



Observatorio de la diversidad lingüística y cultural en la Internet  
Observatoire de la diversité linguistique et culturelle dans l'Internet  
Observatory of the Linguistic and Cultural Diversity in the Internet

<https://OBDILCI.ORG/>

## LESSONS LEARNT FROM MECILDI EXPERIENCE SERIES

### TECHNICAL REPORT #3

#### Web Multilingualism : analysis and typology proposal illustrated by 40 examples

OBDILCI, September 2026

After having applied the MECILDI Multilingual-Aware Computational Framework for Measuring the Language Distribution of Web Content in 30 different web series (gTLD, ccTLD or Tranco), times has come to report on the lessons learnt from that experience. The following technical report is the third of this new series of technical reports elaborated by the Observatory of Linguistic and Cultural Diversity on the Internet (OBDILCI).

The computational framework developed by OBDILCI, with the support of France Ministry of Culture (DGLFLF) and Organisation de la francophonie (OIF), measures language distribution and multilingualism across any collection of websites. It has been applied first to the Tranco series of the million most visited web sites, to 8 gTLDs associated with languages of France, to 10 ccTLDs of member countries of OIF and to some 11 additional ccTLDs, selected as places with local languages counting important speakers' communities. Those measurements have allowed to confirm some previous figures about language distribution in the Web and to compute new indicators of multilingualism.

The first study (OBDILCI, 2026-1) focused the results for **indigenous, vernacular, local or minority** languages and offers a panorama particularly worrying of the tiny prevalence of those languages in the Web, with data much lower than the most pessimistic forecast. In the following, we will use term "local" for any of the mentioned categories of language.

The second study (OBDILCI, 2026-2) focused the global picture of indicators for language distribution and multilingualism indicators and the adjustments required to the OBDILCI model to reflect the lessons learned from the use of MECILDI in some 45 different series of websites (Tranco, ccTLD or gTLDs).

This third study offers the analysis of web multilingualism and from the characterization of the different situations encountered try to provide a first proposal of typology.

## Table of contents

|                                                         |    |
|---------------------------------------------------------|----|
| 1- WEB MULTILINGUALISM INDICATORS .....                 | 4  |
| 2- SUMMARY OF RESULTS .....                             | 6  |
| 3-TYPOLOGY .....                                        | 11 |
| 4- Web Multilingualism topology by examples.....        | 12 |
| 4.1 Luxembourg (.lu) : extreme, trilingual, local.....  | 12 |
| 4.2 Euskadi (.eus): extreme, bilingual, local.....      | 13 |
| 4.3 Monaco (.mc): extreme, bilingual.....               | 15 |
| 4.4 Guadeloupe (.gp): High, quadrilingual, global ..... | 16 |
| 4.5 Andorra (.ad): High, quadrilingual, local.....      | 17 |
| 4.6 Estonia (.ad): High, trilingual, local.....         | 19 |
| 4.7 Paris (.paris): High, bilingual.....                | 20 |
| 4.8 Curacao (.cw): High, bilingual, local.....          | 22 |
| 4.9 TRANCO: Fair/High boundary, global .....            | 22 |
| 4.10 Haiti (.ht): Fair, bilingual, local, global .....  | 24 |
| 4.11 Madagascar (.mg): Fair, bilingual, local .....     | 25 |
| 4.12 Burundi (.bi): Fair, bilingual, global.....        | 26 |
| 4.13 Lebanon (.lb): Fair, bilingual .....               | 27 |
| 4.14 Israel (.il): Fair .....                           | 28 |
| 4.15 Bretagne (.bzh): Medium, local .....               | 29 |
| 4.16 St Lucia (.lc): Medium, global.....                | 30 |
| 4.17 Vietnam (.vn): Medium .....                        | 32 |
| 4.18 Asia (.asia): Low, bilingual.....                  | 32 |
| 4.19 Africa (.africa): Low, global .....                | 33 |
| 4.20 India (.in): Very Low, local .....                 | 35 |

## List of tables

|                                                      |    |
|------------------------------------------------------|----|
| Table 1:Figures per language.....                    | 4  |
| Table 2: Figures per country (ccTLD).....            | 5  |
| Table 3: Figures per gTLD .....                      | 5  |
| Table 4 : Results for all measured TLD- part 1 ..... | 6  |
| Table 5: Results for all measured TLD- part 2 .....  | 8  |
| Table 6: Results for all measured TLD- part 3 .....  | 9  |
| Table 7: Extreme results by indicator 1/2 .....      | 10 |
| Table 8: Extreme results by indicator 2/2 .....      | 10 |

|                                                |    |
|------------------------------------------------|----|
| Table 9: Indicators for Luxembourg (.lu) ..... | 12 |
| Table 10: Indicators for Euskadi (.eus).....   | 14 |
| Table 11: Indicators for Monaco (.mc).....     | 15 |
| Table 12: Indicators for Guadeloupe (.gp)..... | 16 |
| Table 13: Indicators for Andorra (.mc).....    | 18 |
| Table 14: Indicators for Estonia (ee) .....    | 19 |
| Table 15: Indicators for Paris (.paris) .....  | 21 |
| Table 16: Indicators for Curacao (.cw).....    | 22 |
| Table 17: Indicators for Tranco .....          | 23 |
| Table 18: Indicators for Haiti (.ht).....      | 24 |
| Table 19: Indicators for Madagascar (.mg)..... | 25 |
| Table 20: Indicators for Burundi (.bi).....    | 26 |
| Table 21: Indicators for Lebanon (.lb).....    | 27 |
| Table 22: Indicators for Israel (.il).....     | 29 |
| Table 23: Indicators for Bretagne (.bzh).....  | 29 |
| Table 24: Indicators for St Lucia (.lc).....   | 31 |
| Table 25: Indicators for Viet Nam (.vn) .....  | 32 |
| Table 26: Indicators for .Asia .....           | 32 |
| Table 27: Asian languages in .asia.....        | 33 |
| Table 28: Indicators for .africa.....          | 33 |
| Table 29: Indicators for India (.in).....      | 35 |
| Table 30: Indian languages in .in.....         | 36 |

## List of figures

|                                                                                  |    |
|----------------------------------------------------------------------------------|----|
| Figure 1:: Repartition of the measured TLD following the proposed typology ..... | 11 |
| Figure 2: Histogram for Luxembourg AvgL 1/2 .....                                | 13 |
| Figure 3: Histogram for Luxembourg AvgL 2/2 .....                                | 13 |
| Figure 4: Histogram for Euskadi AvgL 1/2 .....                                   | 14 |
| Figure 5: Histogram for Euskadi AvgL 2/2 .....                                   | 14 |
| Figure 6: Histogram for Monaco AvgL 1/2 .....                                    | 15 |
| Figure 7: Histogram for Monaco AvgL 1/2 .....                                    | 16 |
| Figure 8: Histogram for Guadeloupe AvgL 1/2.....                                 | 17 |
| Figure 9: Histogram for Guadeloupe AvgL 2/2.....                                 | 17 |
| Figure 10: Histogram for Andorra AvgL ½.....                                     | 18 |
| Figure 11: Histogram for Andorra AvgL 1/2.....                                   | 19 |
| Figure 12: Histogram for Estonia AvgL 1/2 .....                                  | 20 |
| Figure 13: Histogram for Estonia AvgL 2/2 .....                                  | 20 |
| Figure 14: Histogram for Paris AvgL 1/2 .....                                    | 21 |
| Figure 15: Histogram for Paris AvgL 2/2 .....                                    | 21 |
| Figure 16: Histogram for Curacao AvgL .....                                      | 22 |
| Figure 17: Histogram for Tranco AvgL 1/2.....                                    | 23 |
| Figure 18: Histogram for Tranco AvgL 2/2.....                                    | 23 |
| Figure 19: Histogram for Haiti AvgL 1/2 .....                                    | 24 |
| Figure 20: Histogram for Haiti AvgL 2/2 .....                                    | 24 |
| Figure 21: Histogram for Madagascar AvgL 1/2.....                                | 25 |
| Figure 22: Histogram for Madagascar AvgL 2/2.....                                | 26 |

|                                                |    |
|------------------------------------------------|----|
| Figure 23:Histogram for Burundi AvgL .....     | 27 |
| Figure 24:Histogram for Lebanon AvgL 1/2 ..... | 28 |
| Figure 25:Histogram for Lebanon AvgL 2/2 ..... | 28 |
| Figure 26:Histogram for Bretagne AvgL 1/2..... | 30 |
| Figure 27:Histogram for Bretagne AvgL 2/2..... | 30 |
| Figure 28:Histogram for St Lucia AvgL 1/2..... | 31 |
| Figure 29:Histogram for St Lucia AvgL 2/2..... | 31 |
| Figure 30:Histogram for .africa AvgL 1/2 ..... | 34 |
| Figure 31:Histogram for .africa AvgL 2/2 ..... | 34 |
| Figure 32: African languages in .africa.....   | 34 |

## 1-WEB MULTILINGUALISM INDICATORS

We have chosen 3 indicators to characterize the multilingual character of any series of websites:

**%MULTI**: The percentage of multilingual websites within the series

**MultiR**: The rate of multilingualism of the series, defined as the ratio between the total number of linguistic configurations divided by the total number of websites of the series.

**AvgL**: The average number of languages per multilingual websites.

Note that those are not three independent indicators, since they are linked by the formula:

$$\mathbf{MultiR} = (1 - \%Multi) + AvgL \times \%Multi$$

Note also that 
$$\mathbf{MultiR} = \sum_{L=1}^{L=n} \%Sites(L)$$

Where %SITES (L) is the second main indicator of language distribution of a series defined as the percentage of websites having a linguistic configuration in language L.

The MECILDI experience confirmed that the most suitable indicator to work is %MULTI. The report established from the use of DataProvider.com database of a large proportion of existing websites allowed in (OBDILCI, 2025) to obtain the following results, to sustain the high variability of this indicator:

*Table 1:Figures per language*

| INDICATOR | LANGUAGE   | VALUE |
|-----------|------------|-------|
| Highest   | Basque     | 63.5% |
| Lowest    | Chinese    | 2.1%  |
| Median    | Uzbek      | 15.5% |
| S.Average | Vietnamese | 19.0% |
| W.Average | Russian    | 11.0% |

Table 2: Figures per country (ccTLD)

| INDICATOR | COUNTRY  | VALUE |
|-----------|----------|-------|
| Highest   | Monaco   | 72.0% |
| Lowest    | Zimbabwe | 2.2%  |
| Median    | Hungary  | 15.6% |
| S.Average | Lybia    | 20.4% |
| W.Average | Czechia  | 14.4% |

Table 3: Figures per gTLD

| INDICATOR | gTLD  | VALUE |
|-----------|-------|-------|
| Highest   | .eus  | 100%  |
| Lowest    | .mobi | 1.2%  |
| Median    | .biz  | 8.9%  |
| S.Average | .lat  | 14.1% |
| W.Average | .lol  | 10.9% |

Notes :

- The country data is made on an exhaustive list of countries
- The language data is made on the list of the 88 languages identified by DataProvider.com
- The gTLDs data is far from being exhaustive and is made on 60 different gTLDs.
- The S.Average is the mean of all the values.
- The W.Average is a mean weighted with the number of websites corresponding with each element.
- The W.Averages differ depending of the type of measure because the coverage is different depending of the field in the data base. The global W.Average of the whole database of the websites having the language field informed is 12.4%. The value for .com which represents more than 50% of the total existing domains was 11.1%.
- The values are computed from the percentage of occurrence of the Hreflang field extrapolated on the assumption that 40% of multilingual websites use this technic. The values produced by MECILDI use the same extrapolative assumption, however the definition of the attribution to a language to a website may differ, MECILDI reflecting the multilingualism of websites and DataProvider.com not.

As for the two other indicators, the experience shows that with the same level of Multi%, we could find ccTLDs with AvgL and MultiR in opposite position; as an example, .quebec can have a very high %Multi of 50% while having the lowest AvgL (2.4) of the studied group, this affecting its MultiR, to be relatively low and is the symptom of a domain highly bilingual. In another example, .lc (St Lucia) have the record AvgL of 12 with a moderate %Multi of 18%. This makes the typology not so easy and the goal to reach a simple one out of reach.

MultiR has been chosen as a key indicator because it is interesting to compare it with the equivalent value for humanity, which is 1.44, and is defined as the ratio between L1+L2 speakers by human's population (L1).

The analysis of several cases show that high AvgL are often the results of redirection towards websites with a worldwide vocation, either linked to ecommerce, either some type of global

Internet service. This observation was so recurrent in some TLD, with multiple different websites directing to the same one, that it was decided to release a new version 1.9 to prevent the same site to appear more than once in redirection; note that this report is based on figures obtained before this new release. This could downsize the AvgL in notable proportions and was considered a fair decision to avoid inflating artificially the multilingualism indicators; however, it is not yet implemented.

Paying attention to the distribution of AvgL is required as a mean to better understand the characteristics of multilingual websites, distinguishing the local merits (highly bilingual, trilingual and even in few cases quadrilingual sites) from an apparent multilingualism boosted by global websites with large number of languages. To illustrate the different type of websites the report will therefore present in most of the cases the histograms presenting the distribution of AvgL.

## 2-SUMMARY OF RESULTS

### Legends:

- TLD : Top Level Domain
- cc/g : cc for ccTLD, g for gTLD, mix for other type of series
- Country/Generic : The country associated with the ccTLD or the generic domain.
- %Multi, MultiR, AvgL : the main indicators already defined
- Goot%: the percentage of sites sing Google Translate facilities imbedded
- IDN% the percentage of sites using Internationalized Domains Names.
- 1<sup>st</sup>, 2<sup>nd</sup>, 3<sup>rd</sup>, 4<sup>th</sup> and 5<sup>th</sup> languages in the ranking in terms of %Sites and corresponding percentage.
- MAX : the higher value for the indicator
- MIN: the smaller value for the indicator
- AVG : the simple mean for all the indicators
- W.AVG:T: The mean for all the indicators weighted by total number of domains below
- W.AVG:M: The mean for all the indicators weighted by total number of measured domains below.
- The highest values are set in red the top one in bold.
- The lowest values are set in blue the bottom one in bold.
- The local languages appearing are set in red.

The results which are gathered in a single Excel file for process are copied here below and split in 4 different tables for space reasons.

