The Reach and Richness of Wikinomics: 
Is the free web-based encyclopedia Wikipedia only for rich countries?

Morten Rask, Aarhus School of Business, University of Aarhus, Denmark, mr@morten-rask.dk

In this paper, a model of the patterns of correlation in Wikipedia, reach and richness, lays the foundation for studying whether or not the free web-based encyclopedia Wikipedia is only for developed countries. Based upon data from 12 different Wikipedia language editions, we find that the central structural effect is on the level of human development in the current country. In other words, Wikipedia is in general, more for rich countries than for less developed countries. It is suggested that policy makers make investments in increasing the general level of literacy, education, and standard of living in their country. The main managerial implication for businesses, that will expand their social network applications to other countries, is to use the model of the patterns of correlation in Wikipedia, reach and richness, as a market screening and monitoring model.

Within the last decade, we have been approached with many promises in regards to information and communication technologies’ abilities to lift developing countries out of poverty. We have learned from powerful businesses and political leaders that the Digital Age will leave four qualities that will result in decentralizing, globalizing, harmonizing, and empowering, (Negroponte 1995) because of the “dissolving” of distance and the “flattening” of the world (Cairncross 1997). In other words, the “playing field” is being leveled for companies and especially individuals from all parts of the world, resulting in more frequent and intensive collaboration and competition, benefiting people and companies in poor, emerging, and transition countries (Friedman 2006). This has given rise for “A new art and science of poor, emerging, and transition countries (Friedman 2006). In other words, for poor, emerging, and transition countries, these thoughts sound almost too good to be true. Instead of being at the outskirts of the global economy, the development of the free or low-cost internet-based content and applications makes it possible for underdeveloped countries to compete and collaborate on the same level as the developed countries.

In order to explore these golden promises, the paper will be organized as follows: After a short introduction to Wikipedia and the related research questions, the following section will provide an overview of the empirical foundation. Before the discussion of the findings, the reach and richness of the terminology will be operationalized. The conclusion and implications for policy makers and further research will end the paper.

INTRODUCTION TO THE WIKIPEDIA PHENOMENA

Wikipedia will be the case used in order to challenge the golden promises discussed above. If Microsoft is an icon of the global development of the IT-based society and if Amazon and eBay are icons of global e-commerce, Wikipedia is rapidly becoming a global icon of the new Web 2.0 development. Wikipedia is often used in the description of Web 2.0 compared to Britannica Online which is exemplifying the Web 1.0 application (Baumann 2006; O’Reilly 2005; Tapscott & Williams 2006). Britannica Online content is written, edited, and published by a selected group of people, in much the same way physical books are published. At Wikipedia, anyone can add to the collective pool of knowledge, and anyone can access it for free. Wikipedia is an encyclopedia created by user-added content that is made to change on a continuous basis. Before Web 2.0, the exchange of lexicographical information was mostly one-way. This created much debate about the accuracy of the entries in Wikipedia (Economist 2006a; Economist 2006b; Economist 2006c; Economist 2006d; Giles 2005; Korfiatis, Poulos, & Bokos 2006) where the general dispute has been about whether or not to believe that there is one source of truth coming from the expert or if there are multiple sources of truth coming from the common “herd”. This is also explained briefly on every Wikipedia (2006) article: “As with any community-built reference, there is a possibility for error in Wikipedia’s content — please check your facts against multiple sources and read our disclaimers for more information.” In other words, Wikipedia content has already been a case for discussion, in terms of the quality of the Web 2.0 ideas, because of the openness, peering (peer
production), sharing, and acting globally.

What has not been questioned is: who actually participates? This is the question of the digital divide; the gap between those with regular, effective access to Wikipedia and those without. Who is able to use Wikipedia for their own benefit and who is not? Instead, the enthusiastic messengers of the greatness of Wikipedia simply state that Wikipedia makes “anyone with an Internet connection a potential resource” (Baumann 2006), to participate and contribute because “all one needs is a computer, a network connection, and a bright spark of initiative and creativity to join in the economy” (Tapscott & Williams 2006). So, Wikipedia's promise of the liberation of human knowledge is for “everybody who has access to the internet” (Economist 2006d). However, the Internet only covers approximately 15 percent of the world’s population (International Telecommunication Union 2006) and the 2005 UNCTAD report shows that, while in some developing regions the number of Internet users has grown substantially, overall, the gap between developed and developing countries remains wide. And, the quality of connections is just as important as their numbers. While some countries have seen spectacular growth in broadband access, there are still large variations worldwide (UNCTAD 2005). The broadband access is seen as especially important because the full-scale participatory media, as Wikipedia presumes, is “always-on” and there is always broadband access to it, as opposed to the availability of internet, which is now an “old” phenomenon. (Economist 2006a).

