

Maturity Levels for Cultural Diversity in Wikipedia Language Communities

Marc Miquel, PhD. {marcmiquel@gmail.com}
 U:marcmiquel. Amical Wikimedia / Diversity WG



This is part of [Wikipedia Cultural Diversity Observatory](#), whose goal is to provide data with strategic value and resources to organize and fight for more cultural diversity within each language edition.



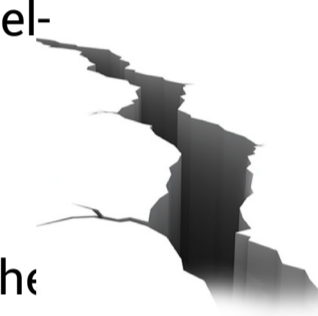
[https://meta.wikimedia.org/wiki/Wikipedia_Cultural_Diversity_Observatory]

Research: we have a problem of cultural diversity representation and sharing.

Taking into account the largest Wikipedia language editions, the content related to the context of each language edition (geography, people, traditions, language, etc.) is in average about a quarter of each Wikipedia (Miquel-Ribé and Laniado 2016).



We have a **problem of representation** considering that the extent of this cultural and geographical context content (CCC) in non-western languages (African and Asian) is on average much smaller (Miquel-Ribé and Laniado, 2019).



Big languages like English or geographically close languages are the ones covering best the smaller languages (Miquel-Ribé and Laniado 2016). About a 60% of the content language gaps is due to cultural and geographical context content (Miquel-Ribé and Laniado 2018). **Culture gap.**

We have a **problem of sharing** considering a Wikipedia language edition cultural diversity as the coverage of all the others' CCC.

How mature is your Wikipedia community when it comes to fostering cultural diversity content?

Being mature at cultural diversity implies that, first, a language edition community represents its cultural context properly (e.g. cities, monuments, leaders, etc.) and, second, it is able to spread it as well as to cover the cultural context content related to the rest of Wikipedia language editions.

A maturity model assesses qualitatively the situation and barriers an organization comes across in order to succeed at developing a particular aspect. In this case, we are going to propose a maturity model with 5 levels for Wikipedia communities to achieve cultural diversity in their language edition' contents. Every Wikipedia language community needs incorporating **Discourse**, **Organization (events and tools)**, **Awareness (metrics and visualizations)** and **Strategy (goals and priorities)**.

Maturity in communities progresses one level at a time. If a community is at a spontaneous level it will not be able to skip ahead to control without moving through organized, gaining the necessary community capacity. Each level requires knowledge, skills, organization and some processes that need to be learn.

Maturity Levels Model

| Level | 01 Unintentional | 02 Spontaneous | 03 Organized | 04 Controlled | 05 Distributed |
|----------------------------------|--|---|---|--|--|
| Situation | Few editors translating general encyclopaedic articles (New York, Mona Lisa, etc). No cultural context representation. | Editors represent their own cultural context and translate articles to cover cultural diversity individually. | Events to represent own context (e.g. Wiki Loves Monuments), spread it (e.g. Catalan Culture Challenge) and cover others' (e.g. Asian Month). | While the use of metrics shows the knowledge gaps but its use is incipient, the community organization is mature and has capacity. | Cultural diversity has dedicated events and is also cross-section. Editors follow the stats on the depth of the gaps and regularly use the tools to bridge them. |
| Incorporated elements | None | Discourse | Discourse Organization | Discourse Organization Awareness | Discourse Organization Awareness Strategy |
| Barriers | Editing barriers i.e. digital divide, sociocultural barriers. | Lack of community building and offline support. | Difficulty to assess the impact and gaps. | Metrics and tools to find gaps are not integrated in the editors workflow. | |
| Tools to reach next level | Lack of editors | Organization | Quantification | Strategic goals | |
| Community Example | African languages | Maltese and Walloon, among others. | Catalan, Spanish, Italian, among others. | CEE languages (e.g. Ukrainian and German) among others. | None yet |

Helping your community reach the next level: you need to work on these aspects.

- **Discourse:** editors recognize the value of representing their cultural context and of translating articles from others' and the need to develop some organization to do it.
- **Organization:** some events are dedicated to represent their own heritage (e.g. Wiki Loves Monuments), to spread it across other languages (e.g. Catalan Culture Challenge), and to cover other languages' cultural context (e.g. Asian Month).
- **Awareness:** the use of metrics to gain understanding of the situation is in incipient. It is useful in order to know the extent of the own cultural context content (% of articles), and the knowledge gaps.
- **Strategy:** cultural diversity is distributed and assumed by all the community members as a value. Some specific goals are set assisted by lists of priority articles (e.g. Top CCC articles list computer-generated by the Cultural Diversity Observatory).

You can read more about the maturity levels at:

[https://meta.wikimedia.org/wiki/Wikipedia_Cultural_Diversity_Observatory/Maturity_Levels]

Wikipedia Cultural Diversity Observatory wants to align the movement (foundation, chapters and communities) to achieve a higher cultural diversity in content.

Top CCC articles: diversity articles every Wikipedia should have.

It is impossible to bridge all the knowledge gaps between languages.

In the Cultural Diversity Observatory we propose every Wikipedia has at least 100 articles about every other language's cultural and geographical content.

This is 28-30 thousand articles to cover a minimum of Wikipedia cultural diversity.

[https://wcd.wmflabs.org/top_ccc_articles]

Top 500 CCC articles list "Men" from Catalan CCC to Standard Chinese

Select the parameters [Download Table \(Excel\)](#)

List: Men Language origin: Catalan (ca) Country origin: Select a country (optional) Language target: Standard Chinese (zh)

| Nº | Catalan Title | Edits | Editors | Pageviews | Bytes | References | Wikidata Properties | Interwiki Links | Inlinks from CCC | Creation Date | Other Languages | Standard Chinese Title |
|----|----------------------------|-------|---------|-----------|--------|------------|---------------------|-----------------|------------------|---------------|--------------------|------------------------|
| 1 | Ramon Llull | 1569 | 449 | 536 | 110.2k | 98 | 82 | 56 | 136 | 2003-11-10 | zh, zh_yue | 拉蒙·柳利 |
| 2 | Antoni Gaudí i Cornet | 1214 | 371 | 144 | 90.2k | 57 | 81 | 115 | 131 | 2003-08-10 | zh, zh_yue | 安东尼·高迪 |
| 3 | Jacint Verdaguer i Santaló | 1081 | 262 | 82 | 93.7k | 41 | 56 | 33 | 95 | 2003-08-11 | zh, zh_yue | 雅辛特·瓦达古埃 |
| 4 | Rafael Casanova i Comes | 1062 | 146 | 2371 | 268.2k | 52 | 27 | 24 | 8 | 2005-09-16 | zh | |
| 5 | Joan Miró i Ferrà | 920 | 227 | 55 | 98.4k | 64 | 114 | 94 | 91 | 2006-01-29 | zh, zh_yue, zh_yue | 胡安·米罗 |
| 6 | Salvador Dalí i Domènech | 880 | 366 | 118 | 49.3k | 52 | 145 | 195 | 46 | 2003-06-10 | zh, zh_yue | 萨尔瓦多·达利 |
| 7 | Artur Mas i Gavarró | 841 | 265 | 873 | 54.0k | 95 | 41 | 50 | 98 | 2004-05-20 | zh, zh_yue | 阿图尔·马斯 |