

**PUNTU
EUS**

**OBSER
VATORY**

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

2016



gure domeinua

Contents

1. FOREWORD.....	3
1.1. Methodology	3
1.2. Information Sources.....	4
1.3. Collaboration.....	5
2. SITUATION OF THE .EUS DOMAIN	6
2.1. Geographical Distribution of the Domain	6
2.2. Typological Distribution of the Domain	7
2.3. Penetration of Basque-language Domain Parking.....	8
2.4. Distribution of the Dominant Languages	9
2.5. Density of the .EUS Domain in the Basque Country	11
2.6. Domain Quality	12
2.7. Distribution between men and women.....	14
3. THE INTERNET IN THE BASQUE COUNTRY	15
3.1. Distribution of the main TLDs	15
3.2. Number of active websites	16
3.3. 2015-2016 evolution.....	17
3.4. Internet Domain Density in the Basque Country.....	18
4. THE BASQUE LANGUAGE ON THE INTERNET IN THE BASQUE COUNTRY	19
4.1. The Presence of Basque in the main TLDs	19
4.2. Density of Websites with Basque in the Basque Country.....	20
4.3. Presence of Basque on Twitter	21
5. LEGAL NOTICE	22
5.1. Permission to use the Data	22
5.2. PuntuEUS Fundazioa	22

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 2016, and reflects the situation of the Internet in the Basque Country in 2016.

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:** Firstly, the HTML website content corresponding to the domain is automatically downloaded, taking those created in JavaScript into consideration as well. Then the text information is extracted from the HTML content and its language is automatically identified by means of language model statistics. This strategy guarantees an accuracy of 0.77. Even if this strategy is

implemented in parallel with the domain processing, it can lead to a lot of traffic when large websites are processed. That is why this strategy is only used for processing low-content websites.

- **Domain level web searches:** This strategy based on Internet search engines (e.g. Google, Bing) is used to process large websites because it generates much less traffic than the crawling process. However, many low-content websites are not fully indexed by the search engines. By running a search comprising the most significant words in each language in each domain, we can calculate the number of content items in the language. The number of pages that have significant words (language filtering words) in each language in the domain is provided by the Internet search engines. The first hits returned by the Internet search engines are automatically classified according to language to check that the filter words are working properly. This strategy guarantees an accuracy of 0.82.

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:

- Basque Autonomous Community Government Sub-Ministry for Language Planning
- 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.

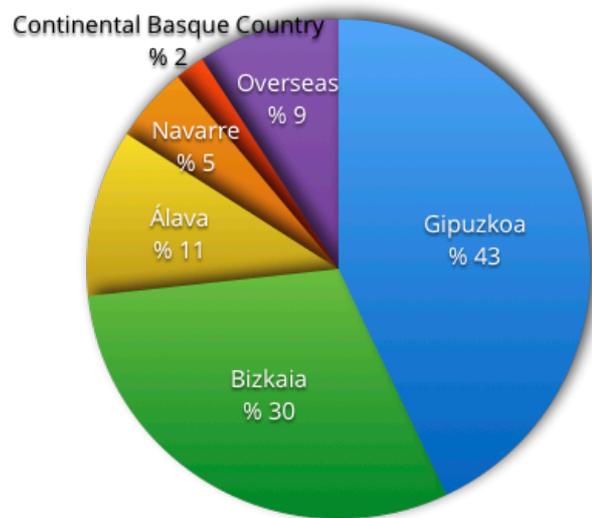


Figure 1: Geographical distribution of the .EUS Domain

	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.

