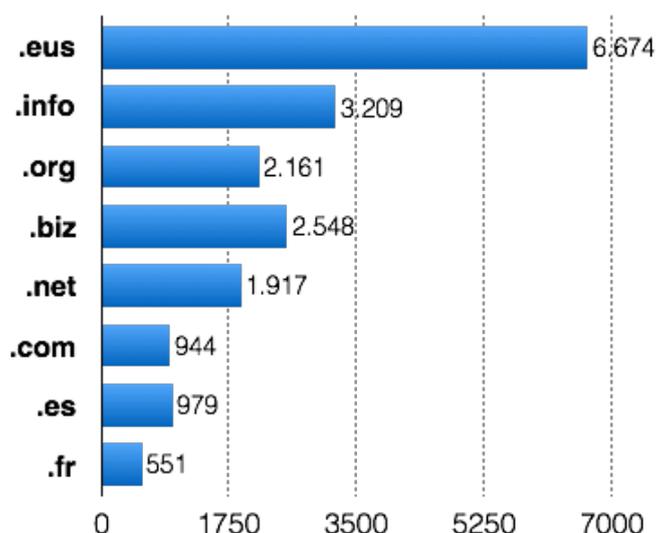


*Figure 5: Density of the .EUS domain names in the Basque Country*

## 2.6. Domain Quality

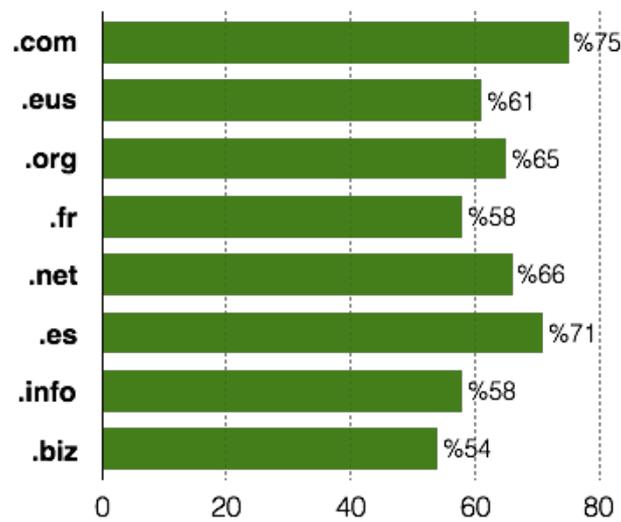
There are no specific criteria for measuring the quality of a TLD. Taking the data of the PuntuEUS Observatory as the basis, we have taken the following criteria into consideration when gauging the quality of TLDs.

- **Domain density:** The contribution made by each domain to the Internet. To measure this, the average number of documents in the websites of each domain were measured. The result clearly shows that among the predominant TLDs in the Basque Country the .EUS domain is the one that makes the biggest content contribution towards the Internet. So it can be said that the .EUS websites post a lot of content thus making a big contribution towards the Internet. That is important in our case because a large proportion of that content is posted in Basque as far as the .EUS domain is concerned.



*Figure 6: Domain density – Average number of website documents for each TLD*

- Number of active domains:** The number of active domains among the registered domains. Many domains are registered defensively, in other words, to prevent anyone else registering the same name, and then they are not subsequently activated. Out of the domains that a TLD has registered, it is important for as many as possible to be active. This criterion can be measured in the case of all TLDs and that is why a comparison is provided below. It can be seen that the .com is the most used among the main TLDs present in the Basque Country, both in terms of the number of registrations and also in terms of the domains that are active.

