



**University of
Zurich**^{UZH}

**Zurich Open Repository and
Archive**

University of Zurich
Main Library
Strickhofstrasse 39
CH-8057 Zurich
www.zora.uzh.ch

Year: 2015

Cultural anthropology through the lens of Wikipedia - A comparison of historical leadership networks in the English, Chinese, And Japanese Wikipedia

Gloor, Peter ; de Boer, Patrick ; Lo, Wei ; Wagner, Stefan ; Nemoto, Keichii

Abstract: In this paper we study the differences in historical worldviews between Western and Eastern cultures, represented through the English, Chinese and Japanese Wikipedia. In particular, we analyze the historical networks of the world's leaders since the beginning of written history, comparing them in the three different language versions of Wikipedia.

Posted at the Zurich Open Repository and Archive, University of Zurich
ZORA URL: <https://doi.org/10.5167/uzh-102327>
Conference or Workshop Item
Accepted Version

Originally published at:

Gloor, Peter; de Boer, Patrick; Lo, Wei; Wagner, Stefan; Nemoto, Keichii (2015). Cultural anthropology through the lens of Wikipedia - A comparison of historical leadership networks in the English, Chinese, And Japanese Wikipedia. In: COINS15, Keio University, Japan, 12 March 2015 - 14 March 2015.

CULTURAL ANTHROPOLOGY THROUGH THE LENS OF WIKIPEDIA – A COMPARISON OF HISTORICAL LEADERSHIP NETWORKS IN THE ENGLISH, CHINESE, AND JAPANESE WIKIPEDIA

Peter A. Gloor¹, Patrick M. de Boer^{1,5}, Wei Lo², Stefan Wagner³, Keiichi Nemoto⁴
 MIT¹, Zhejiang University², FHNW³, Keio University⁴, University of Zurich⁵
 Cambridge USA¹, Hangzhou China², Brugg Switzerland³, Tokyo Japan⁴, Zurich, Switzerland⁵
pgloor@mit.edu, pdeboer@mit.edu, spencer_w_lo@zju.edu.cn, stefan.wagner@fhnw.ch, nemoto76@gmail.com

ABSTRACT

In this paper we study the differences in historical worldviews between Western and Eastern cultures, represented through the English, Chinese and Japanese Wikipedia. In particular, we analyze the historical networks of the world’s leaders since the beginning of written history, comparing them in the three different language versions of Wikipedia.

INTRODUCTION

In this project we are using Wikipedia’s English, Chinese and Japanese localization as a window into the “soul” of different cultures, replacing anthropological fieldwork with statistical analysis of the treatment given by native speakers of a culture to different subjects in Wikipedia.

Wikipedia is structured into categories, one of the most popular ones are people pages, talking about the most important people of all ages. Wikipedians have put together “notability criteria” that clearly define if a person deserves inclusion into Wikipedia or not. In this paper we look at the most prominent people pages over all times in the English, Chinese, and Japanese Wikipedia, leading us towards a comparison between the Western and Eastern worldview.

PEOPLE NETWORK CONSTRUCTION

Our goal was to create a social network of all leaders that every lived, since the beginning of time. As a proxy, we only take people that made it into Wikipedia, fulfilling Wikipedia’s notability criteria. As a second requirement, a link between two people can only exist if both of them were living at the same time. For each language edition of Wikipedia, we start with all pages tagged as “people pages”. The English Wikipedia contains 800,000 pages in this category, for example. In the next step all people pages are dated, by extracting the dates of birth and of death of each individual. Moreover, the links originating and pointing to their Wikipedia page are gathered. Using this information, for each year through history, from 3000 BC to 1950 CE, a link

network is calculated, as shown in figure 1. From all the links originating and pointing back to a particular people page, only the links to and from people living at the same time as the person discussed on that page are included.

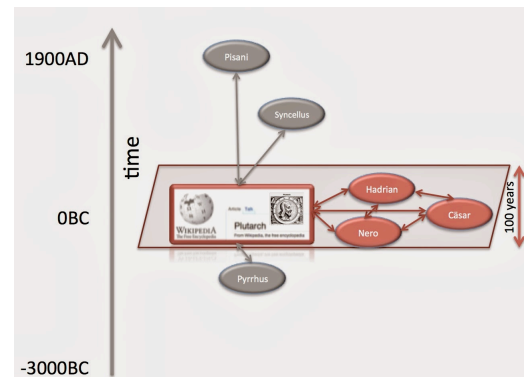


Figure 1: Link construction among contemporary people pages

For instance, in the graph shown in figure 1 above, from all the links to the page about [Plutarch](#), only the links from and to Hadrian, Caesar, and Nero are kept, while the links to Pyrrhus, who died well before Plutarch was born, and the pages to medieval historian Syncellus and modern historian Pisanì are ignored as well. Repeating this process leads to 4900 unique networks for the English Wikipedia (less for the Chinese and Japanese Wikipedia, as their history does not go back as far. For each of these networks, the most central people are determined using the PageRank algorithm. To get a second selection criterion among the influencers, their indegree, i.e. other people pages pointing back to them, is taken.