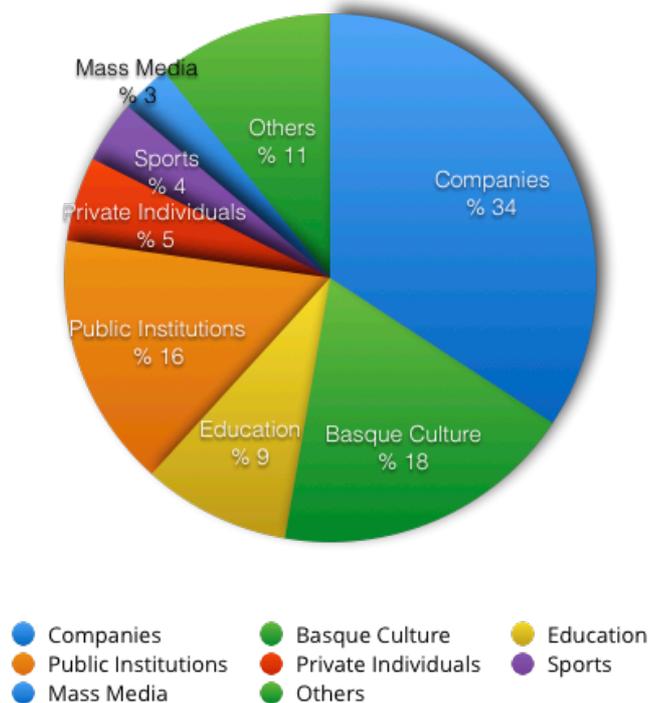


Figure 2: Typological Distribution of the .EUS domain

	Companies	Basque Culture	Education	Public Institutions	Private individuals	Others	Mass Media	Sports
% of domains	34%	18%	9%	16%	5%	11%	3%	4%

2.3. Penetration of Basque-language Domain Parking

When one looks at the whole registered domain parking by the main TLDs¹ 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 presence of our domain can be regarded as significant.

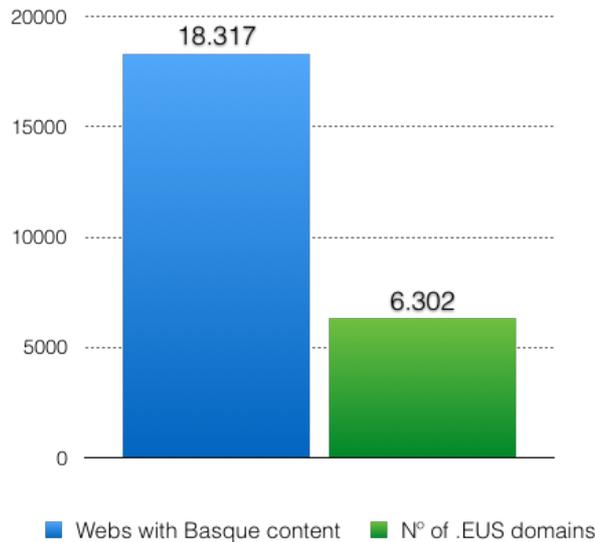


Figure 3: Number of .EUS domains with respect to the Basque-language Internet

	Number of websites with Basque content	Number of .EUS domain names
Number of domains	18.317	6.302

¹ 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 (83%) are mostly used in the websites that use the .EUS domain. They are followed by English (47%) and French (12%) with a very small presence, and other languages (11%).

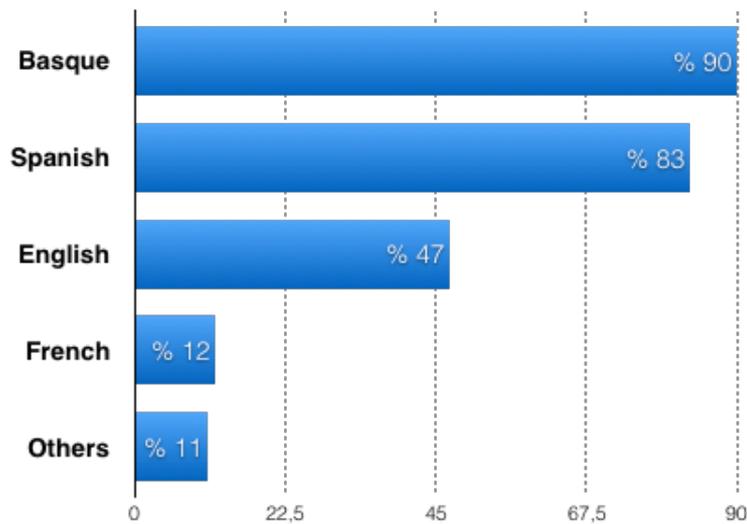


Figure 4: Language distribution in the .EUS domain websites

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

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. As can be concluded from these data, most of the .EUS websites are multilingual ones, the main language combinations being Basque-Spanish-English (29% of the websites) and Basque-Spanish (28% of the websites). Only 9% of the .EUS websites are exclusively in Basque.