In other words, there is a need for research that investigates the reach and richness of Wikipedia with a focus on the contributors of content, in order to understand who is open to work together in the sharing of knowledge? Is Wikipedia only for the rich countries?

Hopefully, the study will contribute to a greater understanding of the relationship between reach and richness and help to comprehend how developing countries could take further advantage of ICT/Internet interactions and specifically participate in wikinomics, in order to promote economic growth and the quality of life (as promoted by UNCTAD (2001; 2002; 2003; 2004; 2005) for some years).

**DATA**

With the aim of understanding the possibilities for poor, emerging, and transition countries to gain access to shared knowledge on the same level as developed countries, 12 countries have been included in the sample out of the 182 active Wikipedia editions/languages (Wikipedia contributors 2007c).

We have chosen to select among the 12 most rapidly developing economies: Brazil, China, The Czech Republic, Hungary, India, Indonesia, Malaysia, Mexico, Poland, Russia, Thailand and Turkey (Aguier et al. 2006). However, some of the countries have not been selected because the language in the country does not relate to the current country alone. This is the case with Brazil, China, Mexico and India.

In order to have an opportunity of comparison with highly developed countries; Japan, Germany and Denmark are included in the sample because these countries have their own more or less exclusive official language. Based on the sample, we are able to gain knowledge about how developing countries could take further advantage of ICT/Internet interactions and participate in Wikinomics, in order to promote economic growth and the quality of life. TABLE 1, showed below, gives an overview of the 12 Wikipedia editions/languages included in the study and the related analytical variables.
TABLE 1: THE 12 WIKIPEDIA EDITIONS/LANGUAGES INCLUDED IN THE STUDY

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Czech Republic</td>
<td>Czech</td>
<td>2002-11</td>
<td>12</td>
<td>0.885</td>
<td>49.97</td>
<td>0.7</td>
<td>1554</td>
<td>14</td>
</tr>
<tr>
<td>Denmark</td>
<td>Danish</td>
<td>2002-02</td>
<td>5.3</td>
<td>0.943</td>
<td>52.55</td>
<td>18.7</td>
<td>1349</td>
<td>15.2</td>
</tr>
<tr>
<td>Germany</td>
<td>German</td>
<td>2001-05</td>
<td>100.1</td>
<td>0.932</td>
<td>45.35</td>
<td>8.4</td>
<td>38959</td>
<td>33.5</td>
</tr>
<tr>
<td>Hungary</td>
<td>Hungarian</td>
<td>2003-07</td>
<td>14.5</td>
<td>0.869</td>
<td>29.71</td>
<td>3.7</td>
<td>1247</td>
<td>16.8</td>
</tr>
<tr>
<td>Indonesia</td>
<td>Indonesian</td>
<td>2003-05</td>
<td>163</td>
<td>0.711</td>
<td>7.18</td>
<td>646</td>
<td>26.3</td>
<td></td>
</tr>
<tr>
<td>Japan</td>
<td>Japanese</td>
<td>2002-09</td>
<td>125</td>
<td>0.949</td>
<td>50.2</td>
<td>14.6</td>
<td>13787</td>
<td>26.3</td>
</tr>
<tr>
<td>Malaysia</td>
<td>Malay</td>
<td>2002-10</td>
<td>23.6</td>
<td>0.805</td>
<td>36.96</td>
<td>1</td>
<td>213</td>
<td>8.7</td>
</tr>
<tr>
<td>Poland</td>
<td>Polish</td>
<td>2001-09</td>
<td>44</td>
<td>0.862</td>
<td>25.95</td>
<td>2.1</td>
<td>7162</td>
<td>13.7</td>
</tr>
<tr>
<td>Russia</td>
<td>Russian</td>
<td>2002-11</td>
<td>167</td>
<td>0.797</td>
<td>15.19</td>
<td>4262</td>
<td>15.2</td>
<td></td>
</tr>
<tr>
<td>Thailand</td>
<td>Thai</td>
<td>2003-03</td>
<td>46.1</td>
<td>0.784</td>
<td>11.03</td>
<td>556</td>
<td>15</td>
<td></td>
</tr>
<tr>
<td>Turkey</td>
<td>Turkish</td>
<td>2002-12</td>
<td>61</td>
<td>0.757</td>
<td>21.86</td>
<td>0.7</td>
<td>1749</td>
<td>14</td>
</tr>
</tbody>
</table>