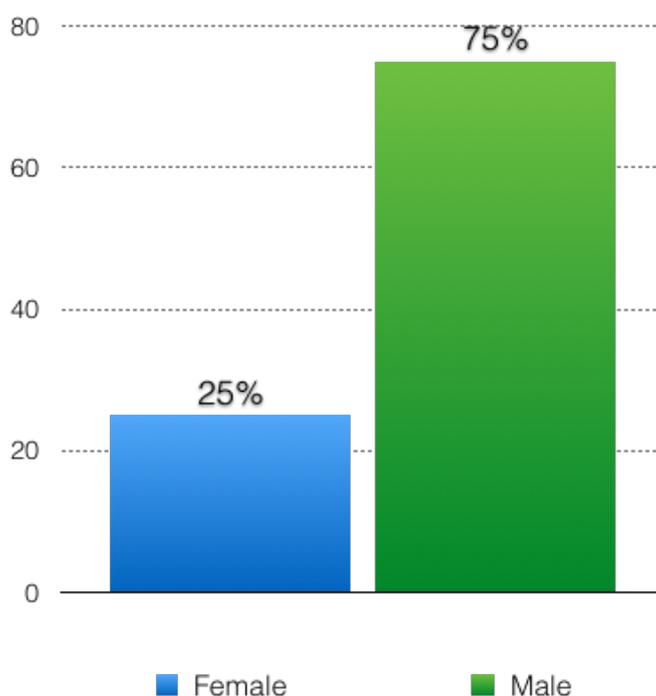


*Figure 7: The number of active domains in each TLD*

## 2.7. Distribution between men and women

According to the European Commission, the presence of women in the field of ICTs is small. [According to research published by the Commission](#), only 30% of the 7 million people who work in ICTs are women and that percentage is even lower in positions of responsibility. Only 19%.

In the area of Internet domains, the owners of .EUS domain names were analysed. Leaving on one side those registered on behalf of an organisation, the conclusion after analysing the cases of men and women owners is that there are fewer women (25%) than men (75%). So the differences can be described as significant in the registration and use of Internet domain names.



*Figure 8: Female and male registrations in the .EUS domain*

### 3. THE INTERNET IN THE BASQUE COUNTRY

#### 3.1. Distribution of the main TLDs

As regards the Basque Country's Internet market as a whole, .COM and .ES are the main TLDs, and together they account for 74% of the market.

The remaining TLDs have a small market share. Among the domain names registered in the Basque Country, so far only 3% are .EUS ones.

