Proposal for:

The accuracy and quality of Wikipedia entries: across subjects and languages

A preliminary study

Dr Naomi Norman, Epic

Dr Chris Davies, Oxford University

12 July 2011

GOLD AWARD
2010 E-LEARNING AWARDS
'Excellence in the production of learning content - Private sector'

2010 NATIONAL TRAINING AWARDS
Regional winner for 'Partnership and collaboration'

2010 BRANDON HALL AWARDS
Bronze for 'Best Custom Content'

2010 INTERNATIONAL E-LEARNING ASSOCIATION AWARDS
Winner 'Business/Professional Mobile Learning'

2010 INSTITUTE OF IT TRAINING AWARDS
Silver Award for 'External Project of the Year'

2010 INTERNATIONAL VISUAL COMMUNICATIONS ASSOCIATION
Highly commended for 'E-learning Project of the Year'

BM TRADA
CERTIFICATION
ISO 14001:2004
UKAS
ENVIRONMENTAL
MANAGEMENT
012

Accredited to ISO 14001:2004
Environmental Management Standard

Epic
52 Old Steine
Brighton BN1 1NH
United Kingdom

+44(0)1273 728686
contactenquiries@epic.co.uk
www.epic.co.uk

Oxford University
Department of Education
15 Norham Gardens
Oxford OX2 6PY
United Kingdom

+44 (0) 1865 274024
general.enquiries@education.ox.ac.uk
www.education.ox.ac.uk

Contents

1.   Background

In December 2005, the scientific journal, *Nature*, reported on a study they had undertaken to compare the accuracy of science entries on Wikipedia with those on the online version of Encyclopaedia Britannica[1]. The key difference between these two encyclopaedias is that Wikipedia relies on voluntary contributors, regardless of proven mastery or qualifications, while Encyclopaedia Britannica uses selected paid expert advisors and editors.

Nature invited independent academic scientists to peer review entries (in the English language) for their particular areas of science expertise, from both Wikipedia and Encyclopaedia Britannica. Without sight of the source, each scientist was asked to identify any inaccuracies.

Forty-two reviews were submitted to *Nature*, revealing, on average, four inaccuracies per Wikipedia article, in contrast to three per Encyclopaedia Britannica article. The general response was one of surprise, with levels of accuracy in Wikipedia being better than expected. However, for Encyclopaedia Britannica, the oldest continuously published reference work in the English language, the results were worse than expected. (Lih, 2009: 214)[2] While Jimmy Wales, the co-founder and promoter of Wikipedia, expressed delight, he also added: 'Our goal is to get to Britannica quality or better'. (Giles, 2005: 900).

Encyclopaedia Britannica refuted the study's findings, stating 'Almost everything about the journal's investigation, from the criteria for identifying inaccuracies to the discrepancy between the article text and its headline, was wrong and misleading' (Encyclopaedia Britannica Inc., March 2006: 1), and called for Nature to issue a 'full and public retraction'. (Encyclopaedia Britannica Inc., March 2006: 2)[3]. Nature responded by rejecting Encyclopaedia Britannica's criticisms, affirming its confidence in the study, and refusing to retract.[4]

At the time of the study, Wikipedia comprised 3.7 million articles in 200 languages and was ranked the 37th most visited website on the internet (Giles, 2005: 900). More than five years on and Wikipedia has become even more significant, now comprising in July 2011 in excess of 19.2 million articles in 281 languages[5] and regularly ranked in the top 10 of the World's most visited websites[6].

---

[1] Giles, J. (2005) 'Internet encyclopaedias go head to head', *Nature*, vol.438, 15 December 2005, pp. 900-901.

[2] Lih, A. (2009) *The Wikipedia Revolution.* London: Aurum Press Ltd.

[3] Encyclopædia Britannica, Inc. (March 2006), *Fatally flawed: refuting the recent study on encyclopaedic accuracy by the journal Nature*, [Online], Available at: http://corporate.britannica.com/britannica_nature_response.pdf [Accessed 11/03/11].

[4] Nature (23 March 2006), *Encyclopaedia Britannica and Nature: a response*, [Online], Available at http://www.nature.com/press_releases/Britannica_response.pdf [Accessed 11/03/11].