*Table 4 : Results for all measured TLD- part 1*

|            |             |                             |               |               |             |              |              |                |                   |
|------------|-------------|-----------------------------|---------------|---------------|-------------|--------------|--------------|----------------|-------------------|
| TRANCO     | mix         | <b>Million most visited</b> | 37.27%        | <b>3.41</b>   | <b>7.47</b> | 1.06%        | <b>0.16%</b> | <b>English</b> | <b>67.52%</b>     |
| <b>TLD</b> | <b>cc/g</b> | <b>Country/Generic</b>      | <b>%Multi</b> | <b>MultiR</b> | <b>AvgL</b> | <b>Goot%</b> | <b>IDN%</b>  | <b>1st L</b>   | <b>%Sites 1st</b> |
| ad         | cc          | <b>Andorra</b>              | <b>56.93%</b> | <b>2.97</b>   | 4.46        | 0.91%        | <b>0.00%</b> | <b>Catalan</b> | 61.68%            |
| africa     | g           | <b>Africa</b>               | <b>10.70%</b> | 1.50          | 5.66        | 0.57%        | <b>0.00%</b> | <b>English</b> | 89.50%            |
| alsace     | g           | <b>Alsace</b>               | 30.57%        | 1.77          | 3.50        | 0.82%        | <b>0.00%</b> | <b>French</b>  | 86.14%            |

| ao      | cc   | Angola          | 22.12% | 1.79   | 4.55  | 5.88% | 0.00% | Portuguese | 91.01%        |
|---------|------|-----------------|--------|--------|-------|-------|-------|------------|---------------|
| asia    | g    | Asia            | 11.97% | 1.53   | 5.58  | 0.00% | 0.10% | Chinese    | 54.74%        |
| bi      | cc   | Burundi         | 21.62% | 1.88   | 5.07  | 1.50% | 0.00% | English    | 53.10%        |
| bo      | cc   | Bolivia         | 14.36% | 2.01   | 8.04  | 0.65% | 0.02% | Spanish    | 91.09%        |
| bzh     | g    | Bretagne        | 16.42% | 1.44   | 3.68  | 1.42% | 0.00% | French     | 82.65%        |
| cd      | cc   | Rep. Dem.Congo  | 18.42% | 1.94   | 6.13  | 2.11% | 0.11% | French     | 78.98%        |
| ci      | cc   | Côte d'Ivoire   | 19.14% | 1.92   | 5.82  | 1.39% | 0.34% | French     | 83.33%        |
| corsica | g    | Corse           | 29.91% | 1.85   | 3.84  | 1.97% | 0.00% | French     | 74.87%        |
| cv      | cc   | Cabo Verde      | 23.98% | 1.79   | 4.30  | 1.97% | 0.00% | English    | 59.18%        |
| cw      | cc   | Curacao         | 38.18% | 1.60   | 2.58  | 0.49% | 0.00% | English    | 77.41%        |
| do      | cc   | Dominican Rep   | 27.65% | 2.26   | 5.55  | 1.52% | 0.00% | Spanish    | 78.66%        |
| ee      | cc   | Estonia         | 55.60% | 2.44   | 3.59  | 0.23% | 1.28% | Estonian   | 88.61%        |
| eus     | g    | Euskadi         | 97.22% | 2.59   | 2.63  | 0.51% | 0.15% | Spanish    | 94.39%        |
| gp      | g    | Guadeloupe      | 60.96% | 6.05   | 9.28  | 0.00% | 0.14% | English    | 58.60%        |
| ht      | cc   | Haiti           | 23.63% | 2.31   | 6.55  | 1.22% | 0.00% | French     | 56.84%        |
| il      | cc   | Israel          | 21.96% | 1.42   | 2.89  | 0.36% | 0.00% | Hebrew     | 92.31%        |
| in      | cc   | India           | 6.52%  | 1.28   | 5.39  | 3.20% | 0.00% | English    | 91.83%        |
| kw      | cc   | Kuwait          | 75.07% | 2.68   | 3.24  | 0.53% | 0.00% | English    | 94.39%        |
| lat     | g    | Latin America   | 15.27% | 1.60   | 4.89  | 0.16% | 0.06% | English    | 45.39%        |
| lb      | cc   | Lebanon         | 24.25% | 1.62   | 3.55  | 0.62% | 0.00% | English    | 80.91%        |
| lc      | cc   | St Lucia        | 17.96% | 3.05   | 12.42 | 0.28% | 0.00% | English    | 70.66%        |
| lu      | cc   | Luxembourg      | 70.36% | 3.02   | 3.87  | 0.40% | 0.05% | French     | 81.09%        |
| ma      | cc   | Maroc           | 31.79% | 1.89   | 3.81  | 0.67% | 0.44% | French     | 81.25%        |
| mc      | cc   | Monaco          | 75.43% | 2.57   | 3.08  | 0.57% | 0.00% | French     | 89.77%        |
| md      | cc   | Moldova         | 74.34% | 2.47   | 2.98  | 0.67% | 0.01% | Romanian   | 84.93%        |
| mg      | cc   | Madagascar      | 25.42% | 1.65   | 3.54  | 0.27% | 0.00% | French     | 77.17%        |
| mq      | cc   | Martinique      | 31.72% | 2.68   | 6.29  | 0.75% | 0.00% | French     | 87.05%        |
| mr      | cc   | Mauritania      | 57.60% | 2.78   | 4.10  | 0.49% | 0.00% | French     | 75.41%        |
| nc      | g    | N. Calédonie    | 20.72% | 1.42   | 3.06  | 0.34% | 0.00% | French     | 93.84%        |
| paris   | g    | Paris           | 46.91% | 2.48   | 4.14  | 0.37% | 0.00% | French     | 63.88%        |
| pe      | cc   | Peru            | 13.63% | 1.54   | 4.97  | 1.78% | 0.04% | Spanish    | 89.25%        |
| pf      | cc   | Fr. Polynesia   | 25.72% | 1.58   | 3.25  | 1.01% | 0.00% | French     | 94.00%        |
| py      | cc   | Paraguay        | 15.43% | 1.90   | 6.85  | 1.24% | 0.00% | Spanish    | 91.96%        |
| qa      | cc   | Qatar           | 49.70% | 2.50   | 4.02  | 1.12% | 0.00% | English    | 91.39%        |
| quebec  | g    | Québec          | 49.98% | 1.72   | 2.45  | 0.35% | 7.88% | French     | 67.41%        |
| ro      | cc   | Romania         | 19.53% | 1.58   | 3.97  | 1.52% | 0.00% | Romanian   | 83.45%        |
| vn      | cc   | Vietnam         | 18.23% | 1.38   | 3.06  | 2.58% | 0.00% | Vietnamese | 92.38%        |
| vu      | cc   | Vanuatu         | 17.57% | 1.61   | 4.45  | 1.28% | 0.00% | English    | 92.35%        |
| TLD     | cc/g | Country/Generic | %Multi | MultiR | AvgL  | Goot% | IDN%  | 1st L      | %Sites<br>1st |
| MAX     |      |                 | 97.22% | 6.05   | 12.42 | 5.88% | 7.88% |            | 94.39%        |
| MIN     |      |                 | 6.52%  | 1.28   | 2.45  | 0.00% | 0.00% |            | 45.39%        |

|         |  |  |        |      |      |       |       |  |        |
|---------|--|--|--------|------|------|-------|-------|--|--------|
| AVG     |  |  | 33.92% | 2.10 | 4.65 | 1.09% | 0.27% |  | 80.17% |
| W.AVG.T |  |  | 18.86% | 1.55 | 4.42 | 1.87% | 0.11% |  | 83.37% |
| W.AVG.M |  |  | 24.30% | 1.68 | 4.14 | 1.37% | 0.16% |  | 81.91% |

Table 5: Results for all measured TLD- part 2

|         |            |               |            |               |            |               |            |               |
|---------|------------|---------------|------------|---------------|------------|---------------|------------|---------------|
| TRANCO  | German     | 21.98%        | French     | 20.73%        | Spanish    | 20.54%        | Italian    | 13.91%        |
| TLD     | 2nd L      | %Sites<br>2nd | 3rd L      | %Sites<br>3rd | 4th L      | %Sites<br>4th | 5th L      | %Sites<br>5th |
| ad      | English    | 58.33%        | Spanish    | 54.53%        | French     | 42.13%        | German     | 8.23%         |
| africa  | French     | 13.21%        | German     | 5.03%         | Spanish    | 3.05%         | Portuguese | 2.73%         |
| alsace  | English    | 22.01%        | German     | 10.33%        | Spanish    | 2.54%         | Dutch      | 2.08%         |
| ao      | English    | 28.08%        | French     | 6.19%         | Spanish    | 3.81%         | German     | 2.33%         |
| asia    | English    | 32.51%        | German     | 6.82%         | French     | 5.09%         | Vietnamese | 5.01%         |
| bi      | French     | 45.10%        | German     | 11.19%        | Spanish    | 6.28%         | Russian    | 6.02%         |
| bo      | English    | 19.25%        | Portuguese | 7.10%         | French     | 4.83%         | German     | 4.34%         |
| bzh     | English    | 19.80%        | German     | 2.67%         | Breton     | 2.19%         | Spanish    | 1.21%         |
| cd      | English    | 30.72%        | German     | 5.47%         | Spanish    | 5.28%         | Portuguese | 4.93%         |
| ci      | English    | 31.48%        | Spanish    | 4.56%         | Portuguese | 3.90%         | Chinese    | 3.61%         |
| corsica | English    | 32.05%        | Italian    | 4.79%         | German     | 4.10%         | Corsican   | 2.08%         |
| cv      | Portuguese | 44.83%        | French     | 10.45%        | Indonesian | 7.93%         | Vietnamese | 6.97%         |
| cw      | Dutch      | 51.02%        | Papiamentu | 17.24%        | Spanish    | 7.24%         | OTHERL     | 1.23%         |
| do      | English    | 40.21%        | French     | 8.04%         | German     | 7.21%         | Portuguese | 6.37%         |
| ee      | English    | 52.27%        | Russian    | 28.36%        | Finnish    | 10.64%        | Latvian    | 5.90%         |
| eus     | Basque     | 94.02%        | English    | 27.69%        | French     | 17.42%        | Catalan    | 2.58%         |
| gp      | French     | 58.42%        | German     | 18.77%        | Spanish    | 15.79%        | Italian    | 12.46%        |
| ht      | English    | 52.72%        | Spanish    | 9.66%         | German     | 9.16%         | Creole     | 7.89%         |
| il      | English    | 25.35%        | Russian    | 4.09%         | Arabic     | 2.40%         | French     | 2.12%         |
| in      | Hindi      | 4.83%         | Vietnamese | 2.61%         | Indonesian | 1.96%         | Spanish    | 1.77%         |
| kw      | Arabic     | 77.09%        | French     | 7.37%         | German     | 5.62%         | Italian    | 5.29%         |
| lat     | Indonesian | 21.69%        | Spanish    | 21.28%        | Portuguese | 10.90%        | Chinese    | 8.40%         |
| lb      | Arabic     | 34.40%        | French     | 12.05%        | German     | 2.18%         | Italian    | 1.87%         |
| lc      | Chinese    | 21.74%        | Spanish    | 19.36%        | German     | 18.26%        | French     | 17.79%        |
| lu      | English    | 64.51%        | German     | 51.51%        | Luxemburg. | 12.83%        | Dutch      | 11.09%        |
| ma      | English    | 34.20%        | Arabic     | 22.62%        | Spanish    | 6.33%         | German     | 3.74%         |
| mc      | English    | 82.65%        | Italian    | 26.49%        | German     | 8.30%         | Russian    | 8.06%         |
| md      | Russian    | 72.54%        | English    | 40.15%        | Italian    | 3.02%         | German     | 2.93%         |
| mg      | English    | 38.48%        | Malagasy   | 7.58%         | German     | 3.75%         | Chinese    | 2.36%         |
| mq      | English    | 40.36%        | Spanish    | 13.06%        | German     | 7.46%         | Italian    | 7.46%         |
| mr      | Arabic     | 61.21%        | English    | 36.17%        | Spanish    | 12.66%        | Portuguese | 9.80%         |
| nc      | English    | 12.89%        | Japanese   | 1.26%         | Spanish    | 0.58%         | Portuguese | 0.39%         |
| paris   | English    | 48.94%        | German     | 7.73%         | Spanish    | 4.83%         | Italian    | 3.53%         |
| pe      | English    | 19.68%        | Portuguese | 3.53%         | French     | 2.75%         | German     | 2.46%         |

|         |         |               |            |               |            |               |            |               |
|---------|---------|---------------|------------|---------------|------------|---------------|------------|---------------|
| pf      | English | 28.21%        | Spanish    | 4.64%         | Japanese   | 2.95%         | Portuguese | 2.67%         |
| py      | English | 17.93%        | Portuguese | 8.77%         | German     | 4.26%         | French     | 2.96%         |
| qa      | Arabic  | 51.78%        | French     | 7.42%         | German     | 5.80%         | Spanish    | 5.59%         |
| quebec  | English | 50.84%        | Spanish    | 1.09%         | German     | 1.09%         | Italian    | 0.51%         |
| ro      | English | 28.61%        | Hungarian  | 4.75%         | German     | 3.91%         | French     | 2.83%         |
| vn      | English | 23.89%        | Japanese   | 2.00%         | Chinese    | 1.90%         | Korean     | 0.98%         |
| vu      | French  | 11.80%        | Spanish    | 5.08%         | Portuguese | 4.79%         | German     | 4.28%         |
| TLD     | 2nd L   | %Sites<br>2nd | 3rd L      | %Sites<br>3rd | 4th L      | %Sites<br>4th | 5th L      | %Sites<br>5th |
| MAX     |         | 94.02%        |            | 54.53%        |            | 42.13%        |            | 17.79%        |
| MIN     |         | 4.83%         |            | 1.09%         |            | 0.58%         |            | 0.39%         |
| AVG     |         | 38.85%        |            | 13.72%        |            | 7.08%         |            | 4.72%         |
| W.AVG.T |         | 21.66%        |            | 8.52%         |            | 4.43%         |            | 3.20%         |
| W.AVG.M |         | 27.68%        |            | 11.47%        |            | 5.30%         |            | 3.74%         |

Table 6: Results for all measured TLD- part 3

| TRANCO  | 1 000 000      | 100 000       |                        | F   |    | 1  |   |                             |
|---------|----------------|---------------|------------------------|-----|----|----|---|-----------------------------|
| TLD     | #SITES<br>TOT. | #SITES<br>M.. | COMMENTS               | MLL | ML | GS | L | TYOLOGY                     |
| ad      | 2 429          | 1 212         |                        | H   | 4  |    | 1 | High + quadrilingual +local |
| africa  | 59 628         | 18 181        |                        | L   |    | 1  |   | Low + global                |
| alsace  | 2 877          | 1 104         |                        | F   |    |    |   | Fair                        |
| ao      | 3 866          | 1 752         |                        | F   | 2  |    |   | Medium + bilingual          |
| asia    | 894 439        | 100 000       |                        | L   | 2  | 1  |   | Low + bilingual             |
| bi      | 1 280          | 532           |                        | F   | 2  | 1  |   | Medium + bilingual          |
| bo      | 5 620          | 2 751         |                        | M   |    | 1  |   | Low + bilingual + global    |
| bzh     | 17 397         | 7 032         |                        | M   |    |    | 1 | Medium + local              |
| cd      | 1 797          | 760           |                        | M   |    | 1  |   | Moderate + global           |
| ci      | 4 699          | 2 012         |                        | M   |    | 1  |   | Fair + global               |
| corsica | 3 058          | 1 170         |                        | F   | 2  |    | 1 | Fair + Local                |
| cv      | 6 827          | 1 220         | Hacked 50%             | F   | 2  |    |   | Medium + bilingual          |
| cw      | 350            | 203           | OtherL =<br>papamiento | F   | 3  |    | 1 | Fair + trilingual local     |
| do      | 12 349         | 5 778         |                        | F   | 2  | 1  |   | Fair + bilingual            |
| ee      | 102 596        | 58 348        |                        | H   | 3  |    |   | High + trilingual           |
| eus     | 31 476         | 8 244         |                        | E   | 2  |    |   | Fair + bilingual            |
| gp      | 2 144          | 570           |                        | H   | 2  | 1  |   | High + quadrilingual.       |
| ht      | 2 119          | 328           |                        | F   |    | 1  | 1 | Medium + local              |
| il      | 127 495        | 71 925        |                        | F   |    |    |   | Medium                      |
| in      | 853 000        | 100 000       |                        | L   | 2  | 1  | 1 | Low + Local                 |
| kw      | 1 392          | 756           |                        | E   | 2  |    |   | Extreme + bilingual         |
| lat     | 324 264        | 100 000       |                        | M   |    |    | 1 | Mediu. + global             |
| lb      | 1 371          | 804           |                        | F   | 2  |    | 1 | Fair + local                |

|         |           |         |          |     |     |    |   |                              |
|---------|-----------|---------|----------|-----|-----|----|---|------------------------------|
| lc      | 1 472     | 362     |          | M   |     | 1  |   | Medium + global              |
| lu      | 47 212    | 22 390  |          | E   | 3   |    | 1 | Extreme + trilingual + local |
| ma      | 32 165    | 14 728  |          | F   | 3   |    | 1 | Fair + trilingual + local    |
| mc      | 1 891     | 1 054   |          | E   | 2   |    |   | Extreme + trilingual         |
| md      | 24 452    | 12 755  |          | E   | 3   |    | 1 | Extreme + trilingual         |
| mg      | 2 459     | 1 131   |          | F   |     |    | 1 | Fair + local                 |
| mq      | 277       | 134     |          | F   | 2   | 1  |   | Fair + global                |
| mr      | 821       | 408     |          | H   | 3   |    |   | High + trilingual            |
| nc      | 3 567     | 2 063   |          | F   |     |    |   | Medium                       |
| paris   | 20 505    | 7 285   |          | H   | 2   |    |   | High + bilingual             |
| pe      | 52 649    | 27 953  |          | M   |     |    |   | Low + global                 |
| pf      | 1 124     | 593     |          | F   |     |    |   | Fair                         |
| py      | 9 658     | 5 574   |          | M   |     | 1  |   | Low + global                 |
| qa      | 6 451     | 3 209   |          | H   | 2   |    |   | High + Bilingual             |
| quebec  | 14 572    | 2 571   |          | F   | 2   |    |   | High + Bilingual             |
| ro      | 377 288   | 100 000 |          | M   |     |    | 1 | Medium + Local.              |
| vn      | 282 269   | 100 000 |          | M   |     |    | 1 | Medium                       |
| vu      | 2 034     | 313     |          | M   |     |    |   | Medium                       |
| TLD     | 2 537 468 | 683 996 | COMMENTS | MLL | MLT | GS | L | TYPOLOGY                     |
| MAX     | 894 439   | 100 000 |          |     |     |    |   |                              |
| MIN     | 277       | 134     |          |     |     |    |   |                              |
| AVG     | 65 063    | 17 538  |          |     |     |    |   |                              |
| W.AVG.T | 455 879   | 84 140  |          |     |     |    |   |                              |
| W.AVG.M | 312 141   | 74 331  |          |     |     |    |   |                              |

The 2 following tables present the TLD corresponding to the most extreme values of the main indicators.