Figure 2 illustrates the Wikihistory application running in a Web browser that we developed based on the network extracted above. It shows the network of most influential people in the English Wikipedia in the year 0 – note that Jesus is not yet part of the network, he will show up in year 1.

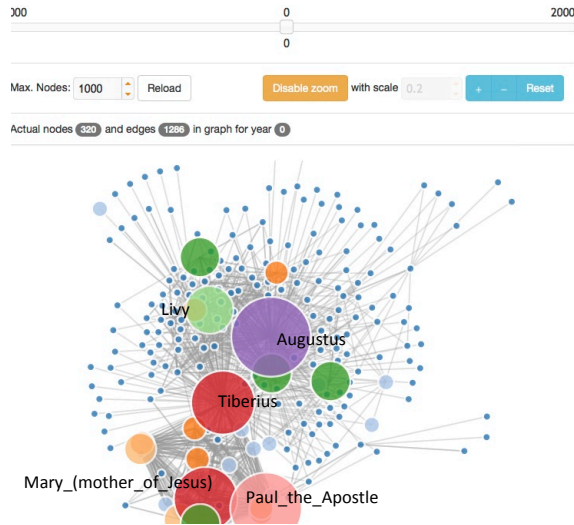


Figure 2: Sample screen shot of Wikihistory at 0 BC

IT'S ALL ABOUT EMPERORS AND WARS

Who are the most important people of all times? The answer to this question is quite different in the US and the UK than it is in China and Japan. Looking at the top ten and top 50 people lists also confirms that most English language Wikipedia editors come from the US and the UK, while Chinese language editors come from Taiwan, Hong Kong, and mainland China.

English	Chinese	Japanese
George_W._Bush	Mao_Zedong	Ikuhiko_Hata
William_Shakespeare	Yuan_Shikai	Tokugawa_Ieyasu
Sidney_Lee	Jay_Chou	Toyotomi_Hideyoshi
Jesus	Oda_Nobunaga	Adolf_Hitler
Charles_II_of_England	Tokugawa_Ieyasu	Oda_Nobunaga
Aristotle	Emperor_Gaorong_of_Tang	Hirohito
Napoleon	Cao_Cao	Tokugawa_Hidetada
Muhammad	Kangxi_Emperor	Tokugawa_Iemitsu
Charlemagne	Emperor_Huizong_of_Song	Chiang_Kai-shek
Plutarch	Yongle_Emperor	Tokugawa_Ienari

Table 1: Top ten most important people in 3 Wikipedias (most important at the top) Red denotes politicians, black religious leaders, blue scientists and artists

As tables 1 and 2 illustrate, in China and Japan only famous warriors and politicians have a chance to make it into the top ten and top fifty – the East seems far less religious than the West – while the English Wikipedia is more balanced with half of the top ten as well as the top fifty of all times being religious leaders or artists or scientists. Historians play a special role. Both Sidney Lee, a relatively minor Victorian professor of English and history, who wrote 800 biographies, and Ikuhiko Hata, a 19 century Japanese military biographer owe their prominent position to their prolific biography writing, as they get many backlinks from the references on the pages of contemporary politicians they wrote about.

	English	Chinese	Japanese
Politicians	26	46	47
Religious Leaders	11	1	0
Artists/Scientists	13	3	3
Cultural Ingroup	10	48	31

Table 2: Distribution of different people categories in 3 Wikipedias among the top 50 people of all times

The second striking difference comes from outgroup leaders included into the top 50. While the English Wikipedia includes 80% non-English leaders among the top 50, just two non-Chinese made it into the top 50 of the Chinese Wikipedia: Napoleon III and Tokugawa Ieyasu. The Japanese Wikipedia is slightly more balanced, with almost 40 percent non-Japanese leaders, half of them Chinese Emperors, the others people like Adolf Hitler, Plato, Cicero, and Augustus.

CONCLUSIONS

The Internet enables researchers to more easily compile rankings of the most important world leaders of all times (Murray 2003, Hidalgo 2014). Our work is unique in that we extract language-specific rankings that allow us to compare the worldview for dozens of different cultures. Probing the historical perspective of many different language-specific Wikipedias gives an X-ray view deep into the historical foundations of cultural understanding of different countries.

REFERENCES

Hidalgo, C. (2014) Pantheon: Mapping Historical Cultural Production. <http://pantheon.media.mit.edu>

Murray, C. (2003). Human accomplishment: The pursuit of excellence in the arts and sciences, 800 BC to 1950. HarperCollins.