[5] Wikipedia (2011) *Lists of Wikipedias*, [Online], Available at: http://meta.wikimedia.org/wiki/List_of_Wikipedias [Accessed 11/07/11].

[6] For example, see Alexa (March 2011) *Top Sites*, [Online], Available at: http://www.alexa.com/topsites [Accessed 11/07/11].

During the intervening period since 2005, Wikipedia has been working to assess the quality of its entries, through its 'featured article' and 'good article' peer review process[7], and more recently through the launch of a pilot to collect feedback, which is ongoing.[8] This involves readers and editors rating articles according to trustworthiness, neutrality, completeness and readability, as well as rating their self-perceived qualification to comment. However, there has never been any attempt to replicate, better, or extend Nature's study, in order to gain a greater understanding of the accuracy and quality of Wikipedia, or make recommendations on how it may be improved.

Recently, Jimmy Wales showed interest in such work being undertaken: a study inspired by the Nature study but that employs greater rigour; one that involves academics and scholars, and examines more than just English language entries and in subjects other than solely science, so that differences in levels of accuracy and quality across languages and subject domains may be identified. This would inform decisions in the future, for example, for editor recruitment efforts and the design of expert feedback mechanisms.

The size of Wikipedia, in terms of article count and language versions, makes it inevitable that such a study will have to be large-scale in scope, and potentially costly in execution. Therefore, it is essential that prior to its commencement, there is a small-scale preliminary project that draws on empirical evidence to determine a sound research methodology. This document sets out a proposal for such a preliminary project, to be undertaken for the Wikimedia Foundation, by Epic, one of Europe's leading commercial organisations specialising in learning technologies, in partnership with Oxford University.

The intention is that the preliminary project will establish the best possible research approach, begin to hypothesize on the best way for Wikipedia to measure and communicate the accuracy and quality of articles, and provide a well-founded justification for seeking funding for a comprehensive study.

---

[7] Wikipedia (2011) *Featured articles*, [Online], Available at http://en.wikipedia.org/wiki/Wikipedia:Featured_articles [Accessed 11/03/11].
[8] Wikipedia (2011) *Article feedback,* [Online], Available at http://www.mediawiki.org/wiki/Article_feedback [Accessed 01/07/11].

2. *Aim and research objectives*

Aim: **To assess the accuracy and quality of a sample of Wikipedia entries**

For a sample of Wikipedia articles across a range of subjects and languages:

Research objective 1: Identify errors, omissions and other quality issues in Wikipedia articles.

Research objective 2: Compare with numbers of errors, omissions and other quality issues about the same subject, but in a popular alternative online version of an encyclopaedia for each language.

Research objective 3: Compare reviews among readers and editors with those of experts, using Wikipedia's article feedback tool design.

3.   Research methodology

3.1 Sampling

3.1.1 Sampling languages

As of July 2011, there are 281 different language versions of Wikipedia.[9] For the purposes of this small-scale preliminary study, three of the most popular World languages will be included. They will be sampled firstly for their popularity in terms of numbers of native speakers[10], and then in terms of numbers of Wikipedia articles[9], with the intention of choosing those with potential for a wide reach.

The top five World languages in order by numbers of native speakers are Mandarin (Standard Chinese), Spanish, English, Hindi-Urdu and Arabic. These appear in the list of number of articles per language version of Wikipedia ordered as follows: English, Spanish, Chinese, Arabic and then Hindi-Urdu. The Chinese Wikipedia is heavily censored, and so may skew the research results.[11] Therefore, the three languages to be sampled are:

English: the de facto language in the UK, Australia, USA, UAE and Malaysia, and the unifying language for countries such as Bangladesh, Botswana, India, Hong Kong, Pakistan, Philippines and Tanzania.

Spanish: the official language of Spain, as well as the de facto or de jure language of a large number of countries in Latin America, among them, Mexico, Argentina, Bolivia, Chile, Colombia, Ecuador, Paraguay and Venezuela. In addition, it is the predominant language in Equatorial Guinea, Africa.

Arabic: the official language of a large number of countries across the Middle East and North Africa, among them Bahrain, Egypt, Kuwait, Oman, Qatar, Saudi Arabia, Algeria and Tunisia. Modern Standard Arabic is based on Classical Arabic and is the literary language used in most current, printed Arabic publications and spoken by the Arabic media.