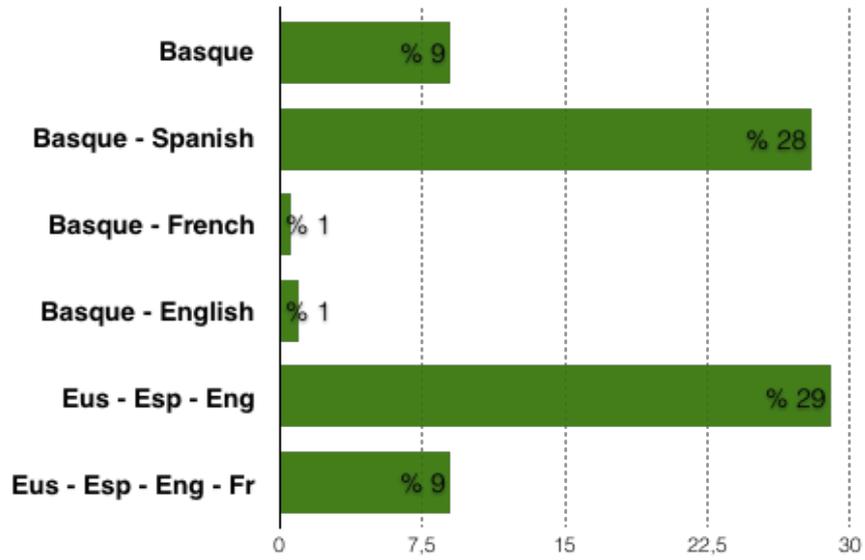


Figure 5: Main language combinations in the websites using the .EUS domain

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.

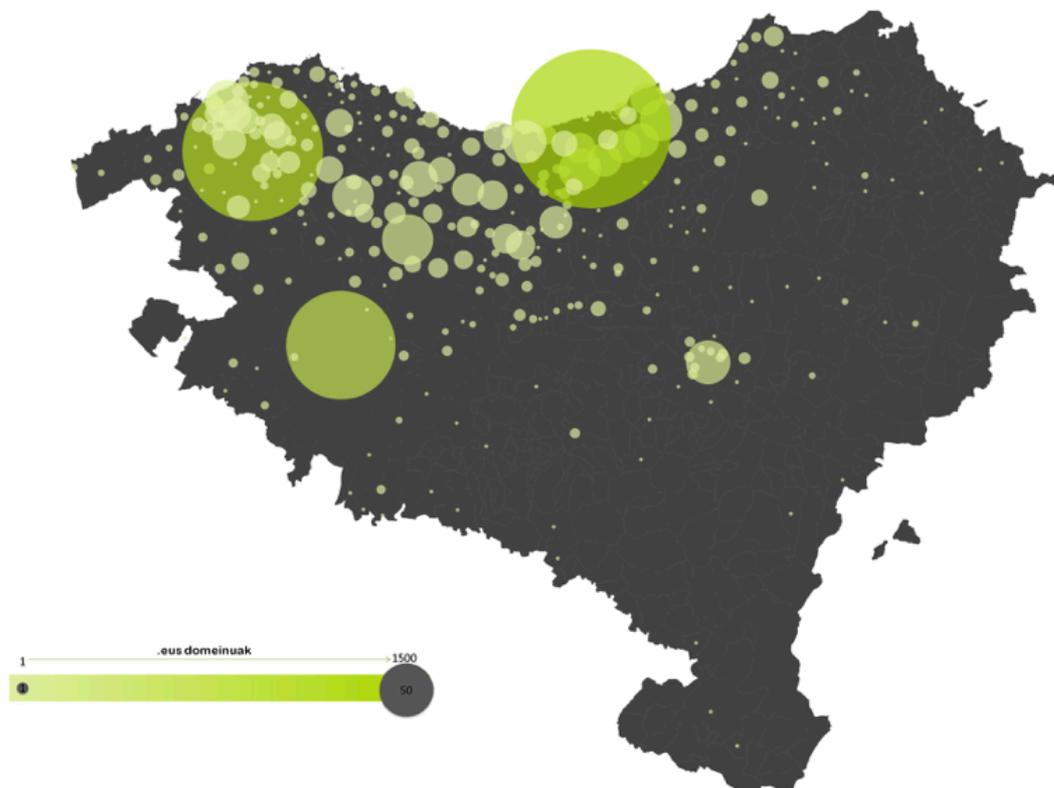


Figure 6: Density of the .EUS domain names in the Basque Country

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.