Table 7: Extreme results by indicator 1/2

| INDICATOR<br>→ | MultiL | Avgl      | Goot%      | IDN%   |
|----------------|--------|-----------|------------|--------|
| Highest        | .eus   | St. Lucia | Angola     | Quebec |
| Lowest         | India  | Quebec    | Guadeloupe | Many   |

Table 8: Extreme results by indicator 2/2

| INDICATOR<br>→ | 1stL          | 2ndL  | 3rdL    | 4thL         | 5thL         |
|----------------|---------------|-------|---------|--------------|--------------|
| Highest        | .eus / Kuwait | .eus  | Andorra | St Lucia     | St Lucia     |
| Lowest         | .lat          | India | Quebec  | N. Caledonia | N. Caledonia |

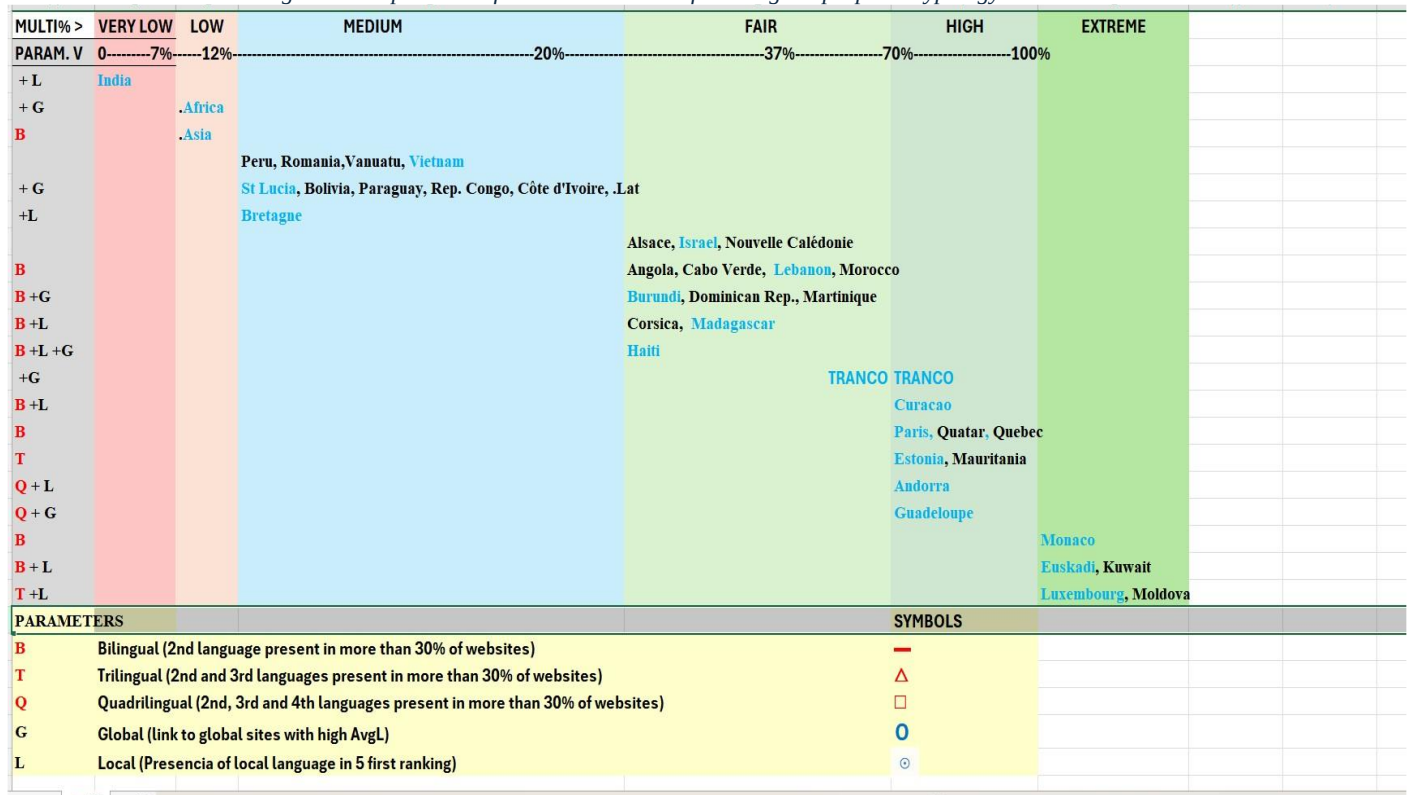
### 3-TYPEOLOGY

The following main parameters has been identified as relevant for a possible typology of the different cases possible of web multilingualism.

- **Percentage of web multilingualism** defined by the following scale:  
**Very Low      Low      Medium      Fair      High      Extreme**  
0-----7%-----12%-----20%-----37%-----70%-----100%  
The threshold between fair and high has been set at 37% which is the percentage found for TRANCO.
- Type of multilingualism : **Bilingual, trilingual, quadrilingual**, set when the last language hold more than 30% of websites
- Amount of multilingualism: **global** when high AvgL pushed by Global highly multilingual websites (often in redirection)
- Presence of a local language in the 5 first results : **local**.

The types will be set by the combination of those 4 parameters. For the whole group of 45 TLDs analyzed we obtained the following split into different types:

Figure 1:: Repartition of the measured TLD following the proposed typology



## 4- WEB MULTILINGUALISM TOPOLOGY BY EXAMPLES

For each identified category, one example is exposed with more details. The presentation is made from the most multilingual to the less multilingual web series.

### 4.1 Luxembourg (.lu) : extreme, trilingual, local

With Moldova, Luxembourg presents the highest profile of web multilingualism of the studied series. Trilingual with French, English and German, with local language Luxembourgish following and Dutch close behind. There are 95 languages additional behind the 11 first ones, a percentage of multilingual sites of 70% well balanced and a relatively high AvgL based on the 3 main languages and some proximity to other European languages rather than highly multilingual global sites, the maximum number of languages for one site reaching however 75. The profile of **Moldova** (.md) is close with a slightly higher %Multi of 74% and a more concentrated trilingual domain with Romanian, Russian and English, followed by Italian, German, French, Ukrainian, Spanish and Portuguese with %Sites between 3% and 2%.

Table 9: Indicators for Luxembourg (.lu)

|                      | %PAGES         | %SITES         |                    |               |
|----------------------|----------------|----------------|--------------------|---------------|
| <b>TOTAL</b>         | <b>100.00%</b> | <b>302.24%</b> | <b>MultiR</b>      | <b>3.02</b>   |
| <b>French</b>        | <b>26.83%</b>  | <b>81.09%</b>  | <b>%Multi</b>      | <b>70.36%</b> |
| <b>English</b>       | <b>21.34%</b>  | <b>64.51%</b>  | <b>AvgL</b>        | <b>3.87</b>   |
| <b>German</b>        | <b>17.04%</b>  | <b>51.51%</b>  | <b>VarCoef</b>     | <b>2.40</b>   |
| <b>Luxembourgish</b> | <b>4.24%</b>   | <b>12.83%</b>  | <b>GooT</b>        | <b>0.40%</b>  |
| <b>Dutch</b>         | <b>3.67%</b>   | <b>11.09%</b>  | <b>IDN</b>         | <b>0.05%</b>  |
| <b>Portuguese</b>    | <b>2.60%</b>   | <b>7.85%</b>   | <b>MAX L/sites</b> | <b>75</b>     |
| <b>Spanish</b>       | <b>2.44%</b>   | <b>7.37%</b>   | <b>%MONO</b>       | <b>29.6%</b>  |
| <b>Italian</b>       | <b>2.38%</b>   | <b>7.18%</b>   | <b>%2LINGUAL</b>   | <b>28.6%</b>  |
| <b>Polish</b>        | <b>1.29%</b>   | <b>3.89%</b>   | <b>%3LINGUAL</b>   | <b>25.0%</b>  |
| <b>Swedish</b>       | <b>1.13%</b>   | <b>3.41%</b>   | <b>%4LINGUAL</b>   | <b>8.3%</b>   |
| <b>Danish</b>        | <b>1.09%</b>   | <b>3.30%</b>   | <b>%LOCAL</b>      | <b>12.8%</b>  |
| <b>95 Lang.</b>      | <b>14.44%</b>  | <b>43.65%</b>  |                    |               |

Figure 2: Histogram for Luxembourg AvgL 1/2

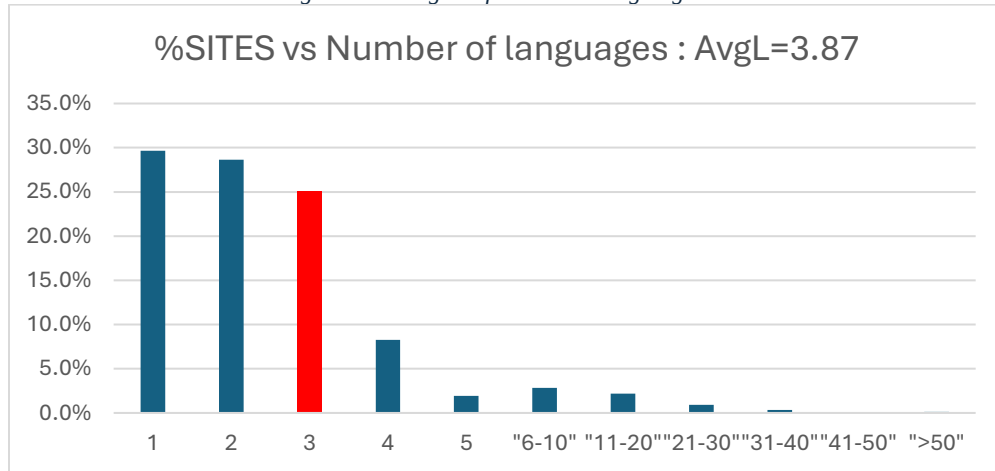
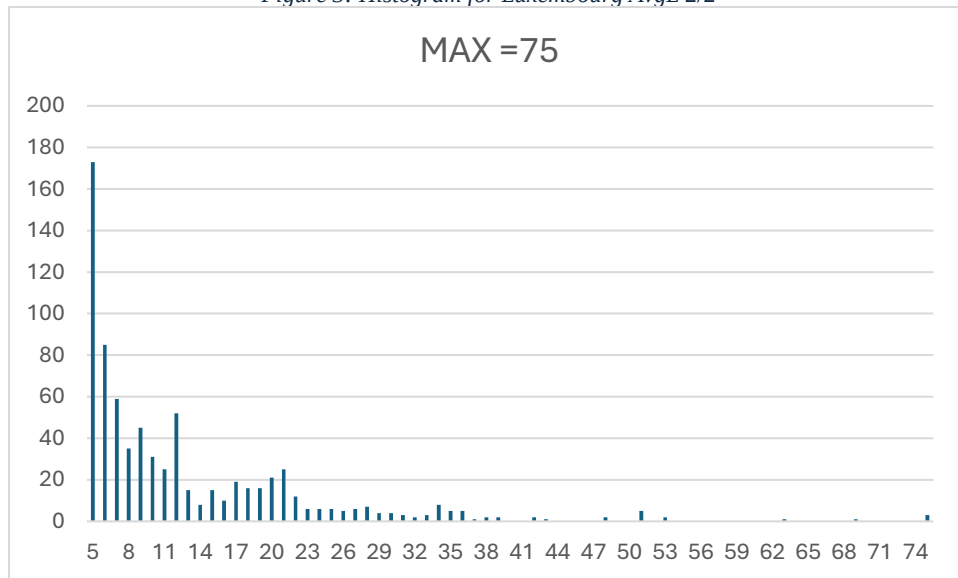


Figure 3: Histogram for Luxembourg AvgL 2/2



## 4.2 Euskadi (.eus): extreme, bilingual, local

The gTLD .eus is the champion of percentage of multilingual sites, with a figure close to 100% and almost perfect bilingualism Spanish, Basque, with more than 94% of sites having both languages, English, with a figure close to classify as trilingual, and French follow with respectively 28% and 17% of sites. The following language is Catalan showing some sort of solidarity between local languages. **Kuwait**, in the same category, with a lower Multi%, presents a strong bilingualism with predominance of English over Arabic.

Table 10: Indicators for Euskadi (.eus)

|                   | %PAGES         | %SITES         |                  |              |
|-------------------|----------------|----------------|------------------|--------------|
| <b>TOTAL</b>      | <b>100.00%</b> | <b>258.51%</b> | <b>MultiR</b>    | <b>2.59</b>  |
| <b>Spanish</b>    | <b>36.51%</b>  | <b>94.39%</b>  | <b>%Multi</b>    | <b>97.2%</b> |
| <b>Basque</b>     | <b>36.37%</b>  | <b>94.02%</b>  | <b>AVGL</b>      | <b>2.63</b>  |
| <b>English</b>    | <b>10.71%</b>  | <b>27.69%</b>  | <b>VAR</b>       | <b>1.62</b>  |
| <b>French</b>     | <b>6.74%</b>   | <b>17.42%</b>  | <b>MAX</b>       | <b>121</b>   |
| <b>Catalan</b>    | <b>1.00%</b>   | <b>2.58%</b>   | <b>%MONO</b>     | <b>2.8%</b>  |
| <b>German</b>     | <b>0.93%</b>   | <b>2.41%</b>   | <b>%2LINGUAL</b> | <b>68.7%</b> |
| <b>Italian</b>    | <b>0.53%</b>   | <b>1.36%</b>   | <b>%3LINGUAL</b> | <b>16.7%</b> |
| <b>Portuguese</b> | <b>0.48%</b>   | <b>1.25%</b>   | <b>%4LINGUAL</b> | <b>8.7%</b>  |
| <b>116 Lang.</b>  | <b>6.72%</b>   | <b>17.38%</b>  | <b>%LOCAL</b>    | <b>94.0%</b> |

Figure 4: Histogram for Euskadi AvgL 1/2

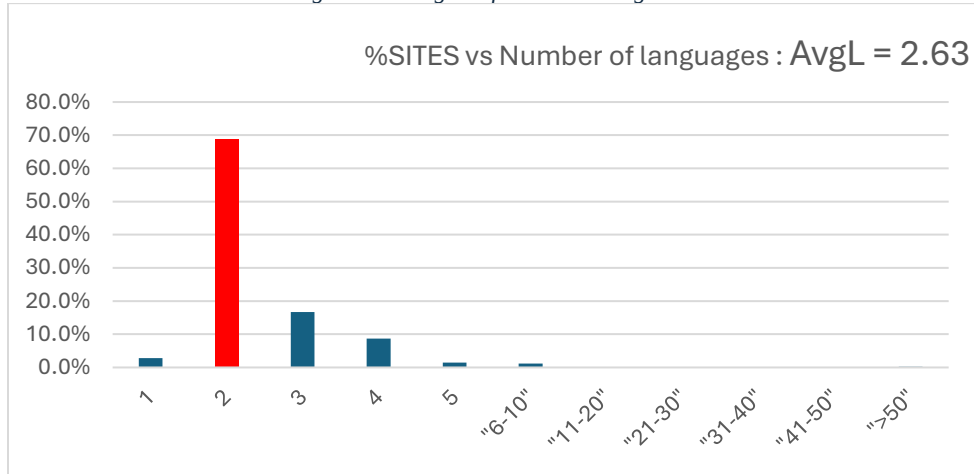
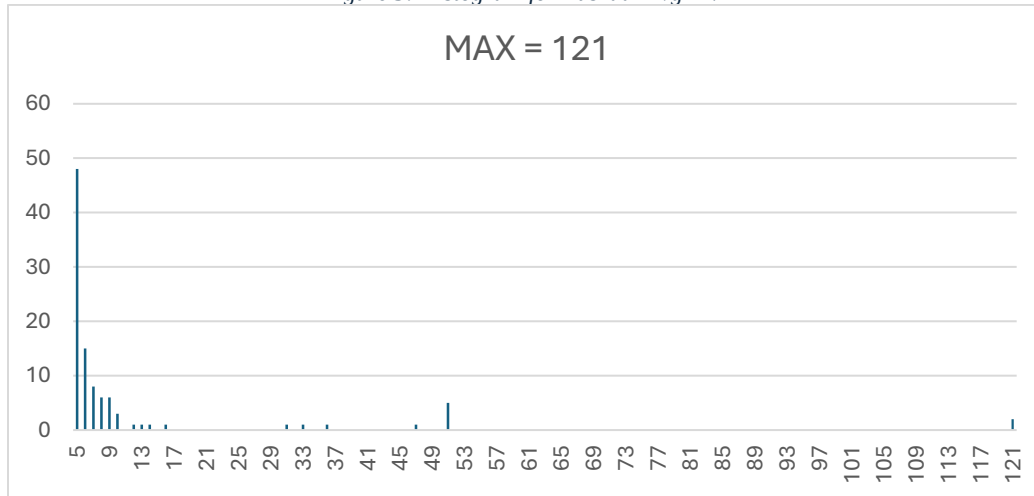


Figure 5: Histogram for Euskadi AvgL 2/2



### 4.3 Monaco (.mc): extreme, bilingual

Monaco is one of the top places in terms of %Multi with more than 75%, with a strong bilingual web ecosystem (French, English) and close to be trilingual with Italian almost at 30%. Monegasque (lij), the official language of the principate and the language of the national anthem, is surprisingly totally absent, in spite of having more speakers than English (22% versus 10%). Monegasque is spoken in Italy under the name of Ligurian by 140 000 persons. Occitan (oci), also spoken by more persons than English in Monaco is absent. The very limited number of additional languages is also characteristic of the .mc.