In addition, these three languages offer a range of numbers of total articles and average edits per article for Wikipedia:

Language	Ranking in the list of Wikipedias ordered by number of articles[9]	Total number of articles[9]	Average number of edits per article (2.s.f)[9]
English	1st	3,680,251	129
Spanish	6th	798,270	64
Arabic	26th	151,850	55

[9] Wikipedia (2011) *Lists of Wikipedias*, [Online], Available at: http://meta.wikimedia.org/wiki/List_of_Wikipedias [Accessed 12/07/11].
[10] Wikipedia (2011) *List of languages by number of native speakers,* [Online], Available at: http://en.wikipedia.org/wiki/List_of_languages_by_number_of_native_speakers [Accessed 16/04/11].
[11] Wikipedia (2010) *Task force/China*, [Online], Available at http://strategy.wikimedia.org/wiki/China_Task_Force [Accessed 01/07/11].

Note: At the data analysis stage, examination of this kind of quantitative data, alongside an assessment of accuracy and quality, may lead to some hypotheses for further study. For example, as numbers of edits per article increase, one may expect an increase in an article's accuracy and quality. However, the number of edits required to reach the same level of accuracy and quality across languages may differ in relation to the consensual or disputatious cultures of the contributors.

3.1.2 Sampling experts

Oxford University is repeatedly ranked in the top ten of Universities worldwide. It has over 1,600 academic staff and 4,600 graduate research students from more than 100 different countries and territories.[12] Each year, the University's Clarendon Fund offers 100 awards to international graduate students, based solely on their academic excellence, enabling the most distinguished scholars worldwide to undertake graduate studies at Oxford.[13]

In the first instance, all Clarendon Fund students, who are native English, Spanish or Arabic speakers, will be invited to express interest in participating in this preliminary study into the quality and accuracy of knowledge sharing in their subject domain. Then, eight will be randomly sampled: two each from different subject domains in the University's four academic divisions of Humanities, Social Sciences, Mathematical, Physical, and Life Sciences and Medical Sciences.

Each Clarendon student will be asked to identify five or six respected academics in their field who share the same native language (who may be from Oxford University or from a former University they have studied at) and invite them to participate in the study as well. The intention is for each student to successfully enlist three, and at the very least two, willing academics. These expert academics will have already demonstrated their academic stature by having a permanent post at a highly rated department at Oxford University or another University. They also should have achieved at least one of the following, in accordance with international measures of excellence in academia: multiple publications in respected journals within their field, multiple citations (as listed in citation databases such as Thomson Reuters Web of Science or Elsevier's Scopus); or a leading investigator role on a National Research Council project. A brief professional biography will be collected in support of their selection to participate in the research, and will be included in the final research report.

---

[12] University of Oxford (2011) *Facts and Figures*, [Online], Available at: http://www.ox.ac.uk/about_the_university/facts_and_figures/index.html#aoxford_colleges [Accessed 21/03/11].
[13] University of Oxford (2011) *About the Clarendon Fund*, [Online], Available at: http://www.clarendon.ox.ac.uk/about/ [Accessed 21/03/11].

In summary, there will be the involvement of 32 experts in total: 8 graduate students and 24 expert academics from across the three languages of English, Spanish and Arabic.

			Language		
			English	**Spanish**	**Arabic**
Division	**Humanities**	**Subject domain, e.g. Philosophy**	4 experts (1 Clarendon student, 3 academics)		
		Subject domain, e.g. Fine Art	4 experts (1 Clarendon student, 3 academics)		
	Social Sciences	**Subject domain, e.g. Economics**	4 experts (1 Clarendon student, 3 academics)		
		Subject domain, e.g. Sociology	4 experts (1 Clarendon student, 3 academics)		
	Mathematical, Physical and Life Sciences	**Subject domain, e.g. Mathematics**	4 experts (1 Clarendon student, 3 academics)		
		Subject domain, e.g. Zoology	4 experts (1 Clarendon student, 3 academics)		
	Medical Sciences	**Subject domain, e.g. Neurology**	4 experts (1 Clarendon student, 3 academics)		
		Subject domain, e.g. Pathology	4 experts (1 Clarendon student, 3 academics)		

Note: There may not be even spread across the languages, but this is not deemed important for the purposes of this preliminary study which aims to begin to formulate hypotheses and test the research methodology for a larger-scale study.