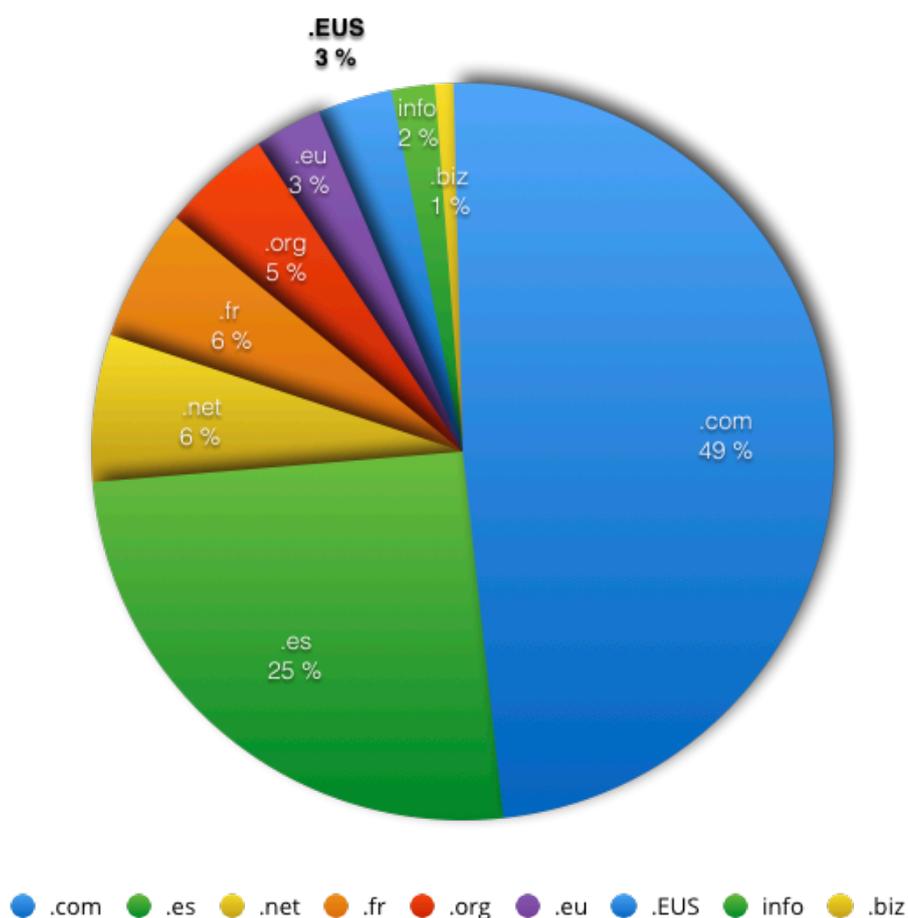


Figure 9: Distribution of main TLDs in the Basque Country

	.COM	.ES	.NET	.FR	.ORG	.EUS	.EU	.INFO	.BIZ
Market %	49%	25%	6%	6%	5%	3%	3%	2%	1%

### 3.2. Number of active websites

Apart from the domain name register, another piece of data that has to be taken into account is the activation level of these domain names. In other words, whether the registered domain name has published content or whether it has a website. In the Basque Country a total of 236,000 domain names were registered in 2017, and 161,837 of them had published some content, 68% of them, in fact

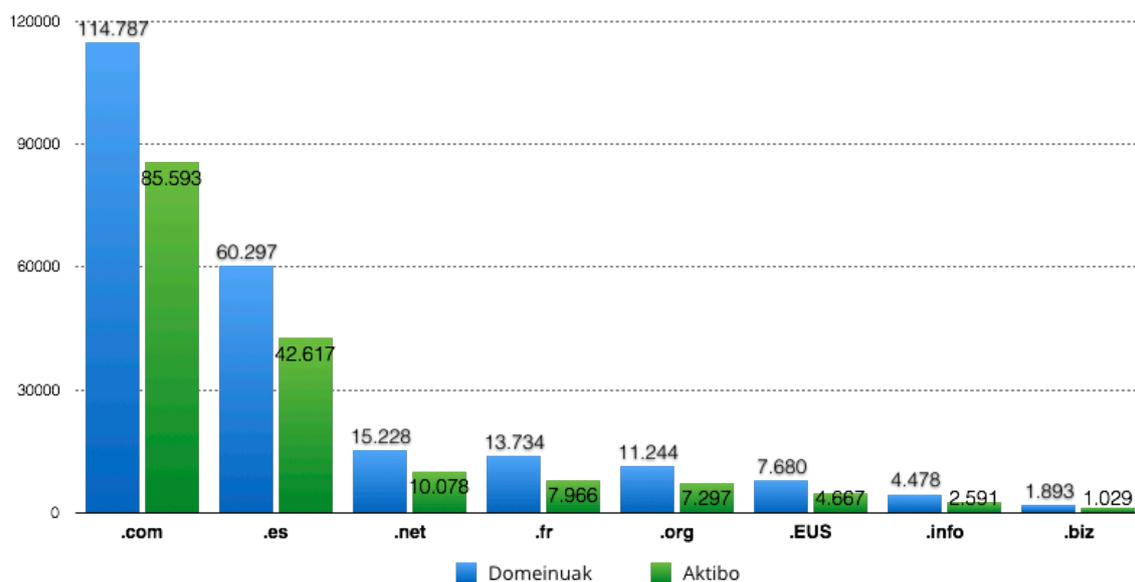


Figure 10: Number of active domains

	.COM	.ES	.NET	.FR	.ORG	.EUS	.INFO	.BIZ
<b>Number of Domains</b>	114.787	60.297	15.228	13.734	11.244	7.680	4.474	1.893
<b>Number of active websites</b>	85.593	42.617	10.078	7.966	7.297	4.667	2.591	1.029
<b>% Active ones</b>	%75	%71	%66	%58	%65	%61	%58	%54

### 3.3. 2016-2017 evolution

The .EUS domain has managed to maintain growth by achieving a growth rate of 22% in 2017; as the graph shows. But it is important to ensure stability as well as growth, and in 2017 the renewal rate of the .EUS domains was 95%, which is much higher than the rate usually enjoyed by Internet domains.