REACH OF AND RICHNESS IN THE WIKIPEDIA NETWORK

In this paper, Wikipedia is understood as being a global network of the new media, which is at the root of cultural expression and public opinion in the Information Age. Normally, language is understood as the most important platform for cultural expression and public opinion (Hollensen 2004; Usunier 2005), so in this paper, we recognize that Wikipedia, in the start of 2007, consisted of 182 networks because Wikipedia is available in 182 active editions/languages (Wikipedia contributors 2007c).

In each language, the Wikipedia network consists of two interconnected layers of network, namely the articles’ network and the contributors’ network (Korfiatis et al. 2006) where the articles represent the contribution that is created and edited by the contributors. Because reach of the network is defined as the number of people exchanging information, (Evans & Wurster 1997) we can define the reach of the Wikipedia network as the number of contributors. The two layers of network are interconnected in the collaborative editing activity, which covers Evans’ & Wurster’s definition of richness as the three information aspects bandwidth, customization and interactivity. When editing alone is an interactive customization activity, we are left with the bandwidth aspect. We can then define the richness in the Wikipedia network as the amount of edits per article. When measuring the relative differences between the different Wikipedia languages, data generated on a monthly basis by Wikimedia (Zachte 2006) can be used, namely edits per article, as a proxy for the richness and the number of contributors as a proxy, for the reach of the network, when contributors are Wikipedians who edited at least 10 times since they arrived and edits per article are the mean number of revisions per article.

In order to recognize the limitations of reach and richness of the Wikipedia networks, the number of speakers of the languages under investigation is a delimiter in itself. Therefore, the Wikipedia penetration rate has been calculated as the contributors per one million speakers. However, several additional structural effects are useful. Firstly, it is recognized that the characteristics of the external environment affects how value is created in e-business (Doern & Fey 2006). Because editing at Wikipedia does not involve transactions, we have only included some of the characteristics suggested by Doern & Fey (2006). We use the Human Development Index (HDI), which “is a comparative measure of life expectancy, literacy, education, and standard of living for countries worldwide... It is used to determine and indicate whether a country is a developed, developing, or underdeveloped country and also to measure the impact of economic policies on quality of life... and has been used since 1993 by the United Nations Development Programme in its annual Human Development Report” (Wikipedia contributors 2007a). We use this index to distinguish between developed, developing, and underdeveloped countries and we use it as a proxy for the level of human development in the current country. Other structural effects are also mentioned by Doern & Fey (2006) as the availability of internet connections and the country’s telecommunication infrastructure. Secondly, structural effect relates therefore to internet penetration because of the clear reason that the contributor needs to have access to the internet in order to edit articles (Baumann 2006; Economist 2006d; Tapscott & Williams 2006). The fourth structural effect is the broadband penetration used as a proxy for the telecommunication infrastructure. The
broadband access is important while full-scale participatory media, as Wikipedia presumes, is not only the availability of the internet as being widespread, “always-on”, and that there is broadband access to it (Economist 2006a).

In the Wikipedia contributor network, the network effect is the characteristic that causes Wikipedia to have a value to a potential contributor which depends on the number of other contributors who edit articles. The network effect is when the value of a resource depends on the number of users of the resource (Shapiro & Varian 1998). Therefore, will the users become an evangelist of the service (Kelly 1997; Kelly 1998)? In Wikipedia, the network effects are simple: More contributors means more interaction, which equals more articles and this will increase the editing activities that will attract more contributors. In other words, the reach, as number of contributors, and the richness, as edits per articles, are expressions of the network effects. Additionally, the interaction over time has also an expression of network effects. In order to understand the opportunities of reach and richness of the Wikipedia networks, the days of operation have been counted from the creation date of the current Wikipedia language edition to the date of March 1, 2007.