3.1.3 Sampling articles

Each of the eight students will be asked to provide a brief description of between 400 and 500 words of their area of specialism. Key words in each description will be used to identify a selection of articles of relevance to their particular expertise that appear in both Wikipedia and in the most popular online alternative encyclopaedia for each language.

For English language the chosen alternative encyclopaedia will be Encyclopaedia Britannica. As well as being the oldest English-language encyclopaedia, it was also the encyclopaedia originally chosen by Nature to compare with Wikipedia.[14] For Spanish, the chosen alternative will be Encicilopedia Universal en Espanol. This is the Spanish version of Encyclopedia Britannica. There are other notable online encyclopedia's in Spanish but many are specialist (for example, Ecured which offers a Cuban perspective on the World, or Vikidia, which is for children).[15] In addition, there is also the popular Enciclopedia Libre, but this was seeded by the original Spanish version of Wikipedia, so may lead to risk of comparing articles from the same original source.[16] For Arabic, the chosen encyclopaedia is Dahsha. There are four general online encyclopaedia alternatives to Wikipedia in Arabic,[17] but according to the web information site, Alexa[18], Dasha has the greatest global traffic, as determined by the number of users who visit a site, as well as the number of pages on the site viewed by those users over a three month period.

---

[14] Giles, J. (2005) 'Internet encyclopaedias go head to head', *Nature*, vol.438, 15 December 2005, pp. 900-901.
[15] Wikipedia (2011) *List of online encyclopedias,* [Online], Available at: http://en.wikipedia.org/wiki/List_of_online_encyclopedias [Accessed 03/07/11].
[16] Wikipedia (2011) *Enciclopedia Libre Univeral en Espanol* [Online], Available at: http://en.wikipedia.org/wiki/Enciclopedia_Libre [Accessed 03/07/11].
[17] Wikipedia (2011) *Arabic encyclopedia* [Online], Available at: http://en.wikipedia.org/wiki/Arabic_encyclopedia [Accessed 03/07/11].
[18] For example, Alexa (2011) *Dahsha.com* [Online], Available at: http://www.alexa.com/siteinfo/dahsha.com# [Accessed 03/07/11]

For each pair of articles, the one from Wikipedia will be categorised according to whether it has been addressed in a 'featured article', 'good article' or one yet to be assessed on Wikipedia.[19] Then, two pairs of articles will be chosen for each student, ensuring the Wikipedia articles are from different categories. (This will establish a spread and enable comparison between expert review and existing quality mechanisms used by Wikipedia for their articles. In addition, these two pairs of articles should be added to Wikipedia's current pilot to collect feedback, if not already included, to enable, in the longer term, comparison between reader and expert review). Undoubtedly, articles will vary in their subject matter, so, for example, some may be about people, others about places, and others about artefacts or phenomena.

First of all, Clarendon student participants will be asked to read through the selected articles to ensure they are of 'reasonable quality' for inclusion in the research study. In particular, students will be asked two key questions about each article – would they describe it as 'very poor quality', and 'does it appear to have been victim of deliberate vandalism' (as can sometimes be the case with online marterial)? If their answer to either of these questions is 'yes' then an alternative pair of articles will be selected of relevance to their expertise, still retaining the spread across 'featured', 'good' and yet-to-be-assessed articles from Wikipedia. This will ensure academic experts' time is not wasted reviewing incontestably inadequate articles as part of the study.

In summary, each student and their selected three academics will examine two pairs of articles (from two different topics of relevance to their particular expertise in their subject domain), amounting to 16 pairs of articles in total.

			Language		
			English	Spanish	Arabic
Division	**Humanities**	Subject domain, e.g. Philosophy	2 pairs of articles		
		Subject domain,e.g. Fine Art	2 pairs of articles		
	Social Sciences	Subject domain, e.g. Economics	2 pairs of articles		
		Subject domain, e.g. Sociology	2 pairs of articles		
	Mathematical, Physical, and Life Sciences	Subject domain, e.g. Mathematics	2 pairs of articles		
		Subject domain, e.g. Zoology	2 pairs of articles		
	Medical Sciences	Subject domain, e.g. Neurology	2 pairs of articles		
		Subject domain, e.g. Pathology	2 pairs of articles		