Table 11: Indicators for Monaco (.mc)

|                 | %PAGES         | %SITES         |                  |               |
|-----------------|----------------|----------------|------------------|---------------|
| <b>TOTAL</b>    | <b>100.00%</b> | <b>256.55%</b> |                  |               |
| <b>French</b>   | <b>34.99%</b>  | <b>89.77%</b>  | <b>MultiR</b>    | <b>2.57</b>   |
| <b>English</b>  | <b>32.22%</b>  | <b>82.65%</b>  | <b>%Multi</b>    | <b>75.43%</b> |
| <b>Italian</b>  | <b>10.33%</b>  | <b>26.49%</b>  | <b>AvgL</b>      | <b>3.08</b>   |
| <b>German</b>   | <b>3.24%</b>   | <b>8.30%</b>   | <b>VAR</b>       | <b>0.79</b>   |
| <b>Russian</b>  | <b>3.14%</b>   | <b>8.06%</b>   | <b>MAX</b>       | <b>21</b>     |
| <b>Spanish</b>  | <b>2.31%</b>   | <b>5.93%</b>   | <b>%MONO</b>     | <b>24.6%</b>  |
| <b>Chinese</b>  | <b>1.85%</b>   | <b>4.74%</b>   | <b>%2LINGUAL</b> | <b>44.1%</b>  |
| <b>Dutch</b>    | <b>1.60%</b>   | <b>4.10%</b>   | <b>%3LINGUAL</b> | <b>18.3%</b>  |
| <b>Japanese</b> | <b>1.48%</b>   | <b>3.80%</b>   | <b>%4LINGUAL</b> | <b>5.2%</b>   |
| <b>Arabic</b>   | <b>1.29%</b>   | <b>3.32%</b>   | <b>%LOCAL</b>    | <b>0.0%</b>   |
| <b>Polish</b>   | <b>1.02%</b>   | <b>2.61%</b>   |                  |               |
| <b>22 Lang.</b> | <b>6.54%</b>   | <b>16.77%</b>  |                  |               |

Figure 6: Histogram for Monaco AvgL 1/2

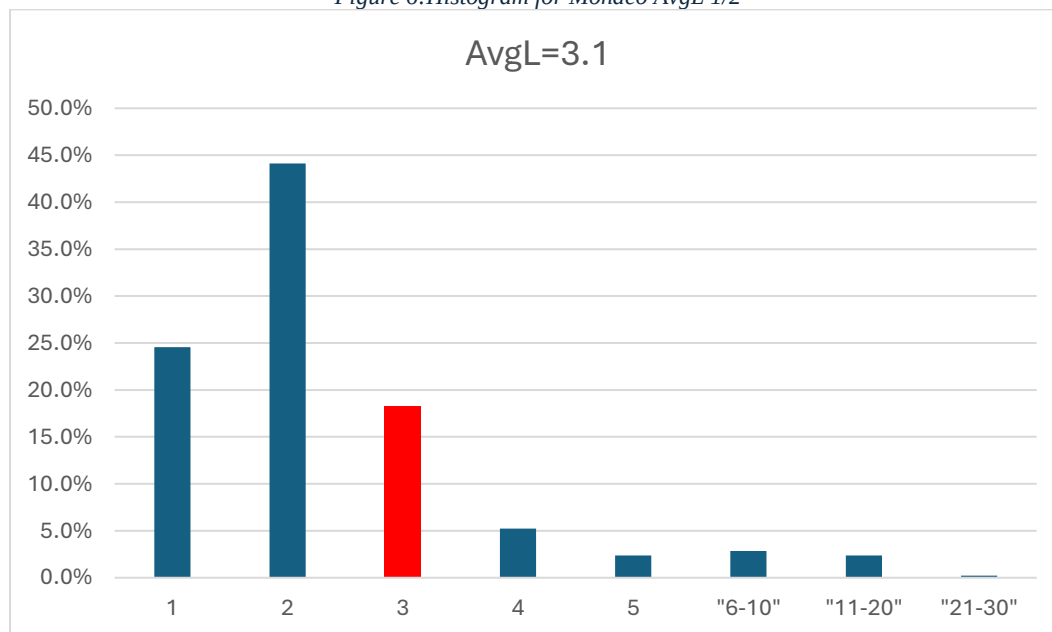
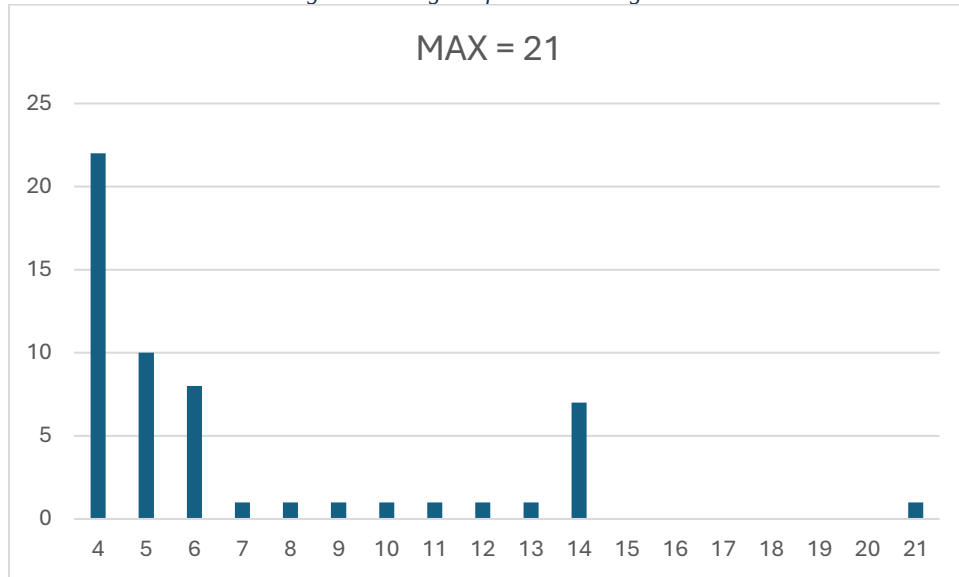


Figure 7: Histogram for Monaco AvgL 1/2



#### 4.4 Guadeloupe (.gp): High, quadrilingual, global

This is an extreme case of multilingualism in terms, not on percentage of multilingual sites, but in terms of languages, tagged as quadrilingual, in fact it reaches 5 languages following our definition. The intense international tourism vocation of the island is the probable explanation of this characteristic, marked by a very high AvgL and the presence of many global websites by redirection from .gp domains. On the negative side, no trace of Guadeloupean creole (some pages may share this local language with French or English but they are not accounted by MECILDI).

Table 12: Indicators for Guadeloupe (.gp)

|                   | %PAGES         | %SITES         | MultiR           | 6.06          |
|-------------------|----------------|----------------|------------------|---------------|
| <b>TOTAL</b>      | <b>100.00%</b> | <b>606.16%</b> | <b>%Multi</b>    | <b>61.18%</b> |
| <b>English</b>    | <b>12.83%</b>  | <b>77.78%</b>  | <b>AvgL</b>      | <b>9.28</b>   |
| <b>French</b>     | <b>12.59%</b>  | <b>76.29%</b>  | <b>VAR</b>       | <b>0.58</b>   |
| <b>German</b>     | <b>7.37%</b>   | <b>44.65%</b>  | <b>MAX</b>       | <b>51</b>     |
| <b>Spanish</b>    | <b>6.42%</b>   | <b>38.91%</b>  | <b>%MONO</b>     | <b>38.7%</b>  |
| <b>Italian</b>    | <b>5.04%</b>   | <b>30.55%</b>  | <b>%2LINGUAL</b> | <b>11.0%</b>  |
| <b>Dutch</b>      | <b>4.15%</b>   | <b>25.18%</b>  | <b>%3LINGUAL</b> | <b>11.8%</b>  |
| <b>Portuguese</b> | <b>3.95%</b>   | <b>23.95%</b>  | <b>%4LINGUAL</b> | <b>1.8%</b>   |
| <b>Japanese</b>   | <b>3.06%</b>   | <b>18.58%</b>  | <b>%LOCAL</b>    |               |
| <b>Polish</b>     | <b>2.90%</b>   | <b>17.61%</b>  |                  |               |
| <b>Chinese</b>    | <b>2.54%</b>   | <b>15.40%</b>  |                  |               |
| <b>Korean</b>     | <b>2.11%</b>   | <b>12.76%</b>  |                  |               |
| <b>66 Lang.</b>   | <b>37.04%</b>  | <b>224.51%</b> |                  |               |

Figure 8: Histogram for Guadeloupe AvgL 1/2

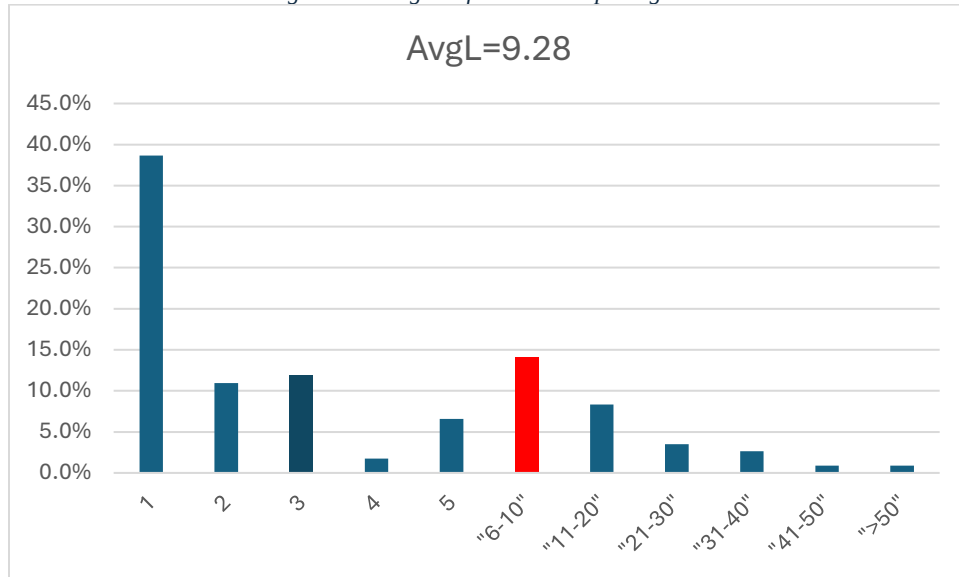
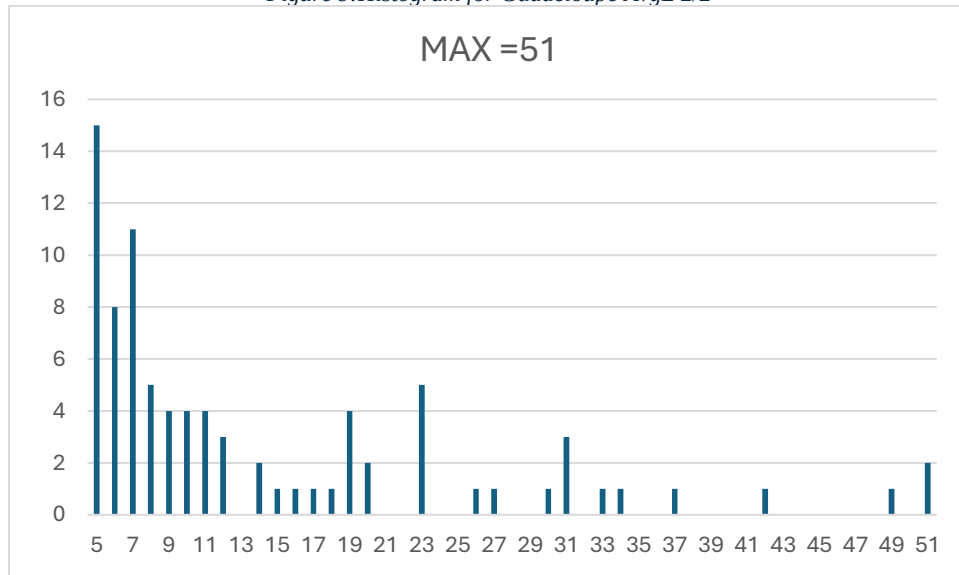


Figure 9: Histogram for Guadeloupe AvgL 2/2



#### 4.5 Andorra (.ad): High, quadrilingual, local

Spectacular quadrilingual domain with relatively high percentages for more languages. High %Multi and AvgL. Domination of Catalan.