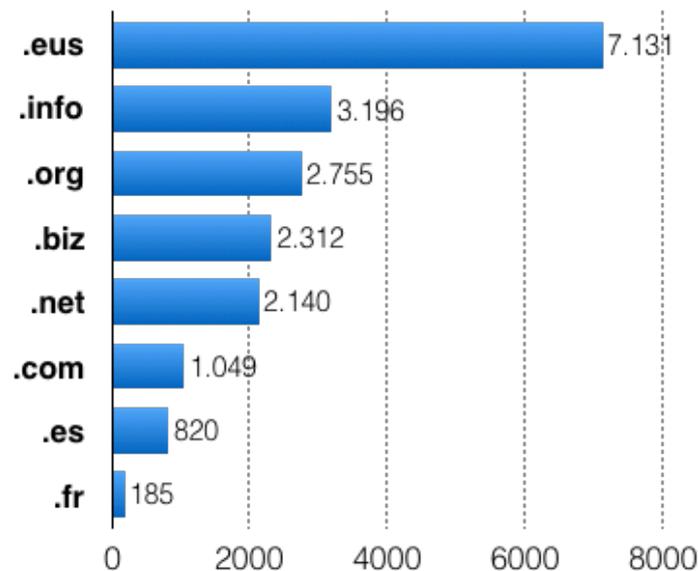


Figure 7: Domain density – Average number of website documents for each TLD

- **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. At the other

extreme would be the .biz TLD; most of the .biz domains registered are reckoned to be defensively registered ones. Among the main TLDs in the Basque Country, after .com, .EUS is the domain that has the most active domains, 58%. That means that the use of the .EUS domain is high among the registrars.

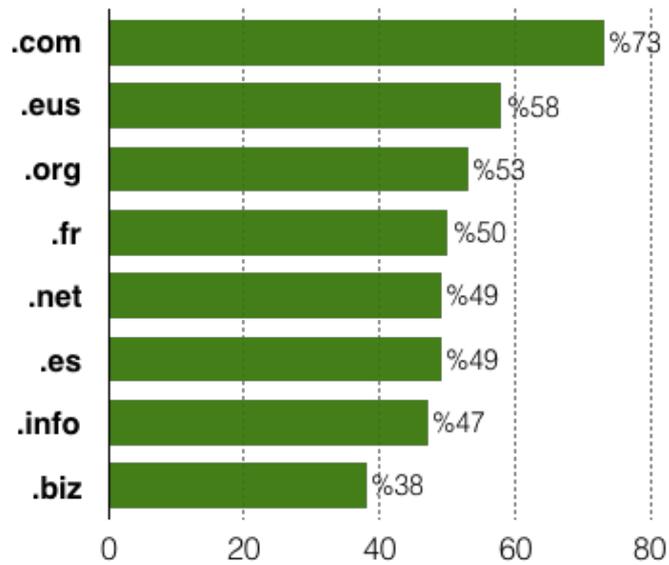


Figure 8: The number of active domains in each TLD

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.

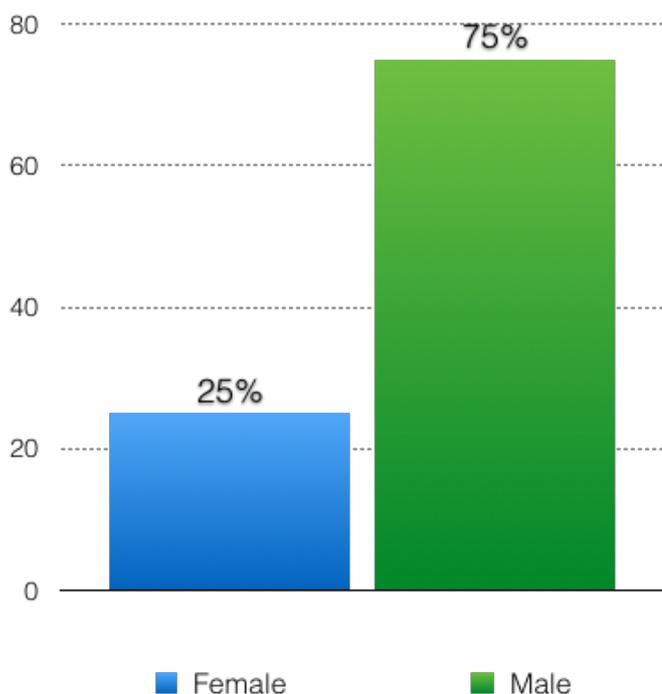


Figure 9: Female and male registrations in the .EUS domain

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 75% 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.