Note: It is anticipated that there will be instances where experts will agree to participate, but then will not be able to fulfil their commitment due to other work priorities. This may mean less expert involvement, or some experts only reviewing one pair of articles, rather than two. The study has been designed to accommodate a 'drop out' rate, with an expected absolute minimum of one pair of articles per each subject domain reviewed by two experts. This would still enable peer review, and also meet the aims and objectives of this preliminary project. This also compares favourably to the Nature study which involved a pair of articles (one from Wikipedia and one from Encyclopedia Brittainica) being peer reviewed by a single expert only.[20]

---

[19] See Background p4 for further information

[20] Nature (14 December 2005) '*Supplementary information to accompany Nature news article "Internet encyclopaedias go head to head" (Nature 438, 900-901; 2005)'* [Online]. Available at: http://www.nature.com/nature/journal/v438/n7070/full/438900a.html [Accessed 5/7/11].

3.2 Data collection

To meet research objectives 1, 2 and 3:

> **Research objective 1:** Identify errors, omissions and other quality issues in Wikipedia articles.
>
> **Research objective 2:** Compare with numbers of errors, omissions and other quality issues about the same subject, but in a popular alternative online version of an encyclopaedia for each language.
>
> **Research objective 3:** Compare reviews among readers/editors with those of experts, using Wikipedia's article feedback tool design.

Expert academics will be informed that they are reviewing pairs of articles, on topics within their expertise, from encyclopaedias. They will not be told which encyclopaedias, and, before embarking on the review work they will be asked to sign a declaration that they will make no attempt to identify the sources and will not use any other encyclopaedia (online or otherwise) to help verify information.

Then the academics will be invited to access articles via a login to an independent wiki, set up specifically for this research study. They will be asked to complete a survey, embedded into the wiki, to rate the articles in terms of their own scholarly judgement. In particular, they will be asked to judge whether they would accept or recommend citation of the article by their students, in terms of whether they would consider the article as a whole constitutes a 'fair' representation of the topic. They will be encouraged to explore this question in detail using the following broad criteria[21]:

- subject coverage/comprehensiveness
  For example, they may consider whether the introduction successfully establishes the topic without key omissions, and then whether the article in its entirety addresses all key points explicitly and thoroughly, providing a sound and broad foundation to understanding the topic.

- accuracy
  For example, they may identify any digression off-topic, as well as any misleading statements, inappropriate emphases and factual inaccuracies.

- verifiability by references
  For example, they may state whether sources that are cited are relevant and given the appropriate significant weighting – if they are important then one might expect them to be substantively, rather than superficially, consulted.

---

[21] criteria adapted from Besiki Stvilia, Michael B. Twidale, Linda C. Smith, Les Gasser (2005) Assessing Information Quality Of A Community-Based Encyclopedia, & (2008) Information Quality Work Organization in Wikipedia p. 993

- arrangement of information

  For example, they may comment on whether the sequential paragraphs enable the clear development of ideas.

- objectivity and non-controversiality

  For example, they may consider whether all notable criticism and controversies are acknowledged, and if so, from all perspectives.

- use of language

  For example, they may comment on the extent to which the language is clear, readable and used in a concise and efficient fashion, with appropriate explanation given to any scholarly terminology or other jargon.

- currency

  For example, they may comment on whether the article includes the most up-to-date information and makes reference to the latest thinking.

We see a major aim of this initial project as developing robust ways of meeting the methodological challenge of capturing professional academics' perspectives on issues of quality and accuracy, and our aim will be to capture and make explicit the kind of serious judgements that academics make all the time in their work. Therefore, in all cases, academic experts will be asked to justify, and cite evidence for, the judgements that they make. In addition, they will be asked to suggest appropriate corrections, providing a rationale, as well as full bibliographic references, where they have used or cited any information from elsewhere.

Where there are instances of experts reviewing the same articles but offer opposing views, there will be an invitation to those experts to provide a written comment on their colleagues' responses. Such comments will be a research finding in itself and will be described in the final research report.

3.3 Data analysis

For quantitative data, such as ratings, statistical analysis will provide frequencies and central tendencies. These will be further understood by qualitative data analysis of edit commentaries and survey comments, which will be coded and organised according to common motifs.