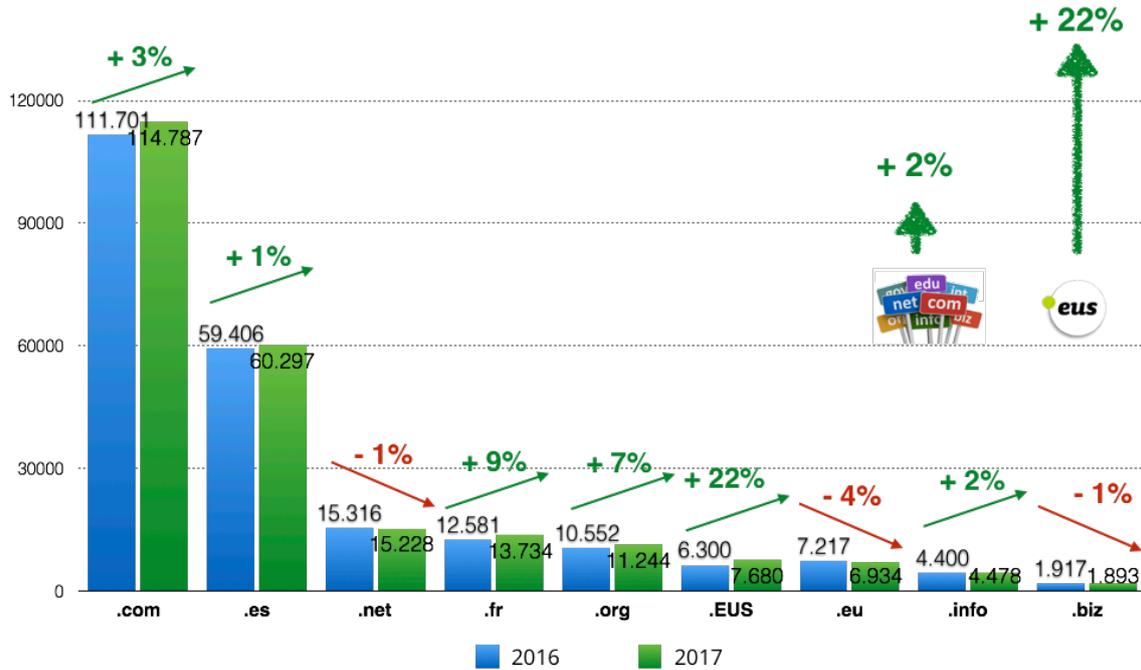
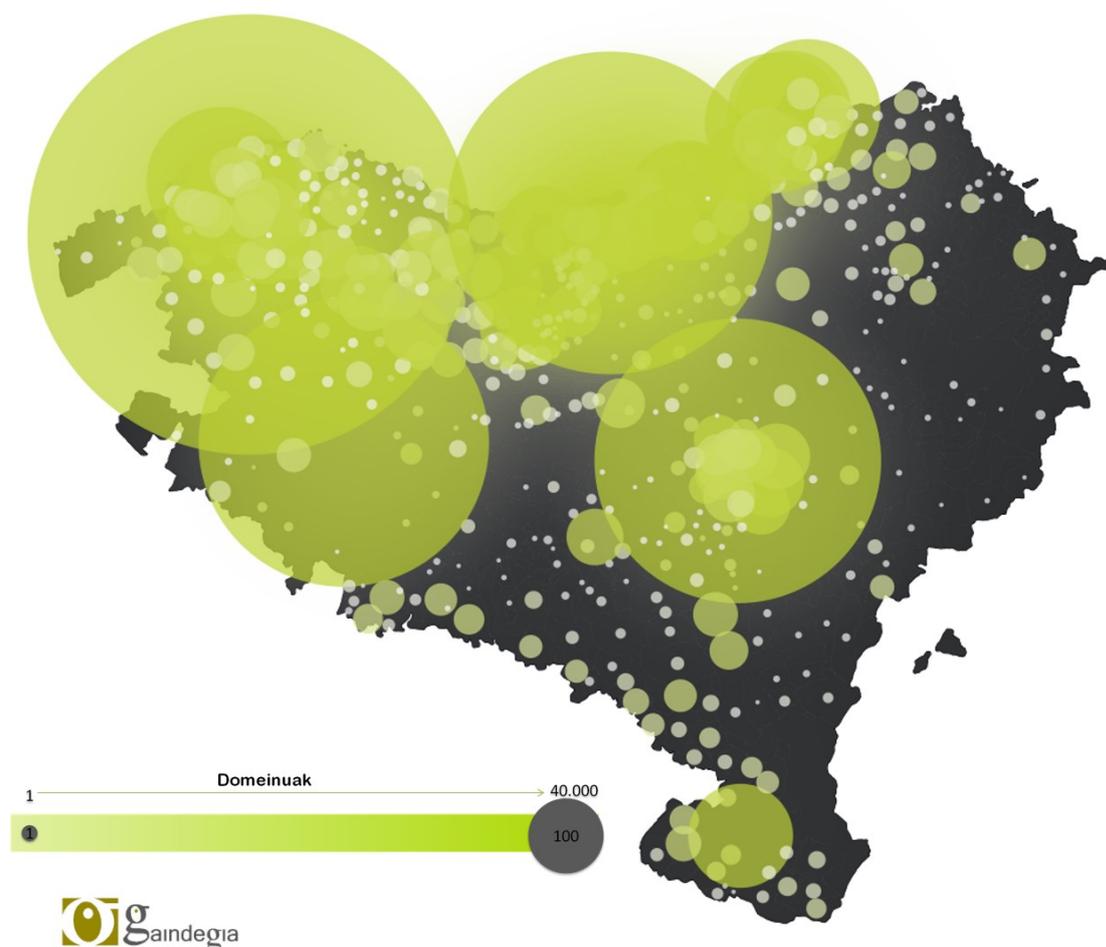