Table 13: Indicators for Andorra (.mc)

|                   | %PAGES         | %SITES         | MultiR    | 2.97          |
|-------------------|----------------|----------------|-----------|---------------|
| <b>TOTAL</b>      | <b>100.00%</b> | <b>297.19%</b> | %Multi    | <b>56.93%</b> |
| <b>Catalan</b>    | <b>20.75%</b>  | <b>61.68%</b>  | GooT      | <b>0.91%</b>  |
| <b>English</b>    | <b>19.63%</b>  | <b>58.33%</b>  | IDN       | <b>0.00%</b>  |
| <b>Spanish</b>    | <b>18.35%</b>  | <b>54.53%</b>  | AVGL      | <b>4.46</b>   |
| <b>French</b>     | <b>14.17%</b>  | <b>42.13%</b>  | VAR       | <b>0.99</b>   |
| <b>German</b>     | <b>2.77%</b>   | <b>8.23%</b>   | MAX       | <b>40</b>     |
| <b>Vietnamese</b> | <b>2.22%</b>   | <b>6.60%</b>   | %MONO     | <b>43.1%</b>  |
| <b>Portuguese</b> | <b>2.17%</b>   | <b>6.44%</b>   | %2LINGUAL | <b>11.1%</b>  |
| <b>Italian</b>    | <b>1.94%</b>   | <b>5.78%</b>   | %3LINGUAL | <b>13.4%</b>  |
| <b>Russian</b>    | <b>1.31%</b>   | <b>3.90%</b>   | %4LINGUAL | <b>23.3%</b>  |
| <b>Dutch</b>      | <b>1.11%</b>   | <b>3.30%</b>   | %LOCAL    | <b>61.7%</b>  |
| <b>Chinese</b>    | <b>1.00%</b>   | <b>2.96%</b>   |           |               |
| <b>49 Lang.</b>   | <b>14.58%</b>  | <b>43.32%</b>  |           |               |

Figure 10: Histogram for Andorra AvgL ½

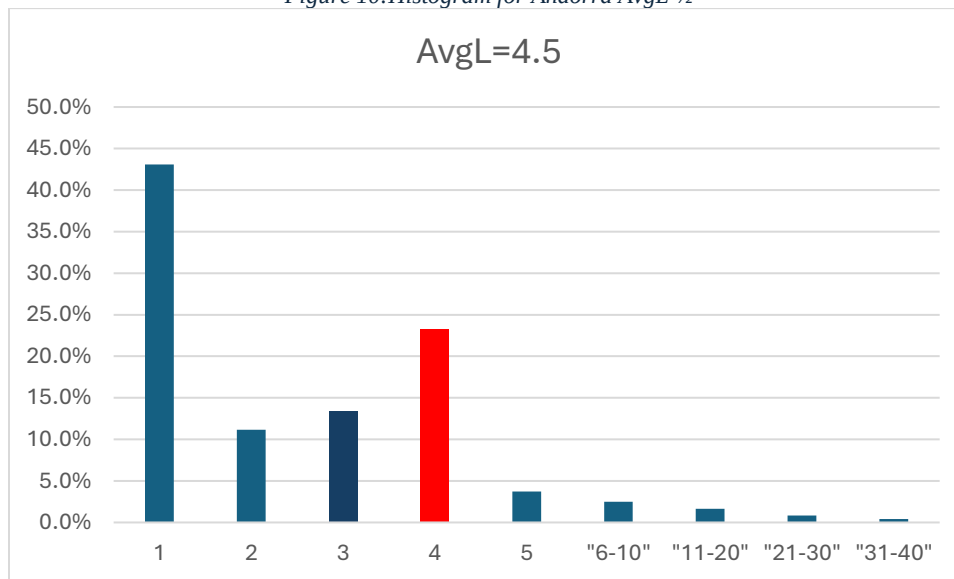
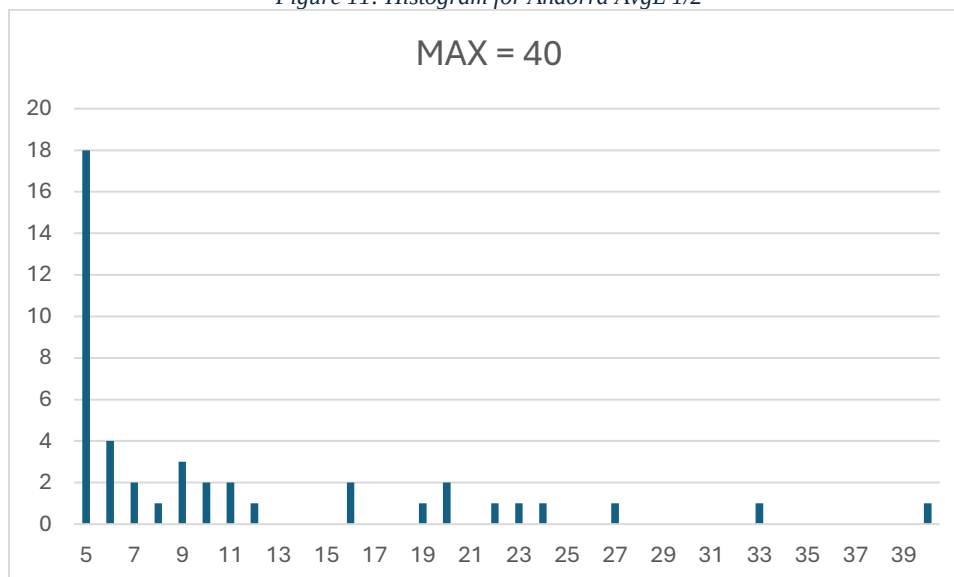


Figure 11: Histogram for Andorra AvgL 1/2



#### 4.6 Estonia (.ad): High, trilingual, local

Estonia presents a profile very close to trilingual, with Russian near 30%. It shows a high percentage of multilingual websites, above 50%. It is noticeable that the top languages are those of neighbor countries.

Mauritania (.mr), the other case in the same category, is trilingual with French (75%), Arabic (61%) and English (36%) and practically the same %Multi (57%). It is interesting to observe than 2 countries so different, one small country leader on information society and the second one some sort of mediator in a region involved in many conflicts could share as much this category.

Table 14: Indicators for Estonia (ee)

|                   | %PAGES         | %SITES         | MultiR           | 2.44          |
|-------------------|----------------|----------------|------------------|---------------|
| <b>TOTAL</b>      | <b>100.00%</b> | <b>243.76%</b> | <b>%Multi</b>    | <b>55.60%</b> |
| <b>Estonian</b>   | <b>36.35%</b>  | <b>88.61%</b>  |                  |               |
| <b>English</b>    | <b>21.44%</b>  | <b>52.27%</b>  |                  |               |
| <b>Russian</b>    | <b>11.63%</b>  | <b>28.36%</b>  | <b>AVGL</b>      | <b>3.59</b>   |
| <b>Finnish</b>    | <b>4.36%</b>   | <b>10.64%</b>  | <b>VAR</b>       | <b>1.04</b>   |
| <b>Latvian</b>    | <b>2.42%</b>   | <b>5.90%</b>   | <b>MAX</b>       | <b>79</b>     |
| <b>German</b>     | <b>1.96%</b>   | <b>4.78%</b>   | <b>%MONO</b>     | <b>44.4%</b>  |
| <b>Lithuanian</b> | <b>1.92%</b>   | <b>4.67%</b>   | <b>%2LINGUAL</b> | <b>29.0%</b>  |
| <b>Swedish</b>    | <b>1.58%</b>   | <b>3.85%</b>   | <b>%3LINGUAL</b> | <b>15.4%</b>  |
| <b>Spanish</b>    | <b>1.22%</b>   | <b>2.98%</b>   | <b>%4LINGUAL</b> | <b>4.7%</b>   |
| <b>French</b>     | <b>1.21%</b>   | <b>2.95%</b>   | <b>%LOCAL</b>    | <b>88.6%</b>  |
| <b>Polish</b>     | <b>1.14%</b>   | <b>2.78%</b>   |                  |               |
| <b>96 Lang.</b>   | <b>12.88%</b>  | <b>31.40%</b>  |                  |               |

Figure 12:Histogram for Estonia AvgL 1/2

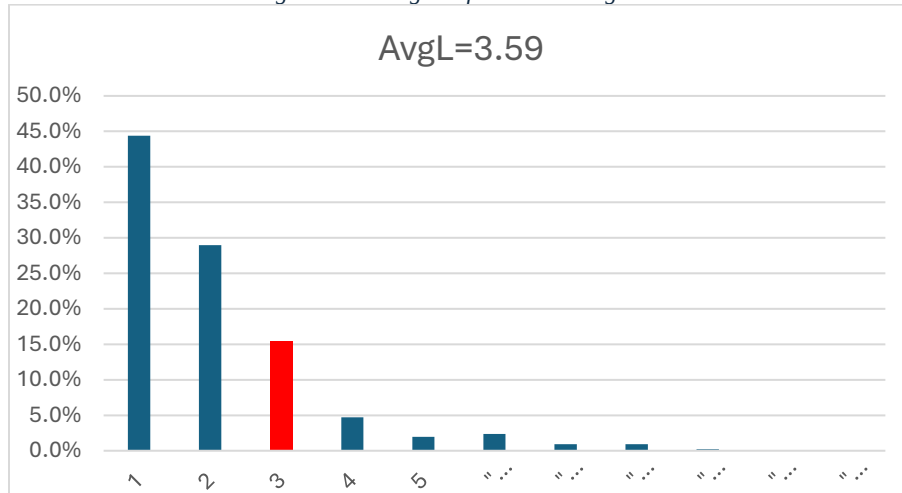
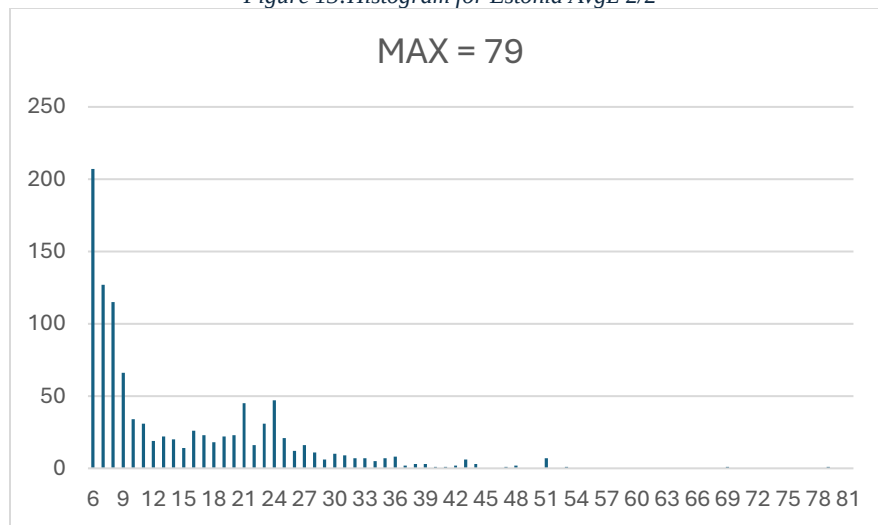


Figure 13:Histogram for Estonia AvgL 2/2



## 4.7 Paris (.paris): High, bilingual

High level of web multilingualism, bilingualism French, English, with 2 languages (German and Spanish) present in high percentages of sites and long list of languages with results higher than 2%. All that characterizing an international tourism vocation. Quebec is probably the most balanced bilingual TLD ever made, and also a champion in the use of IDN. Qatar shows a strong bilingualism with English domination.

Table 15: Indicators for Paris (.paris)

|                     |        |               |
|---------------------|--------|---------------|
| <b>French</b>       | 30.36% | <b>76.36%</b> |
| <b>English</b>      | 26.49% | <b>66.62%</b> |
| <b>German</b>       | 6.44%  | <b>16.20%</b> |
| <b>Spanish</b>      | 4.85%  | <b>12.20%</b> |
| <b>Italian</b>      | 3.60%  | <b>9.06%</b>  |
| <b>Portuguese</b>   | 2.24%  | <b>5.65%</b>  |
| <b>Dutch</b>        | 2.02%  | <b>5.09%</b>  |
| <b>Chinese</b>      | 2.00%  | <b>5.04%</b>  |
| <b>Japanese</b>     | 1.90%  | <b>4.79%</b>  |
| <b>Russian</b>      | 1.37%  | <b>3.45%</b>  |
| <b>Korean</b>       | 1.00%  | <b>2.51%</b>  |
| <b>Polish</b>       | 0.98%  | <b>2.46%</b>  |
| <b>Arabic</b>       | 0.87%  | <b>2.19%</b>  |
| <b>Czech</b>        | 0.84%  | <b>2.12%</b>  |
| <b>Modern Greek</b> | 0.83%  | <b>2.09%</b>  |

|                  |               |
|------------------|---------------|
| <b>MultiR</b>    | <b>2.51</b>   |
| <b>%Multi</b>    | <b>48.31%</b> |
| <b>GooT</b>      | <b>0.38%</b>  |
| <b>IDN</b>       | <b>0.00%</b>  |
| <b>AVGL</b>      | <b>2.51</b>   |
| <b>VAR</b>       | <b>1.50</b>   |
| <b>MAX</b>       | <b>102</b>    |
| <b>%MONO</b>     | <b>51.7%</b>  |
| <b>%2LINGUAL</b> | <b>27.1%</b>  |
| <b>%3LINGUAL</b> | <b>9.6%</b>   |
| <b>%4LINGUAL</b> | <b>2.5%</b>   |

Figure 14: Histogram for Paris AvgL 1/2

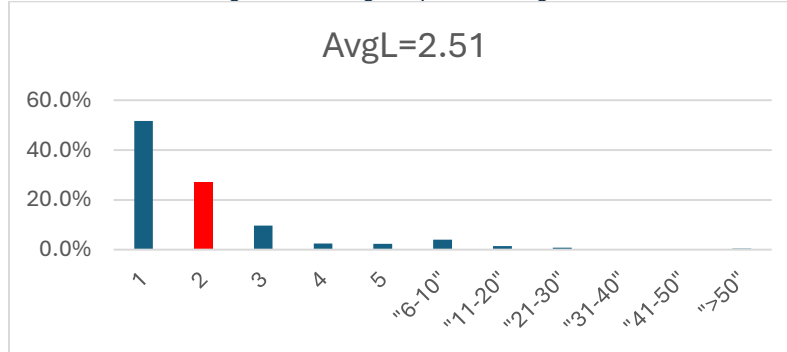
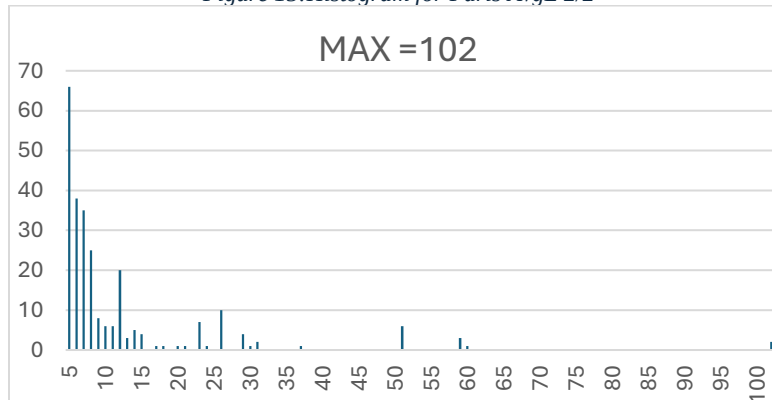


Figure 15: Histogram for Paris AvgL 2/2



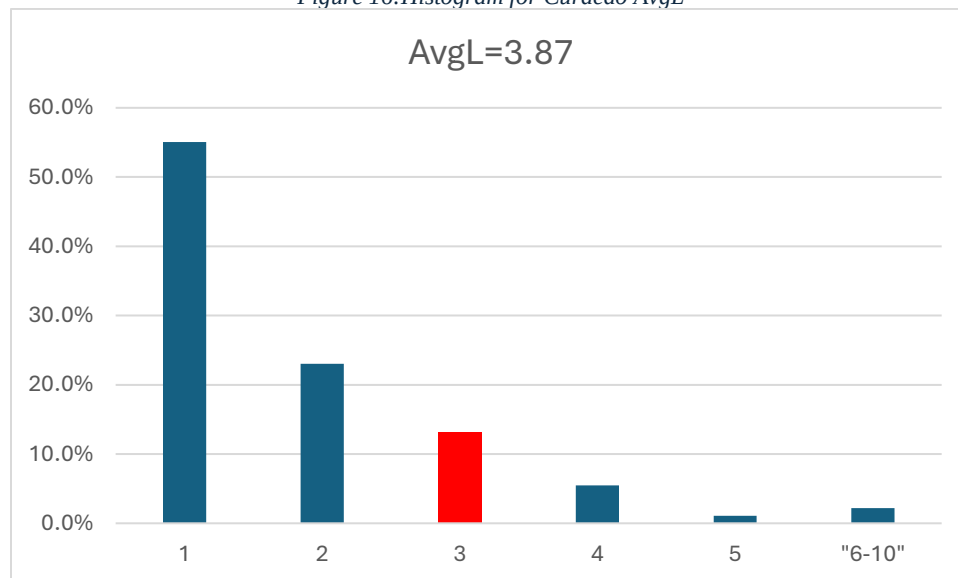
## 4.8 Curacao (.cw): High, bilingual, local

The singularity of Curacao is a very limited number of languages and of the maximum number of languages per website.