All data also will be set in the wider context of information that can be obtained from Wikipedia, for example the number of existing edits to an article when provided to the experts involved in this study, or any existing reader ratings achieved through Wikipedia's pilot to collect feedback.

Preliminary themes will be drawn from the data and compared across languages, subject domains and divisions, and students and academics, leading to possible patterns being identified.

3.4 Future research directions

It is anticipated that this small-scale preliminary project will result in potential hypotheses being formulated, for example, about how best to measure the quality of articles, or how best to collect expert review. In addition, critique of the research methodology will enable informed decisions about how to scale up to include a greater sample of languages, experts and articles. Therefore, it is anticipated that the report that results from this preliminary study will form the sound and rigorous crux to a proposal to seek funding for a follow-on comprehensive larger-scale study.

4.   Expertise

4.1 Research expertise

Epic is a long-standing, award-winning UK market leader in the design and build of bespoke learning solutions for use on a wide range of technology – from PC to mobile to games consoles. Epic frequently engages in research to enable organisations and their people to maximise potential from learning and development, and any investment they make in learning technologies.

Dr Naomi Norman is Epic's Director of Learning and sits on the Company's Board. She leads its research programs. Last year she delivered an evaluation study for Procter and Gamble, examining the success of an e-Portal designed for its global workforce to achieve a unified approach to pricing. More recently, she has delivered a large-scale research study investigating the potential of mandatory training offered via mobile devices to all UK National Health Service staff (the World's third biggest employer).

Naomi has also undertaken a number of research projects prior to joining Epic: as a consultant to the UK Government, appraising a program to embed IT across the curriculum in UK schools; and for Oxford University, independently investigating its admissions process for candidates applying to its Medical School. She also has much experience of working with organisations with an interest in using technology for learning, among them the BBC, the Olympic Games Education Programme and Sesame Street.

> *Dr Norman's contribution to the design, conduct and analysis of interviewers' beliefs and behaviours was critical in the successful outcome of the Oxford Medical School Admissions reforms. Her approach was highly professional, imaginative and technically skilled.*
>
> W.S. James, Professor of Virology
> Sir William Dunn School of Pathology
> University of Oxford

> *What a fantastic report! I can see that it provides us with clear recommendations and information to steer and support our next steps. It has been a pleasure to work with you and I look forward to doing so more in future.*
>
> Helen Bingham
> Workforce, Education & Leadership, NHS

Naomi has a long-standing working relationship with Dr Chris Davies of Oxford University's Department of Education, having studied for her D.Phil (PhD) in the department eleven years ago, and having involved Chris in the design and peer review of both the Procter & Gamble, and then the NHS mobile research projects.

Dr Chris Davies leads the Learning and New Technologies Research Group at the Oxford University Education Department, where he also directs the MSc in Learning and Technology. His main current research activity is as Project Director for the Learner in their Context Project, which investigates the ways in which young learners in school and university use digital technologies for learning outside formal education. This was a £1 million project funded by the UK Government agency Becta.

Previously he was the lead education expert in the collaborative team which conceptualized and designed Immersive Education's *kar2ouche* software, and led the research project investigating its impact within classrooms, for which he was awarded (as a co-recipient) of the British Computer Society 2002 Award for Educational IT first prize.

He was lead author for the UK version of Intel's Teach to the Future programme. He is a Research Associate at the Oxford Internet Institute. As Vice President of Oxford University's Kellogg College he helped to expand provision for part-time study at the University, as well as playing a leading role in the development of a major new site for the College in Oxford. He was a co-organiser of the recent ESRC Research Seminar series, *The educational and social impact of new technologies on young people in Britain* which ran from March 2008 until July 2009.

4.2 Fund-raising expertise

Oxford University has the highest research income from external sponsors of any UK Universities and its research income has doubled in the last five years. In particular, it has a strong record of winning funding from the UK Research Councils and various Trusts, among them, the Economic and Social Research Council (ESRC), the Leverhulme Trust and the Nuffield Trust, in the case of Oxford University's Education Department.

Epic has a Bid Director, Alec Keith, who heads a team of three full-time staff, dedicated to compiling bids and attending pitches to secure Epic funding and work. Alec's team has a long track-record of writing persuasive and engaging proposals and presentations, with excellent win rates. In the past it has been involved in joint bids with Universities and other organisations to secure research funding.