Figure 11: Evolution in the Domain market in the Basque Country 2016-2017

### 3.4. Internet Domain Density in the Basque Country

The figure below shows the density distribution of the main Internet TLDs<sup>†</sup> in the Basque Country; they are expressed in terms of the geographical location of registrations. The circle size indicates the degree of density with respect to the number of domain names registered in each town/city.



*Figure 12: Internet Domain Density in the Basque Country*

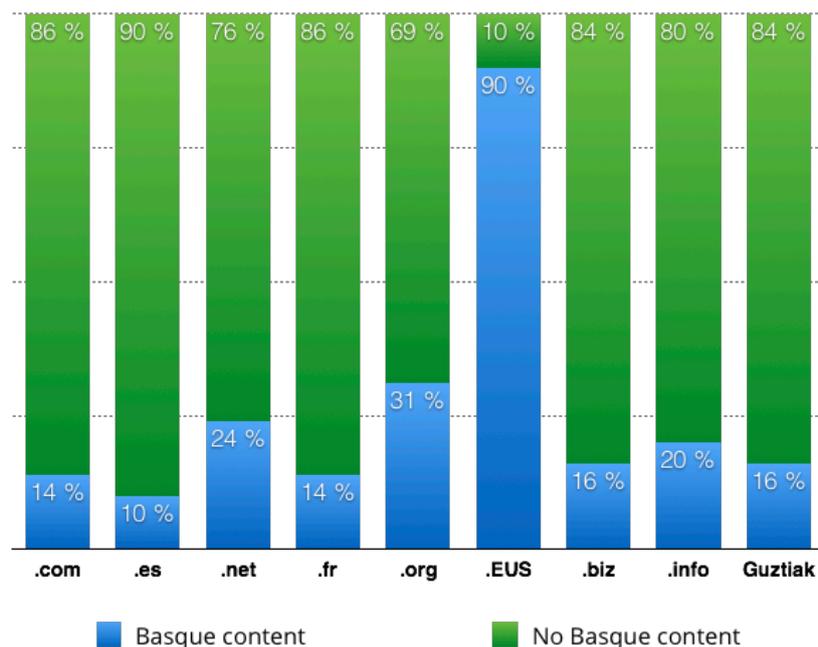
<sup>†</sup> Data on the .EU domain are not included in this analysis.

## 4. THE BASQUE LANGUAGE ON THE INTERNET IN THE BASQUE

### 4.1. The Presence of Basque in the main TLDs

When analysing the percentage of domain names that have content in Basque in the main Internet TLDs<sup>‡</sup> the .EUS domain clearly stands out as the TLD that uses Basque the most. The presence of Basque is prominent in the .EUS domains. In this case, the number of domains that have no Basque account for 10%, while the number of websites with Basque content is 90%. So the .EUS domain can be said to be the digital “lungs” of the Basque language.

As regards the remaining domains, the presence of Basque on their websites is less than 20% in most cases. The .ORG domain is the one with the most Basque with 31% of the websites in Basque but in the remaining cases the presence of Basque is very low.



*Figure 13: The Presence of Basque in the main TLDs*

.COM .ES .NET .FR .ORG .EUS .BIZ .INFO GUZTIAK

<sup>‡</sup> Data on the .EU domain are not included in this analysis.