Table 16: Indicators for Curacao (.cw)

|                   | %PAGES         | %SITES         | MultiR           | 1.60          |
|-------------------|----------------|----------------|------------------|---------------|
| <b>TOTAL</b>      | <b>100.00%</b> | <b>160.34%</b> | <b>%Multi</b>    | <b>38.18%</b> |
| <b>English</b>    | <b>48.28%</b>  | <b>77.41%</b>  | <b>AvgL</b>      | <b>2.58</b>   |
| <b>Dutch</b>      | <b>31.82%</b>  | <b>51.02%</b>  | <b>GooT</b>      | <b>0.49%</b>  |
| <b>Papiamentu</b> | <b>11.52%</b>  | <b>18.47%</b>  | <b>IDN</b>       | <b>0.00%</b>  |
| <b>Spanish</b>    | <b>4.51%</b>   | <b>7.24%</b>   | <b>VAR</b>       | <b>0.48</b>   |
| <b>German</b>     | <b>0.77%</b>   | <b>1.23%</b>   | <b>MAX</b>       | <b>6</b>      |
| <b>Nepali</b>     | <b>0.77%</b>   | <b>1.23%</b>   | <b>%MONO</b>     | <b>55.0%</b>  |
| <b>Chinese</b>    | <b>0.77%</b>   | <b>1.23%</b>   | <b>%2LINGUAL</b> | <b>23.0%</b>  |
| <b>7 Lang</b>     | <b>1.57%</b>   | <b>2.52%</b>   | <b>%3LINGUAL</b> | <b>13.2%</b>  |
|                   |                |                | <b>%4LINGUAL</b> | <b>5.5%</b>   |
|                   |                |                | <b>%LOCAL</b>    | <b>18.5%</b>  |

Figure 16: Histogram for Curacao AvgL



## 4.9 TRANCO: Fair/High boundary, global

Tranco, the mix of the million most visited websites, by definition, is the essence of global and it could serve of a median reference. It is on the high range of the Fair scale with 37% of Multi%, has a large AvgL and show almost equal percentages of mono, bilingual y trilingual websites.

Table 17: Indicators for Tranco

|                 | %PAGES         | %SITES         |  | MultiR    | 3.41          |
|-----------------|----------------|----------------|--|-----------|---------------|
| <b>TOTAL</b>    | <b>100.00%</b> | <b>341.48%</b> |  | %Multi    | <b>37.27%</b> |
| English         | 19.77%         | 67.52%         |  | AvgL      | <b>7.47</b>   |
| German          | 6.44%          | 21.98%         |  | VarCoef   | <b>1.39</b>   |
| French          | 6.07%          | 20.73%         |  | MonoL     | <b>62.73%</b> |
| Spanish         | 6.02%          | 20.54%         |  | GooT      | <b>1.06%</b>  |
| Italian         | 4.07%          | 13.91%         |  | IDN       |               |
| Portuguese      | 3.85%          | 13.15%         |  | MAX       | <b>75</b>     |
| Russian         | 3.55%          | 12.11%         |  | %MONO     | <b>29.6%</b>  |
| Dutch           | 3.22%          | 10.98%         |  | %2LINGUAL | <b>28.6%</b>  |
| Japanese        | 2.97%          | 10.14%         |  | %3LINGUAL | <b>25.0%</b>  |
| Chinese         | 2.90%          | 9.91%          |  | %4LINGUAL | <b>8.3%</b>   |
| Polish          | 2.63%          | 8.98%          |  | %LOCAL    | <b>12.8%</b>  |
| <b>138 Lang</b> | <b>38.51%</b>  | <b>131.52%</b> |  |           |               |

Figure 17: Histogram for Tranco AvgL1/2

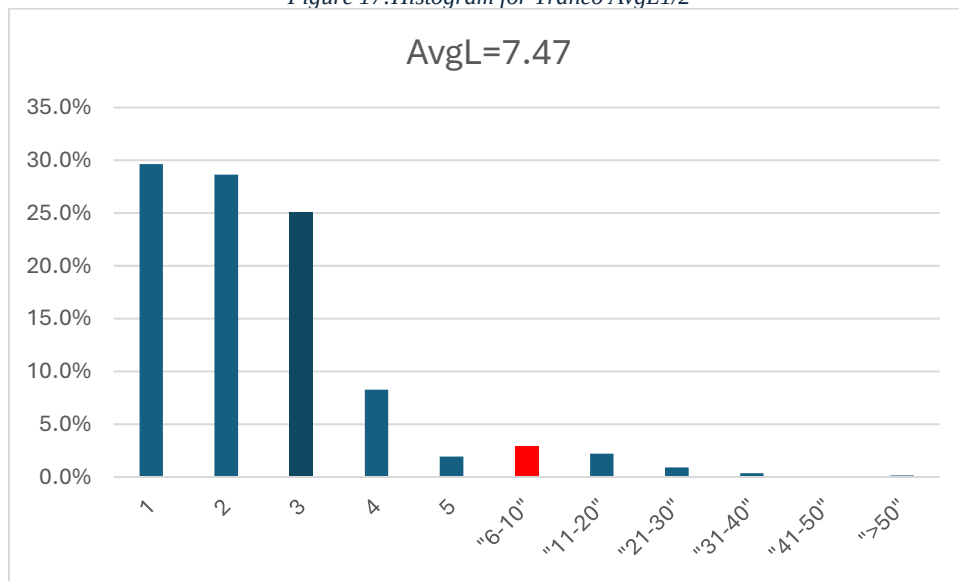
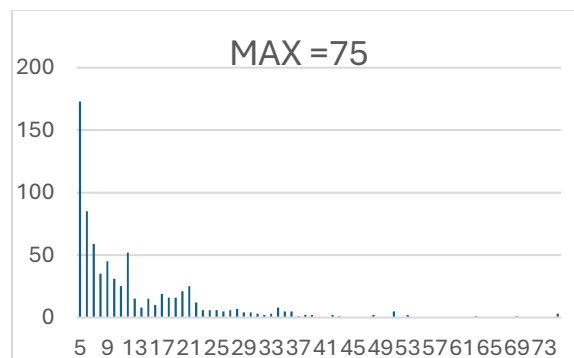


Figure 18: Histogram for Tranco AvgL2/2



## 4.10 Haiti (.ht): Fair, bilingual, local, global

Haiti shows a balanced bilingual pattern between English and French, a high AvgL pushed by global websites and a local presence in position 5.

Table 18: Indicators for Haiti (.ht)

|                       | %PAGES         | %SITES         |                  |               |
|-----------------------|----------------|----------------|------------------|---------------|
| <b>TOTAL</b>          | <b>100.00%</b> | <b>231.10%</b> | <b>MultiR</b>    | <b>2.31</b>   |
| <b>français</b>       | <b>24.59%</b>  | <b>56.84%</b>  | <b>%Multi</b>    | <b>23.63%</b> |
| <b>anglais</b>        | <b>22.81%</b>  | <b>52.72%</b>  | <b>MonoL</b>     | <b>76.37%</b> |
| espagnol              | 4.18%          | 9.66%          | <b>GooT</b>      | <b>1.22%</b>  |
| allemand              | 3.97%          | 9.16%          | <b>IDN</b>       | <b>0.00%</b>  |
| <b>créole haïtien</b> | <b>3.41%</b>   | <b>7.89%</b>   | <b>AVGL</b>      | <b>6.55</b>   |
| italien               | 2.97%          | 6.86%          | <b>VAR</b>       | <b>4.06</b>   |
| portugais             | 2.97%          | 6.86%          | <b>MAX</b>       | <b>37</b>     |
| chinois               | 2.97%          | 6.86%          | <b>%MONO</b>     | <b>75.8%</b>  |
| japonais              | 2.31%          | 5.34%          | <b>%2LINGUAL</b> | <b>9.8%</b>   |
| néerlandais           | 1.98%          | 4.58%          | <b>%3LINGUAL</b> | <b>4.5%</b>   |
| <b>38 Lang</b>        | <b>27.83%</b>  | <b>64.32%</b>  | <b>%4LINGUAL</b> | <b>1.5%</b>   |
|                       |                |                | <b>%LOCAL</b>    | <b>7.9%</b>   |

Figure 19: Histogram for Haiti AvgL1/2

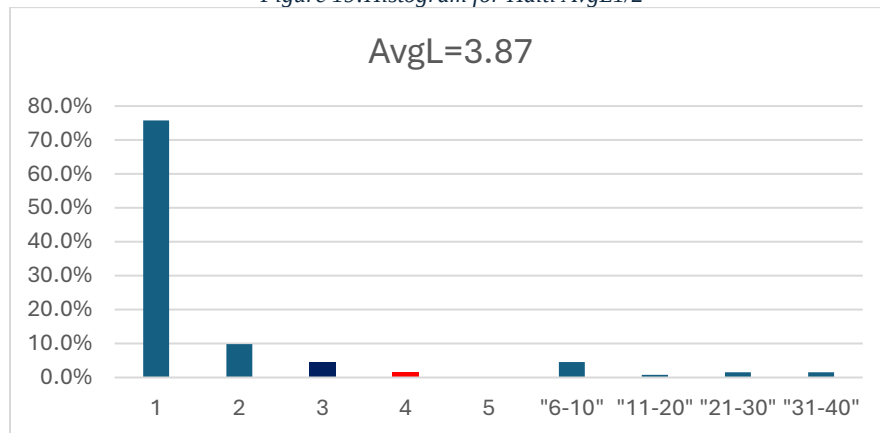
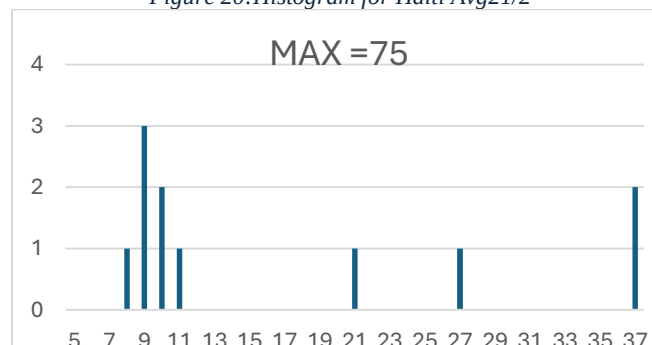


Figure 20: Histogram for Haiti Avg21/2



## 4.11 Madagascar (.mg): Fair, bilingual, local

Madagascar presents a bilingualism with French, English strong predominance of French and the local language in third position. Corsica, in the same category, have a lower local presence of 2% and in position 5, but share the same type of bilingualism.

Table 19: Indicators for Madagascar (.mg)

|              | %PAGES         | %SITES         | MultiR    | 1.65   |
|--------------|----------------|----------------|-----------|--------|
| <b>TOTAL</b> | <b>100.00%</b> | <b>164.54%</b> | %Multi    | 25.42% |
| français     | 46.90%         | 77.17%         | GooT      | 0.27%  |
| anglais      | 23.38%         | 38.48%         | IDN       | 0.00%  |
| malgache     | 4.61%          | 7.58%          | AvgL      | 3.54   |
| allemand     | 2.28%          | 3.75%          | VAR       | 1.75   |
| chinois      | 1.47%          | 2.43%          | MAX       | 37     |
| portugais    | 1.43%          | 2.36%          | %MONO     | 74.4%  |
| espagnol     | 1.34%          | 2.21%          | %2LINGUAL | 17.9%  |
| 50 Lang.     | 18.58%         | 30.57%         | %3LINGUAL | 4.9%   |
|              |                |                | %4LINGUAL | 0.7%   |
|              |                |                | %LOCAL    | 7.6%   |

Figure 21: Histogram for Madagascar AvgL1/2

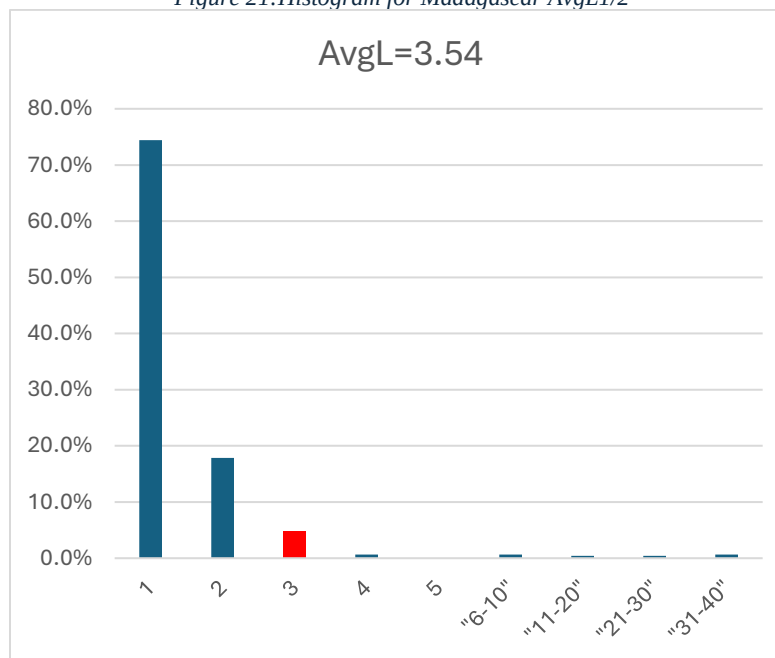
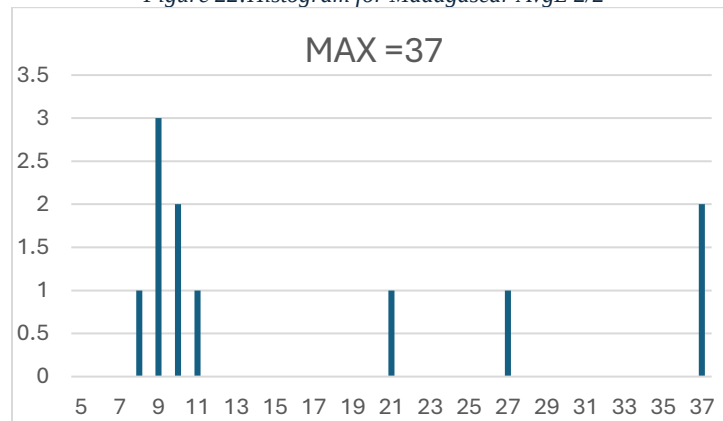


Figure 22:Histogram for Madagascar AvgL 2/2



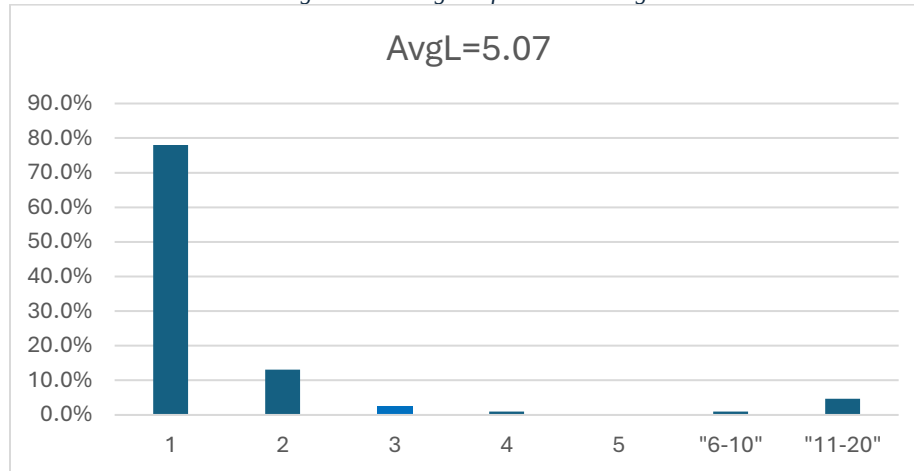
#### 4.12 Burundi (.bi): Fair, bilingual, global

Burundi is on the low range of the scale fair with a little more than 20%, a high AvgL indicates a large number of redirects towards global websites with many languages, although the total of different languages remains low. Bilingualism English, French, with German a strong third language. **Dominican Rep.** has a close profile with much more languages (111 versus 39) indicative of the Tourism orientation of the country. **Martinique** shows a still higher %Multi and AvgL witnessing a solid international tourism vocation, although not as reflected as sister island Guadeloupe.