In addition, Epic's chairman, Andrew Brode, is a well-respected UK businessman with many contacts in the City of London and a background in publishing.

Andrew qualified as a Chartered Accountant and worked with Arthur Andersen and Rothschilds. He entered the business publishing arena as Managing Director of Croner Publications which he built up and sold to Wolters Kluwer. He was subsequently appointed CEO of Wolters Kluwer (UK) PLC and expanded this to UK sales of £80 million in the 1980's.

He acquired Eclipse Publications via an MBI in 1990 which became a leading journals and conference organiser, being sold to Reed Elsevier in 2000. Today, he is the Chairman and major shareholder in RWS Holdings PLC, Europe's largest technical translations group which floated on AIM in late 2003, and is Chairman or a Director of a number of listed and private venture capital-backed media companies. In June 2008 he acquired Epic Group PLC, the UK market leader in e-learning.

He is a member of the Court of the Worshipful Company of Stationers and Newspaper Makers, one of the oldest Livery Companies of the City of London, and its current Treasurer. He has been a Fellow of the Royal Society of Arts since 1995.

5. Research timetable

	Date	Activity	Responsibility	Stage
Finalisation of proposal				
D1	**By Fri 17 Jun**	Submit updated proposal to Wikimedia Foundation	Dr Naomi Norman, Epic Dr Chris Davies, Oxford University	Complete
M1	**By Fri 24 Jun**	Wikimedia Foundation sign off on proposal	Dario Taraborelli, Wikimedia	Complete
Contractual agreements				
D2	**By Fri 7 Oct**	Grant agreement between Wikimedia Foundation and Epic is signed	Dr Naomi Norman, Epic Dario Taraborelli, Wikimedia	
D3	**By Fri 7 Oct**	Consultancy agreement between Epic and Oxford University is signed	Dr Naomi Norman, Epic Dr Chris Davies, Oxford University	
University co-operation				
T	**By Fri 1 Jul**	Meet with Clarendon student Council	Dr Naomi Norman, Epic Dr Chris Davies, Oxford University	Complete
Appointment of research assistant				
T	**By Fri 14 Oct**	Identify/interview potential research assistants	Dr Naomi Norman, Epic Dr Chris Davies, Oxford University	
M2	**By Mon 17 Oct**	Appoint research assistant	Dr Naomi Norman, Epic Dr Chris Davies, Oxford University	
Sampling				
T	**By Fri 21 Oct**	Send letters/emails to Clarendon students with English, Spanish and Arabic as their first language, asking for expression of interest to participate in this research study	Research assistant Dr Naomi Norman, Epic Dr Chris Davies, Oxford University	
T	**By Fri 28 Oct**	Send follow-up letters/emails to interested Clarendon students, with accompanying questionnaires seeking information on their discipline, area of specialism and known academic experts in their field	Research assistant	
T	**By Fri 4 Nov**	Sample and recruit two Clarendon students from each of the four academic disciplines: Humanities, Social Sciences, Mathematical, Physical and Life Sciences, and Medical Sciences	Research assistant Dr Naomi Norman, Epic Dr Chris Davies, Oxford University	
T	**By Wed 23 Nov**	Clarendon students confirm two willing academic experts from their home country	Research assistant Dr Naomi Norman, Epic Dr Chris Davies, Oxford University	

T	**By Mon 28 Nov**	Identify pairs of articles, of relevance to expertise of students and academics, in Wikipedia and popular alternative encyclopedia for each language	Research assistant Dr Naomi Norman, Epic Dr Chris Davies, Oxford University	
M3	**By Wed 30 Nov**	Complete sampling	Research assistant Dr Naomi Norman, Epic Dr Chris Davies, Oxford University	
D4	**By Fri 2 Dec**	Send document clarifying sample of research participants and articles to Wikimedia	Research assistant Dr Naomi Norman, Epic Dr Chris Davies, Oxford University	

Desk research				
T	**By Mon 21 Nov**	Familiarise with relevant literature to place research in appropriate context	Research assistant Dr Naomi Norman, Epic Dr Chris Davies, Oxford University	
D5	**By Wed 23 Nov**	Send list of references of articles reviewed to Wikimedia	Research assistant	