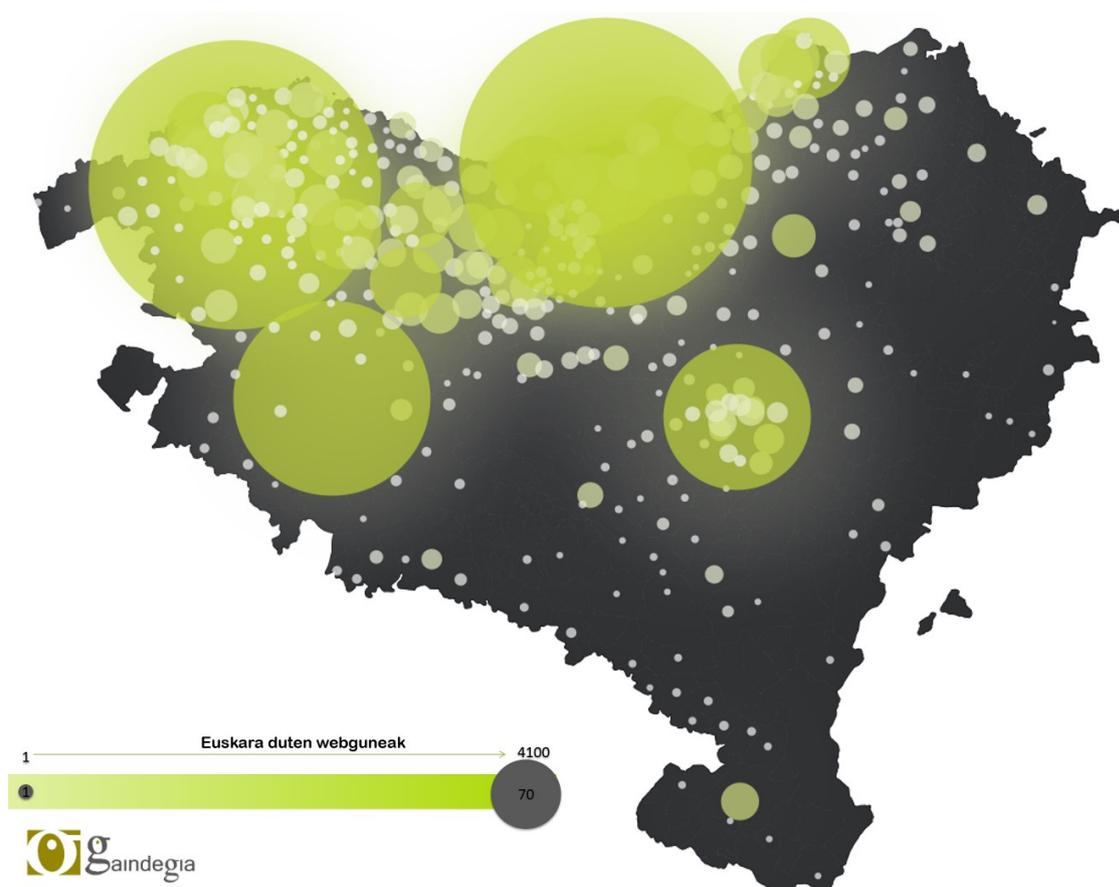
# PuntuEUS Observatory

Analysis of the Basque Country's Internet Situation 2017

% of Websites	14%	10%	24%	14%	31%	90%	16%	20%	16%
---------------	-----	-----	-----	-----	-----	-----	-----	-----	-----

## 4.2. Density of Websites with Basque in the Basque Country

The figure below shows the distribution of the density of websites<sup>§</sup> in the Basque Country that have Basque content; it is expressed according to the geographical location of registrations. The circle size indicates the degree of density with respect to the number of domain names registered in each town/city.



*Figure 14: Density of Websites with Basque in the Basque Country*

<sup>§</sup> Data on the .EU domain are not included in this analysis.

## 4.3. Presence of Basque on Twitter

In response to a request made by users, Twitter produced a Basque version in 2012 and since then there has been an option to use the Twitter interface in Basque.

UMAP is the website that gathers and filters the activity of Basque-speaking Twitter users. According to it, there were a total of 13,500 active users using Basque in 2017. As far as the tweets are concerned, these active users in the community sent 6,480,109 tweets (2,400,000 in Basque). The number of Twitter users writing mainly in Basque (more than 75% of the content in Basque) has been 2,2302, 17% of the total.