**DISCUSSION OF FINDINGS**

In order to understand the linear relationships among the different dimensions, a simple correlation analysis has made the basis for the analysis of the patterns of correlation in the Wikipedia, reach and richness, which is illustrated in figure 1.

![Figure 1: Patterns of Correlation in Wikipedia Reach and Richness](image)

The main finding in this study is that the central structural effect is the level of human development in the current country, which means that Wikipedia is, in general, more for the rich than for the poorer countries. However, much more can be learned about reach and richness of the Wikipedia network.

First of all, we find a linear relationship between reach and richness so we can confirm that more contributors mean more interaction in terms of more articles and therefore, more editing that will attract more contributors. In regards to network effects, we also find a very strong relationship between the history and reach, which seems natural because the older the Wikipedia network is, the more contributors will have to be (self) recruited.

Besides the linear relationship with reach, the richness of the Wikipedia network is only a subject for correlation with the level of human development. The structural effect on richness can be argued as the higher the level of literacy, education, and standard of living in the country of the contributor, the better equipped the contributor is for editing. Vice versa, the more the articles are edited, the better affect these articles may have on enhancing the level of literacy, education, and standard of living in the country.

This is also true for the understanding of reach of the Wikipedia network. The structural effect on reach can similarly be argued as the higher the level of literacy, education, and standard of living in the country of the contributor, the better the chance for attracting more capable contributors and vice versa, the more contributors enhance their skills and knowledge by editing, the greater the chance is that these contributors will collectively enhance the level of literacy, education, and standard of living in the country.

An important note is that there seems not to be a correlation between the history network effect and the internet penetration structural effect. In other words, there is no reason to believe that the level of internet usage in a current country initialises an early start of a Wikipedia network in a language related to the current country.

Based on this analysis of the correlation patterns of the Wikipedia, reach and richness, there are reasons to
believe that developing countries could take further advantage of ICT/Internet interactions, especially by participating in wikinomics, in order to promote economic growth and better quality of life. However, the level of human development is the basis. This means that countries with a higher level of human development have a competitive advantage over countries with a lower level. So, the development of the free or low-cost internet-based content and applications does not necessarily make it possible for underdeveloped countries to compete and collaborate on the same level as the highly developed countries.

So what can be done? Figure 1 illustrates that investing in technological infrastructure is not sufficient, if the policy makers in a country want to facilitate participation in the Wikipedia network. Although there is a linear relation between the level of internet penetration and reach of the Wikipedia network, there is a stronger linear relationship between the level of human development and internet penetration. Also, there is a linear relationship between the level of human development broad band and penetration. All things considered, the overall picture is that the level of human development is crucial for participating in sharing knowledge at a global scale, in order to promote economic growth and improve quality of life. Access to low-cost internet-based content and applications will not be sufficient enough. So, the most important investment which policy makers can pursue is increasing literacy, education, and the standard of living in their country.

CONCLUSION AND IMPLICATIONS

Based upon the model of the patterns of correlation in Wikipedia, reach and richness, we can draw some conclusions and implications from this study. This study aimed to explore whether or not the free web-based encyclopedia Wikipedia was only for rich countries. We found that countries with a high level of human development, which normally are rich, highly developed countries, have a greater chance to participate, than countries with a lower level of human development, which are the poor, emerging, and transition countries. However, we also found that there are reasons to believe that participation in Wikipedia, as the illustrative case of wikinomics demonstrates, can enhance the level of human development in a current country. In answering this study’s research question regarding whether or not Wikipedia is only for rich countries, it was found that there is reason to believe that poorer countries can benefit from the Wikipedia network. But, it will be the richer countries that will gain and benefit the most because of their more frequent and intensive collaboration and competition in wikinomics.

These findings contradict the golden promises (Cairncross 1997; Friedman 2006; Negroponte 1995) discussed in the beginning of the paper. In contrast to the enthusiastic messengers’ statements about the greatness of Wikipedia, and that an Internet connection is the only precondition for participation, (Baumann 2006; Economist 2006d; Tapscott & Williams 2006) we can conclude that sufficient technological infrastructure is not sufficient enough alone. The high quality broadband connection especially, does not have a direct relationship with reach and richness.