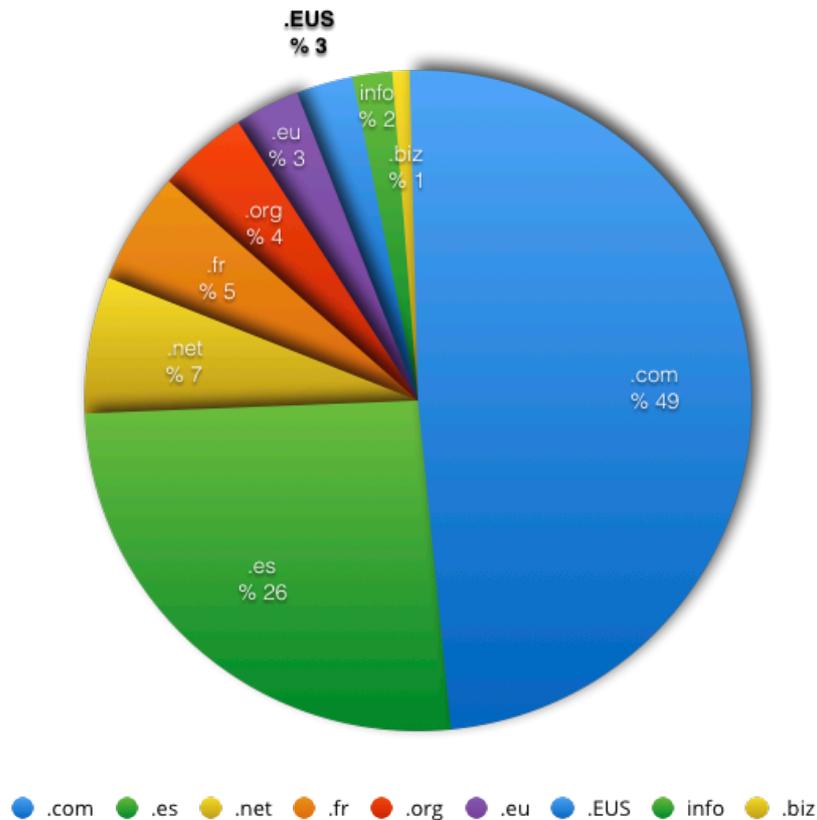


Figure 10: Distribution of main TLDs in the Basque Country

	.COM	.ES	.NET	.FR	.ORG	.EUS	.EU	.INFO	.BIZ
Market %	49%	26%	7%	5%	4%	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 230,000 domain names were registered in 2016, and 136,351 of them had published some content, 59% of them, in fact.

It is clear that .com is the most used among the main TLDs present in the Basque Country both in terms of number of registrations and also in terms of domains that are active. At the other extreme would be the .biz TLD, and most of the .biz domains registered are reckoned to be defensively registered ones. Among the main TLDs in the Basque Country, .EUS is, after .com, the domain that has the highest number of active domains, 58%. That means that the use of the .EUS domain is high among the registrars.