For further information: <http://umap.eus/>

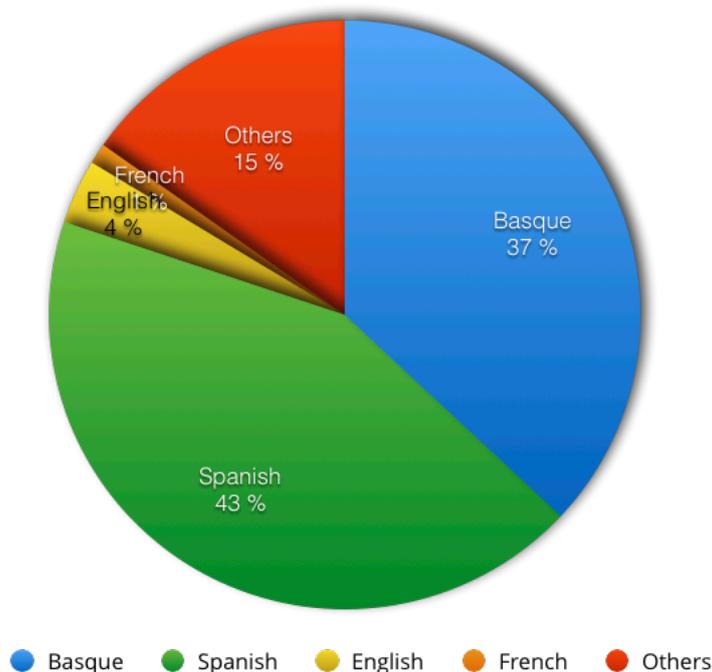


Figure 15: Language use among Basque-speaking Twitter users

	Basque	Spanish	French	English	Others**
Language %	37%	43%	1%	4%	15%

\*\* Other tweets: messages without language, bilingual messages or tweets in other languages

## 5. PUNTUEUS FUNDAZIOA

The PuntuEUS Foundation is a private non-profit organization, and its objective is the promotion of the Basque language and culture and the cohesion of the community of this language and culture.

The PuntuEus Foundation presented, in May 2012, the application for the approval of the .EUS domain and signed the contract with ICANN in December 2013. For this reason, the PuntuEus Foundation is the entity responsible for the control, administration and management of the .EUS domain.

The .EUS domain is the top-level Internet domain of the Community of the Basque language and Culture. It seeks to promote the Basque language and Basque culture and to build the Community of the Basque language and Culture. So .EUS is a tool which puts Basque on the same level as other languages; it will assist in the normalisation of Basque; and provides the country of the Basque language with international recognition.

### 5.1. PuntuEUS 2020

The PuntuEUS Foundation has drawn up its Strategic Plan 2016-2020 with the participation of the community. The strategic bases and aims for the coming years are laid down in this Strategic Plan, as are the challenges facing the .EUS domain and the Foundation.

The PuntuEUS Foundation has specified the following mission in this Strategic Plan:

**To make the Basque language valid, useful and visible in the digital world through the .EUS domain**

By developing this mission, we are aiming to fulfil the vision of the PuntuEUS Foundation as follows:

**Basque as a first-class language in the digital environment**

We have identified four strategic challenges to turn this into reality:

- **DIGITAL ENVIRONMENT:** To strengthen Basque use in the digital environment
- **DOMAIN:** To turn the .EUS domain into the number 3 domain in the Basque Country
- **COMMUNITY:** To achieve a broader, bigger and more lively community

- **BRAND:** To make .EUS the leading reference in the Basque Digital Environment

And the Foundation's values are our basis when it comes to tackling these challenges:

- **COURAGE:** Courageous when tackling challenges and innovative in working methods and results
- **COMMITMENT:** Committed towards Basque and the people of the Basque language
- **COLLABORATION:** Starting with the structure and in any initiative so that collaboration and transparency can be turned into the starting point and aim.
- **INCLUSIVENESS:** The .EUS community is diverse and we need to get this characteristic accepted and promoted naturally
- **PRESTIGE:** We have to ensure a successful outcome in whatever we do

The following graph sums up the 2016-2020 Strategic Plan of the PuntuEUS:







**gure domeinua**



**PuntuEUS Fundazioa 201**

**[behatokia.domeinuak.eus](http://behatokia.domeinuak.eus)**