These findings relate to the question of digital divide and this study is very much inline with earlier research which has emphasized the importance of the external environment in e-business (Doern & Fey 2006) and that e-policy making (Barzilai-Nahon 2006) is not only a question of the country’s level of technological infrastructure but that socioeconomic factors also need to be considered.

The implications for the policy makers are relatively straightforward. Besides investment in a sufficient technological infrastructure, in order to facilitate access to low-cost internet-based content and applications such as Wikipedia, and other social network applications, the most important investment is one which focuses on increasing the general level of literacy, education, and standard of living in their country.

The managerial implications for businesses that will expand their social network applications to other countries, is that the model of patterns of correlation in Wikipedia, reach and richness, can be useful as a market screening and monitoring model. From this point of view, the level of human development in the current country will be the most important market ranking factor because the higher the level of human development, the higher the chance of gaining high reach and richness of the social network application in question. The technological infrastructure can also be seen as vital, but will not be a complete market screening factor. In other words, in a global world of wikinomics, a traditional market screening model based upon political, economical, socio-cultural and technological factors can still do the job.

In this study, we have joined language and country together as the same phenomena covering the same group of people. This is one limitation to the study. Wikipedia does not categorize the different language editions by countries or nation-states. There is not necessarily a correlation between nation-states and languages. In fact, Ohmae (2005) argues that “it is a misfortune that language has become tied so intimately to the nation-state”. However, available structural variables are categorized by countries. In order to overcome this barrier of comparison, based upon structural effect, we have chosen countries with an exclusive language. It was also a limitation that many of the most developed Wikipedia networks are based upon languages that are not exclusive to one country and we were not able to analyze these. It is particularly
paradoxical that the English version of Wikipedia, being the far most comprehensive edition, consists of almost half the total Wikipedia network. English is “the lingua franca of the global economy and the de facto standard in cyberspace” (Ohmae 2005). However, “it does not mean that the appeal of operating globally removes the obligation to localize” (Ohmae 1989). Additionally, it can also be argued that “the wider the reach, the greater the number of regional and national preferences will encounter... The differences that persist throughout the world despite its globalization affirm an ancient dictum of economies – that things are driven by what happens at the margin, not at the core” (Levitt 1983). Altogether, there is a need for further research which investigates the relationship between the language and country of origin, in social networks like Wikipedia and similar Web 2.0 applications.

There is no doubt that languages are an important identity in social networks when social networks are understood as global networks of the new media. The language is understood as the most important platform for cultural expression and public opinion (Hollensen 2004; Usunier 2005). Therefore, an important avenue of research, which has not been taken up in this paper, is to investigate the relationship between the different Wikipedia languages and the corresponding cultural identities as low- and high-context cultures (Hall 1968) and/or work related national cultural differences (Hofstede 1984; Hofstede 1994; Hofstede 2006). It could be hypothesized that contributors from low-context as well as small power distance and individualistic cultures would be more suited for expressing their opinions explicitly in the Wikipedia network, where people from high-context as well as from high power distance and collectivistic cultures would avoid these kinds of explicit expressions.

Furthermore, other social networks could be investigated in order to strengthen the understanding of the suggested model of patterns of correlation in Wikipedia reach and richness. In this way, we could determine if the findings in this study are exclusive to Wikipedia. Other interesting cases that are closely connected to Wikipedia, in terms of professional knowledge sharing, are the professional career network, like LinkedIn (2007) and the research dissemination network, like the Social Science Research Network (SSRN 2007). More consumer oriented social networks, like MySpace (2007), and YouTube (YouTube 2007) could also be interesting cases to investigate.

Finally, we have shown that there are reasons to believe that developing countries could take further advantage of ICT/Internet interactions by participating in the Wikipedia knowledge network, in order to promote economic growth and quality of life. But, we have also shown that the level of human development is the basis. Internet penetration is not the only complete and sufficient variable as Negroponte once said: “that is the Internet that will and has become the greatest agent of change” (Negroponte 1995). With that said, let us conclude by looking towards the future with an open question: Does Wikipedia build bridges and forge alliances to the benefit of developing nations? (paragraph).

REFERENCES