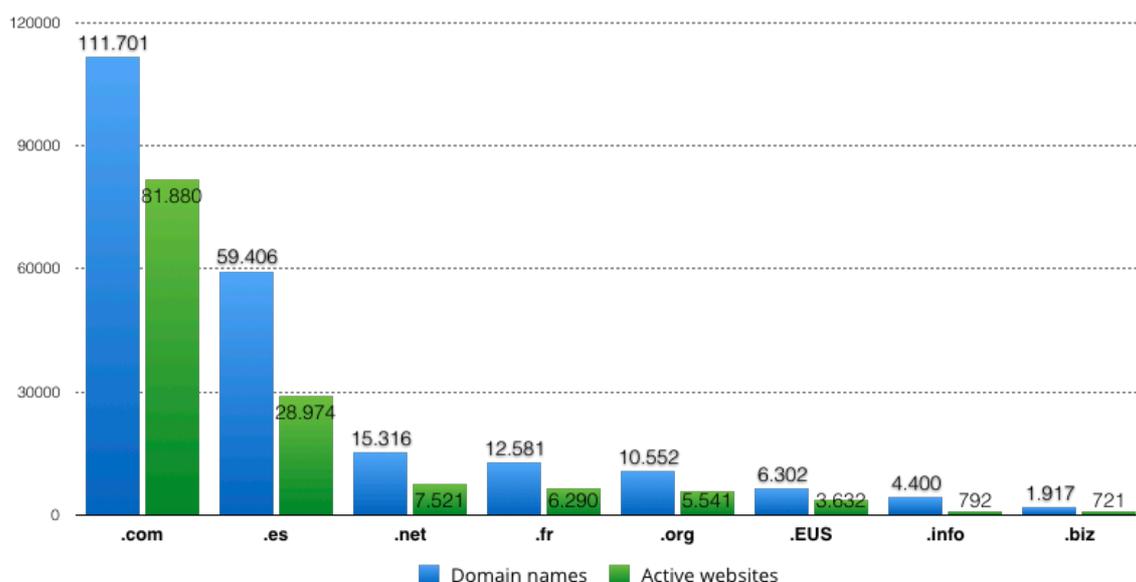


Figure 11: Number of active domains

	.COM	.ES	.NET	.FR	.ORG	.EUS	.INFO	.BIZ
Number of Domains	111.071	59.406	15.316	12.581	10.552	6.302	4.400	1.917
Number of active websites	81.880	28.974	7.521	6.290	5.541	3.632	792	721
% Active ones	%73	%49	%49	%50	%53	%58	%41	%38

3.4. Internet Domain Density in the Basque Country

The figure below shows the density distribution of the main Internet TLDs² 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.

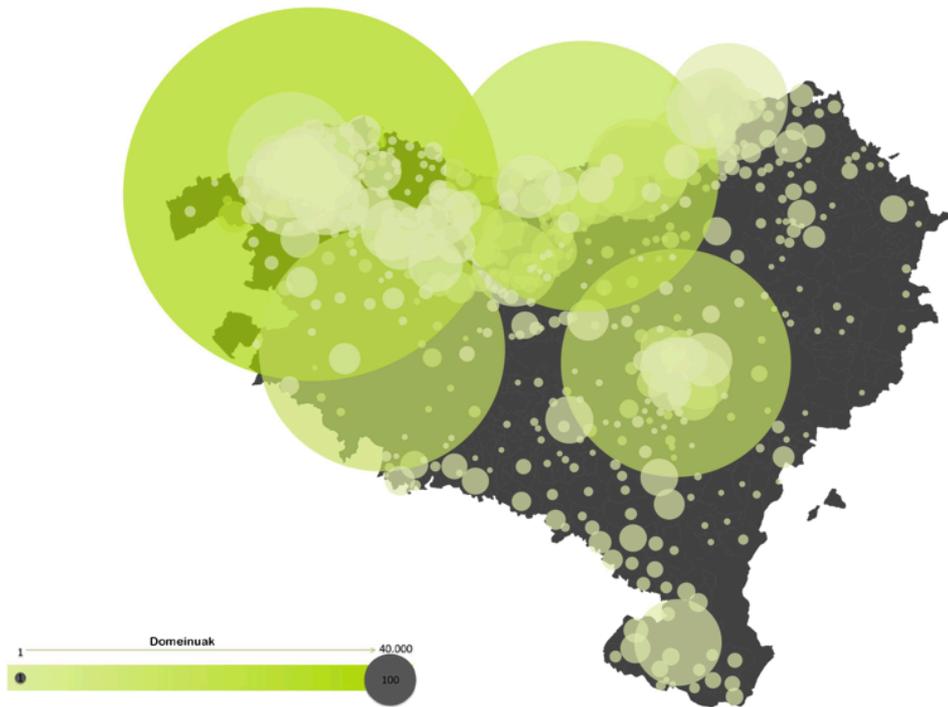


Figure 13: Internet Domain Density in the Basque Country

² Data on the .EU domain are not included in this analysis.

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

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 TLDs³ 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 34% of the websites in Basque but in the remaining cases the presence of Basque is very low.