Table 20:Indicators for Burundi (.bi)

|                 | %PAGES         | %SITES         | MultiR    | 1.88   |
|-----------------|----------------|----------------|-----------|--------|
| <b>TOTAL</b>    | <b>100.00%</b> | <b>187.88%</b> | %Multi    | 21.62% |
| anglais         | 28.26%         | 53.10%         | GooT      | 1.50%  |
| français        | 24.00%         | 45.10%         | IDN       | 0.00%  |
| allemand        | 5.96%          | 11.19%         | AVGL      | 5.07   |
| espagnol        | 3.34%          | 6.28%          | VAR       | 2.66   |
| russe           | 3.20%          | 6.02%          | MAX       | 18     |
| chinois         | 3.20%          | 6.02%          | %MONO     | 78.4%  |
| portugais       | 3.09%          | 5.80%          | %2LINGUAL | 13.2%  |
| italien         | 3.00%          | 5.64%          | %3LINGUAL | 2.3%   |
| japonais        | 2.01%          | 3.77%          | %4LINGUAL | 0.9%   |
| néerlandais     | 1.93%          | 3.63%          | %LOCAL    | 0.6%   |
| polonais        | 1.84%          | 3.45%          |           |        |
| indonésien      | 1.75%          | 3.29%          |           |        |
| coréen          | 1.75%          | 3.29%          |           |        |
| vietnamien      | 1.60%          | 3.00%          |           |        |
| <b>26 Lang.</b> | <b>15.06%</b>  | <b>28.29%</b>  |           |        |

Figure 23:Histogram for Burundi AvgL



#### 4.13 Lebanon (.lb): Fair, bilingual

Bilingual, English, Arabic with clear domination of English. French not close enough to qualify trilingual, limited number of languages. Angola is another in the same bilingual category with strong dominance of Portuguese over English, similar AvgL with more languages involved. Cabo Verde, not very different but with English dominating Portuguese, Finally, Morrocco, with a higher Multi%, French dominance over English and Arabic not so far to make it trilingual.

Table 21:Indicators for Lebanon (.lb)

|              | %PAGES         | %SITES         | MultiR    | 1.62   |
|--------------|----------------|----------------|-----------|--------|
| <b>TOTAL</b> | <b>100.00%</b> | <b>161.88%</b> | %Multi    | 24.25% |
| anglais      | 49.98%         | 80.91%         | GooT      | 0.62%  |
| arabe        | 21.25%         | 34.40%         | IDN       | 0.00%  |
| français     | 7.44%          | 12.05%         | AVGL      | 3.55   |
| allemand     | 1.34%          | 2.18%          | VAR       | 2.52   |
| italien      | 1.15%          | 1.87%          | MAX       | 31     |
| espagnol     | 1.15%          | 1.87%          | %MONO     | 75.5%  |
| 38 Lang      | 17.68%         | 28.62%         | %2LINGUAL | 17.7%  |
|              |                |                | %3LINGUAL | 3.4%   |
|              |                |                | %4LINGUAL | 0.3%   |
|              |                |                | %LOCAL    |        |

Figure 24: Histogram for Lebanon AvgL 1/2

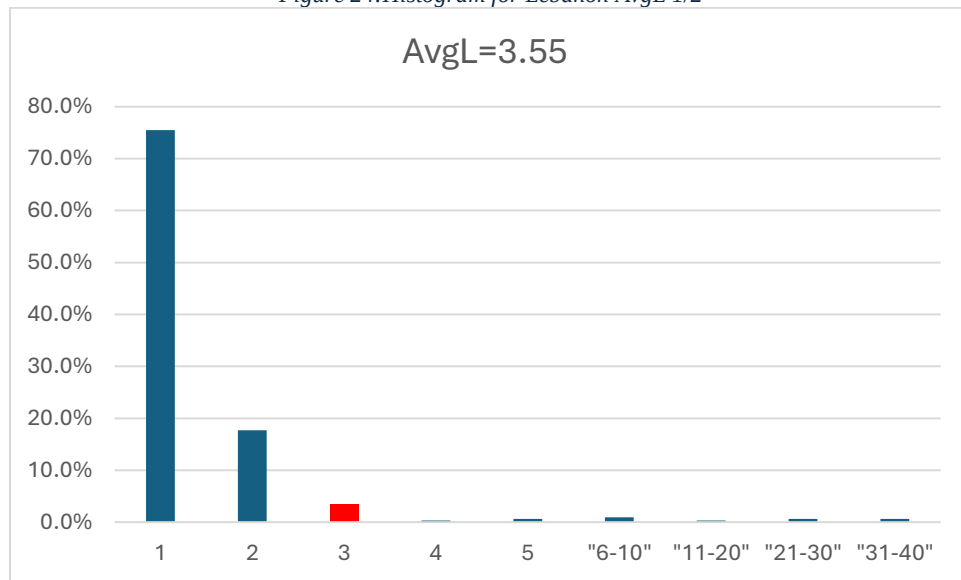
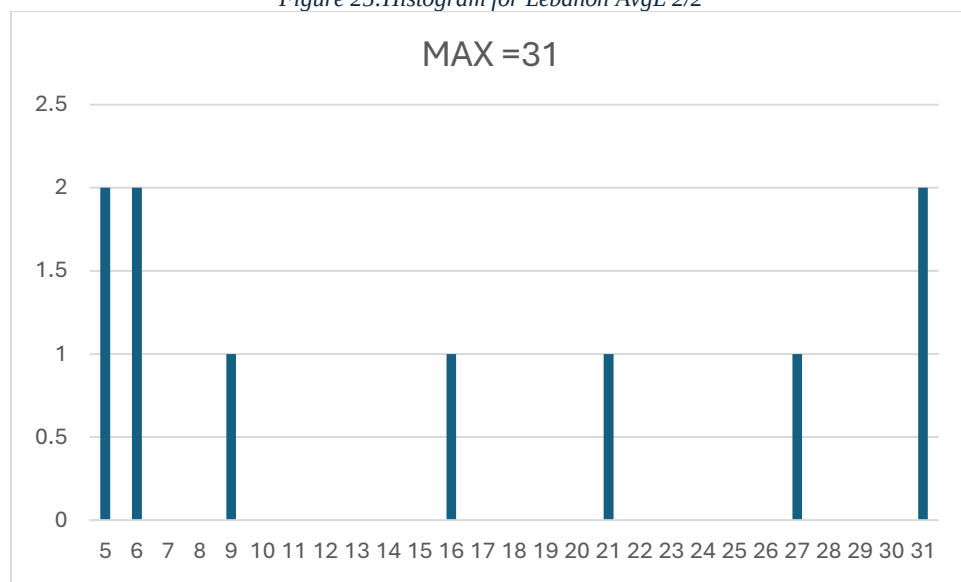


Figure 25: Histogram for Lebanon AvgL 2/2



#### 4.14 Israel (.il): Fair

In the bottom level of the fair category, with 22%, Israel shows a near bilingualism of Hebrew with English, followed by Russian above Arabic and French. The category is shared by .alsace and :nc (Nouvelle Calédonie).

Table 22: Indicators for Israel (.il)

|                 | %PAGES         | %SITES         | MultiR | 1.42   |
|-----------------|----------------|----------------|--------|--------|
| <b>TOTAL</b>    | <b>100.00%</b> | <b>141.57%</b> | %Multi | 21.96% |
| <b>Hebrew</b>   | <b>65.76%</b>  | <b>92.31%</b>  | AvgL   | 2.89   |
| <b>English</b>  | <b>17.96%</b>  | <b>25.35%</b>  | GooT   | 0.36%  |
| <b>Russian</b>  | <b>2.87%</b>   | <b>4.09%</b>   | IDN    | 0.35%  |
| <b>Arabic</b>   | <b>1.68%</b>   | <b>2.40%</b>   |        |        |
| <b>French</b>   | <b>1.48%</b>   | <b>2.12%</b>   |        |        |
| <b>Spanish</b>  | <b>0.93%</b>   | <b>1.36%</b>   |        |        |
| <b>86 Lang.</b> | <b>9.33%</b>   | <b>13.96%</b>  |        |        |

#### 4.15 Bretagne (.bzh): Medium, local

The ,bzh gTLD, with medium percentage of multilingual sites, show English as second language, close to bilingualism threshold, a low AvgL and the presence of local language in position 4.

Table 23: Indicators for Bretagne (.bzh)

|                 | %PAGES         | %SITES         | MultiR    | 1.44   |
|-----------------|----------------|----------------|-----------|--------|
| <b>TOTAL</b>    | <b>100.00%</b> | <b>144.09%</b> | %Multi    | 16.48% |
| <b>French</b>   | <b>58.60%</b>  | <b>84.43%</b>  | GooT      | 1.43%  |
| <b>English</b>  | <b>18.49%</b>  | <b>26.65%</b>  | IDN       | 0.00%  |
| <b>German</b>   | <b>4.02%</b>   | <b>5.79%</b>   | AVGL      | 1.44   |
| <b>Breton</b>   | <b>2.92%</b>   | <b>4.21%</b>   | VAR       | 2.37   |
| <b>Spanish</b>  | <b>2.09%</b>   | <b>3.01%</b>   | MAX       | 121    |
| <b>Italian</b>  | <b>1.40%</b>   | <b>2.01%</b>   | %MONO     | 83.5%  |
| <b>Dutch</b>    | <b>1.22%</b>   | <b>1.76%</b>   | %2LINGUAL | 9.2%   |
| <b>116 Lang</b> | <b>11.26%</b>  | <b>16.23%</b>  | %3LINGUAL | 4.2%   |
|                 |                |                | %4LINGUAL | 1.2%   |
|                 |                |                | %LOCAL    | 4.2%   |

Figure 26:Histogram for Bretagne AvgL 1/2

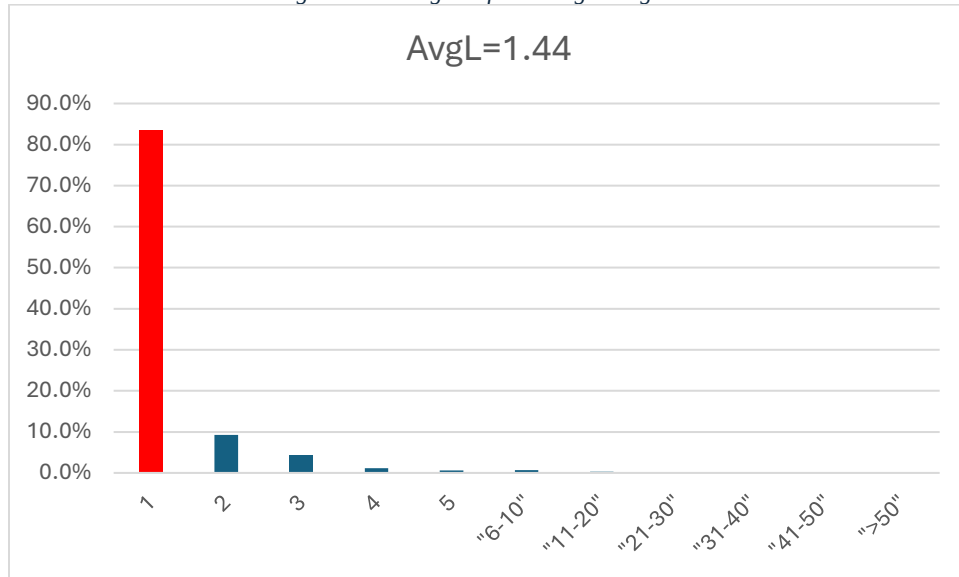
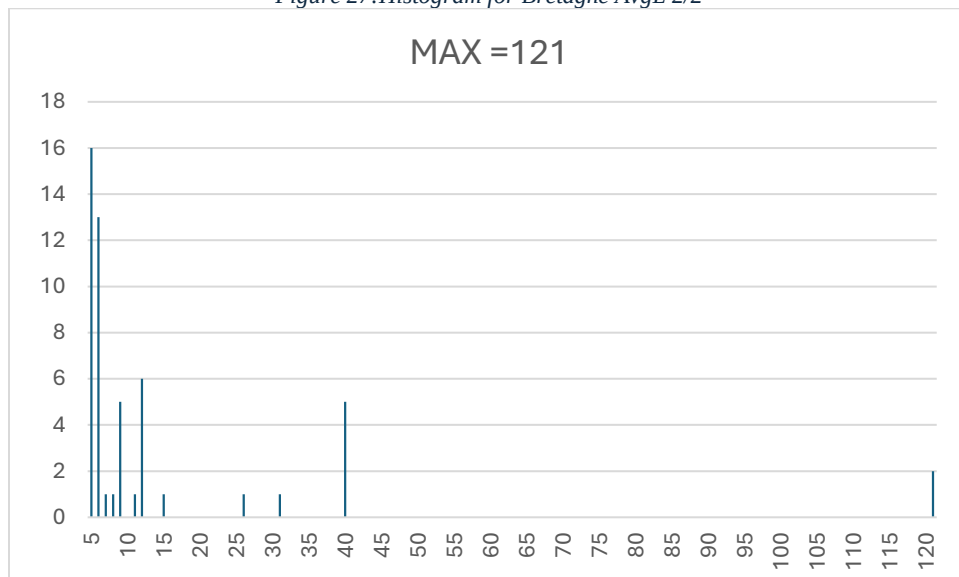


Figure 27:Histogram for Bretagne AvgL 2/2



#### 4.16 St Lucia (.lc): Medium, global

St Lucia has a medium %Multi together with an impressive AvgL reaching 12.4, a top value well above average. This is the result of the presence of global websites with heavy multilingualism. This is the category the most loaded in our 50 TLD sampling, with the presence of Bolivia, Paraguay, Rep. Congo, Côte d'Ivoire and .Lat, all having high AvgL as a consequence of the global push.

Table 24: Indicators for St Lucia (.lc)

|                   | %PAGES         | %SITES         | MultiR | 3.05          |
|-------------------|----------------|----------------|--------|---------------|
| <b>TOTAL</b>      | <b>100.00%</b> | <b>324.73%</b> | %Multi | <b>17.96%</b> |
| <b>English</b>    | <b>21.76%</b>  | <b>70.66%</b>  | AvgL   | <b>12.42</b>  |
| <b>Chinese</b>    | <b>6.69%</b>   | <b>21.74%</b>  | Var    | <b>0.83</b>   |
| <b>Spanish</b>    | <b>5.96%</b>   | <b>19.36%</b>  | GooT   | <b>0.28%</b>  |
| <b>German</b>     | <b>5.62%</b>   | <b>18.26%</b>  | IDN    | <b>0.00%</b>  |
| <b>French</b>     | <b>5.48%</b>   | <b>17.79%</b>  |        |               |
| <b>Italian</b>    | <b>5.19%</b>   | <b>16.87%</b>  |        |               |
| <b>Portuguese</b> | <b>5.05%</b>   | <b>16.39%</b>  |        |               |
| <b>Arabic</b>     | <b>4.42%</b>   | <b>14.34%</b>  |        |               |
| <b>Japanese</b>   | <b>3.58%</b>   | <b>11.64%</b>  |        |               |
| <b>49 Lang.</b>   | <b>36.24%</b>  | <b>117.68%</b> |        |               |

Figure 28: Histogram for St Lucia AvgL 1/2

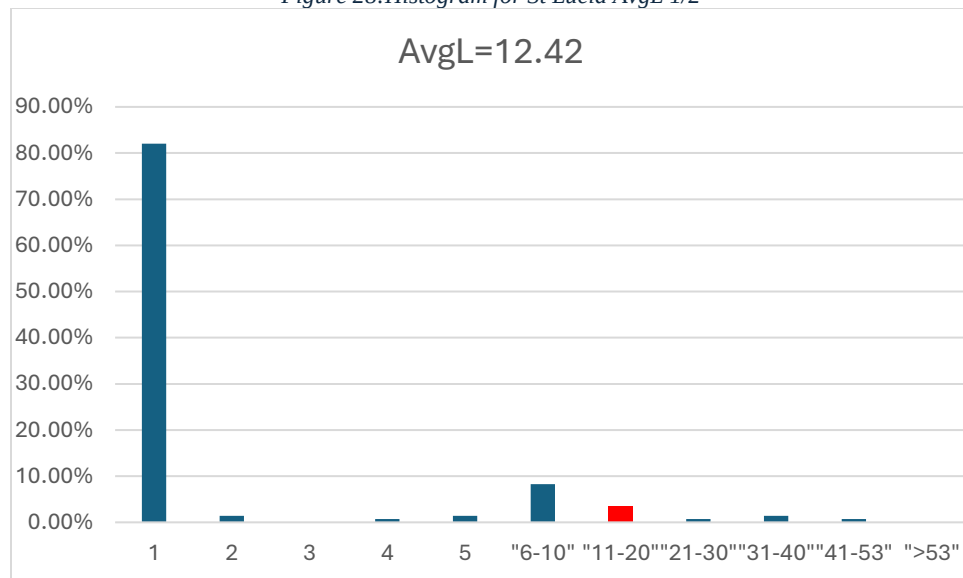
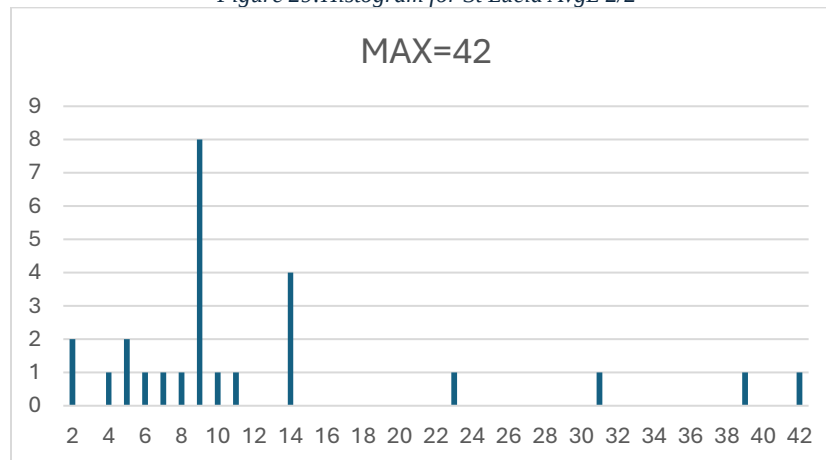


Figure 29: Histogram for St Lucia AvgL 2/2



## 4.17 Vietnam (.vn): Medium

Medium %multi, close to the up boundary, with a low AvgL. The score of English is close to classify the TLD as bilingual. This category is shared by Peru, Romania and Vanuatu, all of them except Vanuatu, being close to bilingualism.