'Article review' wiki				
T	**By Fri 2 Dec**	Write 'article review' wiki specification	Research assistant Dr Naomi Norman, Epic Dr Chris Davies, Oxford University	
M4	**By Thu 22 Dec**	Build 'article review' wiki, incl. articles	Epic	
D6	**By Fri 9 Jan**	'Article review wiki' released to Wikimedia for comment	Research assistant Dr Naomi Norman, Epic Dr Chris Davies, Oxford University	

Data collection				
T	**By Mon 30 Jan**	Collate relevant data from Wikipedia to inform data analysis	Research assistant Dr Naomi Norman, Epic Dr Chris Davies, Oxford University	
M5	**By Mon 30 Jan**	Complete collecting articles reviews	Clarendon Students Academic experts	

Data analysis				
T	**By Mon 13 Feb**	Measure frequencies and tendencies and code data	Research assistant Dr Naomi Norman, Epic Dr Chris Davies, Oxford University	
T	**By Mon 13 Feb**	Draw themes from data	Research assistant Dr Naomi Norman, Epic Dr Chris Davies, Oxford University	
M6	**By Wed 22 Feb**	Complete data analysis	Research assistant Dr Naomi Norman, Epic Dr Chris Davies, Oxford University	

Reporting				
M7	**By Wed 14 Mar**	Write draft report, including recommendations for next steps/wider study	Research assistant Dr Naomi Norman, Epic Dr Chris Davies, Oxford University	
D7	**By Thu 15 Mar**	Draft Report released to Wikimedia for comment	Research assistant Dr Naomi Norman, Epic Dr Chris Davies, Oxford University	

D8	**By Mon 19 Mar**	Present findings via WebEx	Research assistant Dr Naomi Norman, Epic Dr Chris Davies, Oxford University	
Knowledge sharing				
D9	**By Fri 30 Mar**	Submit final report and upload to Wikipedia	Research assistant	
D10	**By Mon 30 Apr**	Present research at: Wikimania Oxford University Department of Education Oxford Internet Institute	Research assistant Dr Naomi Norman, Epic Dr Chris Davies, Oxford University	
D11	**By Mon 30 Apr**	Submit paper to a conference, e.g. WikiSym, American Educational Research Association annual conference, and a journal, e.g. Oxford Review of Education and an open access journal	Research assistant Dr Naomi Norman, Epic Dr Chris Davies, Oxford University	

6. Costs

Activity	Time	Cost
Dr Naomi Norman or Dr Chris Davies to recruit and induct research assistant	2 days	£1,800
Research assistant to undertake research activities to include: • Identifying, securing and then communicating with Clarendon students • Communicating with students' selected academics • Sampling articles according to student expertise in subject domains • Building (with some input from Wikipedia) and uploading articles to a wiki, including embedded survey • Engage in 'communications campaign' to keep research participants interested and encourage them to undertake their tasks • Collating relevant data from Wikipedia that may be used for data analysis • Undertake first pass data analysis under guidance • Produce detailed notes to inform first draft research report	10 weeks	£7,500
Arabic and Spanish natives to aid with identifying the relevant articles in other language encyclopaedias.	4 days @£300/day	£1,200
Payment to 8 Clarenden students for: • Identifying and securing two academics each • Reviewing two pairs of articles each	12 days @£280/day (1.5 days per student)	£3,360
Payment to 24 academics for review of two pairs of articles each	12 days @800/day (0.5 day per academic)	£9,600
Some translation to aid data analysis	2 days @£300/day	£600
Moderation and peer review throughout; then final write up by Dr Naomi Norman and Dr Chris Davies	10 days @£900/day	£9,000
	Total	**£33,060**

Assumptions: All costs exclude VAT; any travel and sustenance expenses to be billed additionally at cost.

Notes: Typical day rates for Dr Naomi Norman and Dr Chris Davies are in excess of £1,000 per day. We have agreement to reduce them to £900 per day in support of the work of Wikipedia. Experts who do not complete their task, i.e. who only review one pair of articles rather than two, will be paid proportionally.

Rate card

Description	Standard day charge	Number of project days	Total staff cost
Research academics to lead the project	£900	12	£10,800
Research assistant	£150	50	£7,500
Native speakers/translators for article selection	£300	4	£1,200
Graduate students	£280	12	£3,360
Academics	£800	12	£9,600
Translators for data analysis	£300	2	£600
			£33,060