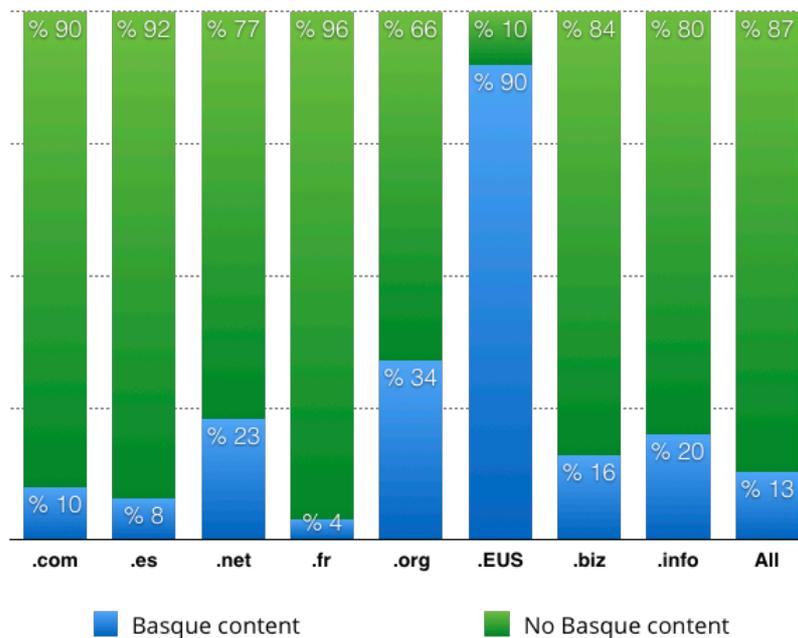


Figure 14: The Presence of Basque in the main TLDs

	.COM	.ES	.NET	.FR	.ORG	.EUS	.BIZ	.INFO	ALL
% of Websites	10%	8%	23%	4%	34%	90%	16%	20%	13%

³ Data on the .EU domain are not included in this analysis.

4.2. Density of Websites with Basque in the Basque Country

The figure below shows the distribution of the density of websites⁴ 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 15: Density of Websites with Basque in the Basque Country

⁴ 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 14,069 active users using Basque in 2016. As far as the tweets are concerned, these active users in the community sent 6,003,534 tweets (2,457,045 in Basque). The number of Twitter users writing mainly in Basque (more than 80% of the content in Basque) has been 2,073, 15% of the total.

Almost half of the Twitter accounts correspond to individuals, of which 66.5% correspond to men and 33.5% to women.

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

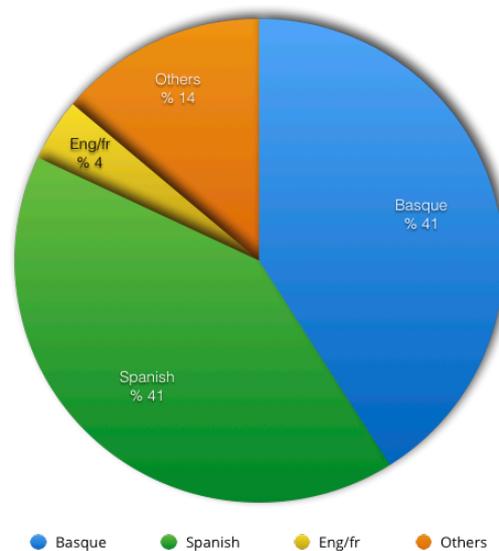


Figure 16: Language use among Basque-speaking Twitter users

	Basque	Spanish	English/French	Others*
Language %	41%	41%	4,2%	13,8%

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

5. 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.

5.1. Permission to use the Data

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.

5.2. PuntuEUS Fundazioa

The .EUS domain is a 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.

The .EUS domain project took its first steps in 2007 and the PuntuEus Association was set up in 2008 for the purpose of setting up the Community of the Basque Language and Culture.

An expansion process took place between 2008 and 2012 and a total of 58 associations and institutions now make up the PuntuEus Association. In 2012, the PuntuEus Association set up the PuntuEus Fundazioa (Foundation), which submitted the request for the .EUS domain to the ICANN in May 2012. This request was considered and on 14 June 2013 we were informed that the .EUS domain had been approved. After completing the necessary work to get it going, the .EUS domain has been visible on the Internet since the spring of 2014, although it was not until 3 December 2014 that it was opened up to all users.



gure domeinua



PuntuEUS Fundazioa 2017

behatokia.domeinuak.eus