Table 25: Indicators for Viet Nam (.vn)

|                   | %PAGES        | %SITES        | MultiR | 1.38         |
|-------------------|---------------|---------------|--------|--------------|
|                   | 100.00%       | 137.84%       | %Multi | 18.23%       |
| <b>vietnamien</b> | <b>67.92%</b> | <b>92.38%</b> | AvgL   | <b>3.06</b>  |
| <b>anglais</b>    | <b>17.43%</b> | <b>23.89%</b> | GooT   | <b>2.58%</b> |
| <b>japonais</b>   | <b>1.42%</b>  | <b>2.00%</b>  | IDN    | <b>0.23%</b> |
| <b>chinois</b>    | <b>1.35%</b>  | <b>1.90%</b>  |        |              |
| <b>coréen</b>     | <b>0.84%</b>  | <b>1.20%</b>  |        |              |
| <b>français</b>   | <b>0.67%</b>  | <b>0.98%</b>  |        |              |
| <b>83 Lang.</b>   | <b>10.36%</b> | <b>15.50%</b> |        |              |

## 4.18 Asia (.asia): Low, bilingual

Low %Multi, although on the high side, close to world average. Bilingualism Chinese, English with dominance of Chinese. Relatively high AvgL and limited use of IDN.

Table 26: Indicators for .Asia

|                   | %PAGES        | %SITES        | MultiR | 1.53         |
|-------------------|---------------|---------------|--------|--------------|
|                   | 100.00%       | 153.42%       | %Multi | 11.97%       |
| <b>Chinese</b>    | <b>36.74%</b> | <b>54.74%</b> | AvgL   | <b>5.58</b>  |
| <b>English</b>    | <b>21.65%</b> | <b>32.51%</b> | GooT   | <b>0.00%</b> |
| <b>German</b>     | <b>4.40%</b>  | <b>6.82%</b>  | IDN    | <b>0.10%</b> |
| <b>French</b>     | <b>3.25%</b>  | <b>5.09%</b>  |        |              |
| <b>Vietnamese</b> | <b>3.24%</b>  | <b>5.01%</b>  |        |              |
| <b>Thai</b>       | <b>3.03%</b>  | <b>4.62%</b>  |        |              |
| <b>Spanish</b>    | <b>2.75%</b>  | <b>4.35%</b>  |        |              |
| <b>Japanese</b>   | <b>2.25%</b>  | <b>3.55%</b>  |        |              |
| <b>Indonesian</b> | <b>2.16%</b>  | <b>3.37%</b>  |        |              |
| <b>Portuguese</b> | <b>1.96%</b>  | <b>3.12%</b>  |        |              |
| <b>Italian</b>    | <b>1.94%</b>  | <b>3.11%</b>  |        |              |
| <b>80 Lang.</b>   | <b>16.64%</b> | <b>27.13%</b> |        |              |

The following table shows the score of first ranked Asian languages,

Table 27: Asian languages in .asia

|             | %PAGES | %SITES |
|-------------|--------|--------|
| Chinese     | 36.74% | 54.74% |
| Vietnamese  | 3.24%  | 5.01%  |
| Thai        | 3.03%  | 4.62%  |
| Japanese    | 2.25%  | 3.55%  |
| Indonesian  | 2.16%  | 3.37%  |
| Bengali     | 1.37%  | 2.03%  |
| Korean      | 1.24%  | 2.00%  |
| Burmese     | 1.03%  | 1.53%  |
| Turkish     | 0.66%  | 1.10%  |
| Malay       | 0.54%  | 0.85%  |
| Hindi       | 0.39%  | 0.63%  |
| Tagalog     | 0.07%  | 0.11%  |
| Tamil       | 0.07%  | 0.11%  |
| Khmer       | 0.06%  | 0.09%  |
| Kazakh      | 0.06%  | 0.09%  |
| Marathi     | 0.03%  | 0.05%  |
| Telugu      | 0.03%  | 0.05%  |
| Azerbaijani | 0.03%  | 0.05%  |
| Panjabi     | 0.03%  | 0.05%  |
| Lao         | 0.02%  | 0.03%  |

#### 4.19 Africa (.africa): Low, global

The gTLD .Africa gathers some 60 000 domains; it has a low %Multi with a strong predominance of English witnessing a higher presence in anglophone countries compared to francophone countries, as French is not present in the proportion of African speaker's presence. The rather large AvgL is the results of the presence of many global websites , most often by redirection.

Table 28: Indicators for .africa

|                | %PAGES         | %SITES         |                  |               |
|----------------|----------------|----------------|------------------|---------------|
| <b>TOTAL</b>   | <b>100.00%</b> | <b>149.80%</b> | <b>MultiR</b>    | <b>1.50</b>   |
| <b>English</b> | <b>59.74%</b>  | <b>89.50%</b>  | <b>%Multi</b>    | <b>10.70%</b> |
| French         | 8.81%          | 13.21%         | <b>GooT</b>      | <b>0.57%</b>  |
| German         | 3.36%          | 5.03%          | <b>IDN</b>       | <b>0.00%</b>  |
| Spanish        | 2.04%          | 3.05%          | <b>AVGL</b>      | <b>5.66</b>   |
| Portuguese     | 1.82%          | 2.73%          | <b>VAR</b>       | <b>1.14</b>   |
| Vietnamese     | 1.47%          | 2.20%          | <b>MAX</b>       | <b>121</b>    |
| Italian        | 1.31%          | 1.96%          | <b>%MONO</b>     | <b>89.3%</b>  |
| Arabic         | 1.19%          | 1.78%          | <b>%2LINGUAL</b> | <b>5.4%</b>   |
| Dutch          | 1.15%          | 1.72%          | <b>%3LINGUAL</b> | <b>1.6%</b>   |

|                   |               |               |                  |             |
|-------------------|---------------|---------------|------------------|-------------|
| <b>Indonesian</b> | 1.04%         | 1.55%         | <b>%4LINGUAL</b> | <b>0.6%</b> |
| <b>128 Lang</b>   | <b>18.07%</b> | <b>27.07%</b> | <b>%LOCAL</b>    |             |

Figure 30:Histogram for .africa AvgL 1/2

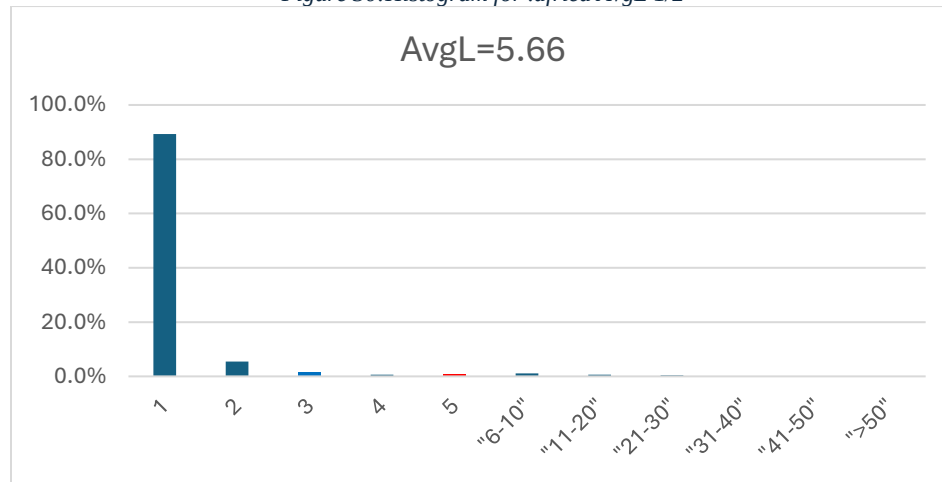
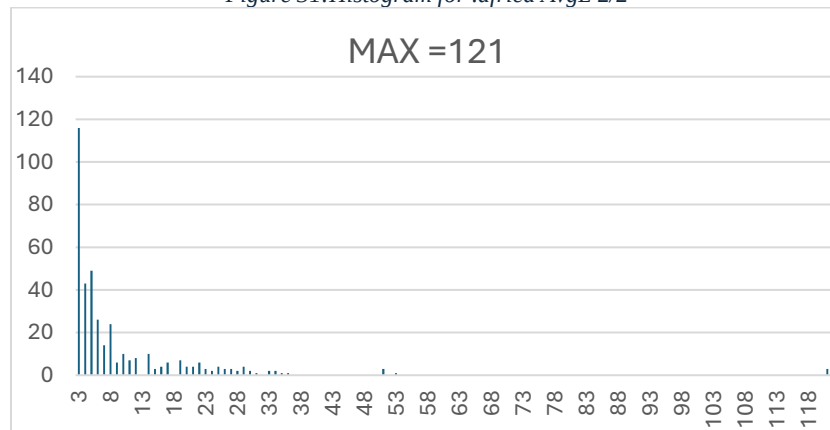


Figure 31:Histogram for .africa AvgL 2/2



The following table shows the percentages of first ranked African languages.

Figure 32: African languages in .africa

|                    | <b>%PAGES</b> | <b>%SITES</b> |
|--------------------|---------------|---------------|
| <b>Arabic</b>      | 1.19%         | <b>1.78%</b>  |
| <b>Swahili</b>     | 0.77%         | <b>1.15%</b>  |
| <b>Afrikaans</b>   | 0.34%         | <b>0.51%</b>  |
| <b>Zulu</b>        | 0.17%         | <b>0.25%</b>  |
| <b>Hausa</b>       | 0.16%         | <b>0.24%</b>  |
| <b>Amharic</b>     | 0.14%         | <b>0.21%</b>  |
| <b>Xhosa</b>       | 0.10%         | <b>0.15%</b>  |
| <b>Yoruba</b>      | 0.10%         | <b>0.15%</b>  |
| <b>Igbo</b>        | 0.09%         | <b>0.14%</b>  |
| <b>Kinyarwanda</b> | 0.09%         | <b>0.14%</b>  |

|                |       |              |
|----------------|-------|--------------|
| Oromo          | 0.06% | <b>0.08%</b> |
| Somali         | 0.06% | <b>0.08%</b> |
| Southern Sotho | 0.05% | <b>0.07%</b> |
| Swati          | 0.05% | <b>0.07%</b> |
| Tswana         | 0.05% | <b>0.07%</b> |
| Lingala        | 0.04% | <b>0.06%</b> |
| Tsonga         | 0.04% | <b>0.06%</b> |
| Venda          | 0.04% | <b>0.06%</b> |
| Ganda          | 0.03% | <b>0.04%</b> |
| South Ndebele  | 0.03% | <b>0.04%</b> |
| Rundi          | 0.03% | <b>0.04%</b> |
| Wolof          | 0.03% | <b>0.04%</b> |
| Malagasy       | 0.02% | <b>0.03%</b> |
| Shona          | 0.02% | <b>0.03%</b> |
| Tigrinya       | 0.02% | <b>0.03%</b> |

## 4.20 India (.in): Very Low, local

Very low %Multi for India, a strong English monolingualism with Hindi in second position but with low percentage.

Table 29: Indicators for India (.in)

|            | %PAGES  | %SITES  | MultiR | 1.28  |
|------------|---------|---------|--------|-------|
|            | 100.00% | 128.45% | %Multi | 6.52% |
| English    | 73.00%  | 91.83%  | AvgL   | 5.39  |
| Hindi      | 3.81%   | 4.83%   | GooT   | 3.20% |
| Vietnamese | 2.04%   | 2.61%   | IDN    | 0.00% |
| Indonesian | 1.53%   | 1.96%   |        |       |
| Spanish    | 1.33%   | 1.77%   |        |       |
| German     | 1.30%   | 1.74%   |        |       |
| French     | 1.19%   | 1.60%   |        |       |
| Portuguese | 1.04%   | 1.39%   |        |       |
| Italian    | 0.91%   | 1.23%   |        |       |
| Tamil      | 0.79%   | 1.05%   |        |       |
| 124 lang   | 13.06%  | 18.45%  |        |       |

The following table shows the percentages of first ranked Indian languages.

Table 30: Indian languages in .in

|           | %PAGES | %SITES | % POP  |
|-----------|--------|--------|--------|
| Hindi     | 3.81%  | 4.83%  | 41.86% |
| Tamil     | 0.79%  | 1.05%  | 5.32%  |
| Marathi   | 0.55%  | 0.71%  | 6.88%  |
| Malayalam | 0.40%  | 0.53%  | 2.48%  |
| Bengali   | 0.36%  | 0.48%  | 7.41%  |
| Telugu    | 0.29%  | 0.40%  | 6.53%  |
| Gujarati  | 0.27%  | 0.37%  | 4.17%  |
| Kannada   | 0.24%  | 0.34%  | 4.07%  |
| Urdu      | 0.14%  | 0.20%  | 4.40%  |
| Panjabi   | 0.13%  | 0.19%  | 2.36%  |
| Oriya     | 0.07%  | 0.09%  | 2.64%  |
| TOTAL     | 7.05%  | 9.18%  | 88.13% |

## 5- CONCLUSIONS

Studies conducted in 2024 using DataProvider.com have revealed that the Web was already a multilingual realm with growth of that property but a huge variance depending of countries, languages and type of websites.

The first studies conducted in 2026, using the new tool MECILDI, developed by OBDILCI, have allowed more precise measurements while confirming the 2024 outcomes. The precision obtained have allowed to understand that the multilingualism of the Web is concentrated in a limited set of languages, mainly major European languages and hardly beneficiate local languages which presence is extremely limited.

This new study reveals that web multilingualism could be either related to a genuine will to provide bilingual, trilingual and even sometimes quadrilingual websites, between the most spoken languages in the place associated to the top-level domain, either to the presence, often by redirection, to global websites having a large number of linguistic versions which somehow artificially boost the average number of languages.

Based on those observations this report propose a typology for apprehending multilingualism in websites series and apply it to some 50 ccTLDs or gTLDs which has been analyzed as regard to their linguistic diversity.

## REFERENCES:

- MECILDI pages including access to all published results: <https://obdilci.org/projects/mecildi/>
- Pimienta D., Field M., 2026, *MECILDI: A Multilingual-Aware Computational Framework for Measuring the Language Distribution of Web Content*, in peer review stage in Plos One. A preliminary version is still accessible in: <https://doi.org/10.13140/RG.2.2.33908.10886>
- OBDILCI, 2025, Web multilingualism studies. <https://www.obdilci.org/projects/other/mlreports/>
- OBDILCI, 2026, First report on the lessons learned by MECILDI : The Desert Web of Indigenous, Vernacular, Local, and Minority Languages, <https://www.obdilci.org/wp-content/uploads/2026/09/WebLocal.pdf>
- OBDILCI 2026, Second report on the lessons learned by MECILDI : Revising the actual figures for languages presence in the Web. <https://www.obdilci.org/wp-content/uploads/2026/09/WebAfterMecildi.pdf>