

**Toward Global Open Scholarship – Access to Research in Development
and Globalization**

by

Arif Jinha

Thesis submitted to the
Faculty of Graduate and Postdoctoral Studies
In partial fulfillment of the requirements
For the M.A, degree in
Globalization and International Development

School of International Development and Global Studies
Faculty of Social Sciences
University of Ottawa

©Arif Jinha, Ottawa, Canada, 2012

Abstract

Two centuries after the printing press was invented, the first scholarly journal appeared in 1665. Less than two decades after the journal went online, the digital format is reshaping scholarly communication rapidly. We are moving quickly towards an open system of scholarship, and from a Western heritage of print scholarship to a future of global knowledge, a shift driven by the communications revolution. This thesis provides data describing the size and growth of the universe of scholarship, its global reach, how much of it is accessible free of charge on the internet and the rate at which that share is growing. Open Access together with development programs aimed at reducing price barriers to subscription journals have vastly increased the possibilities for accessing research in the South. The relevance to globalization and development is explored conceptually and revealed in the results.

Acknowledgements

Much gratitude to my brother and fellow scholar Azim Jinha, who has helped in numerous ways, from assisting with math to copy-editing.

Many thanks to my readers Gilles Breton and Leslie Weir, an interested and kind jury who have given excellent feedback and encouragement. Thank you to many other knowledgeable and interested people who discussed this thesis with me.

Thank you to my supervisor Michael Geist for razor-sharp guidance and allowing me time and freedom to continue this work to its completion.

Many thanks to family and friends who have pulled me through and given me inspiration.

Finally, thanks to the young people around the globe, who are fighting the odds to create a generation of greater openness and equity in the world.

Foreword

In 2006, I was hired to coordinate the International Health (IHP) at the University of Ottawa (now the IHDP – International Health and Development Program). My staff were 3 dynamic students who coordinated the energy of a dozen volunteers. The program use a peer-based education approach. My staff had been doing some excellent work on their own through the summer, and their chosen theme for the fall was Access to Medicines. I reflected that such access was obviously more empowering than lack of access, but it didn't get at the root problem. Why was there no indigenous research and development on health and medicine established within the universities of the South that created appropriate health interventions, of which medicine is but one, for diseases affecting their populations?

I coined a term in my head '*access to knowledge*' and began thinking about it as a 2-way, street contemplating the loss of benefits to the South of exclusion from global literature, and contemplating the weakness of understanding in the North and in the development community stemming from the minimal production of Southern literature, as well as the loss of diversity in global literature. This I felt went all the way to the root in philosophy of science and knowledge, and the limitations and bias towards what can enter the world a accredited knowledge. This is a loss to me in my efforts towards a global education, a loss to such programs as IHP whose mandate is to promote 'global citizenship'. This is a loss with regard to solutions to 'global' and 'development' problems, and it creates the asymmetrical relationships in development assistance. This imbalance drives a great deal of cynicism about development which is too often critiqued for its paternalism and structural inequality. How long, I thought, will the wonderful ideal of self-determination in development codified in the Right to Development go unrealized?

As is often the case when I think up a term, it already exists and this was discovered by 'googling' it. There was an A2K movement already! As I looked into it further, I learned about that aspect that was most specific to what we do at university was Open Access to research literature and Open Educational Resources. During this time, I met with an African colleague, a PhD nursing student on exchange coming out of University of Ottawa's Faculty of Health Sciences partnership with the Tropical Institute for Community Health in Kenya – now the Great Lakes University of Kisumu (GLUK). She gave me the idea for what became the research proposal for this thesis. It turned out that two of GLUK's key interests were working together towards 2 goals. 1. Broadband wireless internet access throughout GLUK's campus. 2. Access to digital library resources.

The ideas grew with conversations with Dr. Diack over Skype, into the Development Equation (Deq). The Deq seeks to articulate a vision for 21st century development based on addressing the barriers of access to the internet itself – the ICT4D dimension, barriers to access to content for development – the Openness dimension –

OA, OER, FLOSS (free/libre open source software) and the participatory web (web 2.0 or read/write web). With regard to universities, the vision is to exploit digital opportunity toward networked partnerships (NP) across universities in the North and South.

This work-in-progress is viewable on the web on WikiEducator under the title 'Learning and Knowledge in the 21st century' with the goal of eventually publishing an academic book to elaborate this vision of the future of universities in the South. Dr. Diack invited me to present my early thesis work at UCAD in spring of 2008. Having initiated our relationship digitally, this was our first in-person meeting!

By then I was putting the final touches on an article for publication. One of the empirical research questions to act as a starting point for the big picture, was 'how many articles exist?' Having figured out a way to answer this question, I was able to publish that article by 2010 in *Learned Publishing*. I re-worked my M.A. thesis to focus on a few key indicators laid out when we constructed the Deq as a theoretical model. These are focused on the size, dynamics and growth rate of the scholarly system, the global distribution of participants, the share of articles that are OA, and the trend in the growth of that share. The discussion of the transition from print to digital as an aspect of globalization is unavoidable. Change is what defines development, and though the subject is not typical of International Development Studies (IDS), the subject both strikes at the core of human development and at the activity of IDS as a scholarly field rooted in social science, engaged in research, education and publishing research. The thesis fits into study of globalization and international development and has a degree of self-reflexivity. This is a work of Open Access scholarship.

I thank you for reading about a subject that is less dramatic than many discussions in development, those dealing with poverty, war, climate change, injustice etc. Though less dramatic, the development that this thesis deals what appears to a more benign and optimistic change, simply a trend towards dramatically more fair and better access and participation in the world of peer-reviewed journal articles throughout the world. There is a long way to go and plenty of drama and conflict among those involved in academic publishing, and thankfully no violence.

I hope though, you find in here something close to what we do that is quite positive for development, something you can be engaged in yourself as a student or researcher. There are actionable things to do at your university, and if are referencing this for your own publication, you have many options to make your work available globally without price barriers in support of your moral and material interests as a scholar.

Table of Contents

Abstract.....	ii
Foreword.....	iii
Table of Contents.....	v
List of Tables.....	vii
List of Figures.....	viii
Chapter 1. Introduction.....	9
1.1: In the making - a short history of OA.....	11
1.2: Development as Change - Scholarship in Transition	12
1.3: Definitions – the peer-reviewed scholarly article	13
1.4: Open Access Definitions:.....	14
1.4.1: OA publication charges – a distinction left out of the definition	16
1.4.2: Operational Definitions versus Discussion and Concepts	16
1.5: Definitions and Distinctions - Globalization? Development?.....	19
1.5.1: Development	19
1.5.2: Globalization	21
1.6: ICT4D and K4D not included.....	22
Chapter 2. Rationale, Context and Theory – Literature Review.....	24
2.1: Development Context - Universities in Sub-Saharan Africa.....	24
2.2: Review of Key Theoretical Literature - Development as Economic Change.....	26
2.2.1: Development as Freedom – Amartya Sen.....	27
2.2.2: Knowledge as a Global Public Good - Joseph Stiglitz.....	28
2.2.3: Path Dependence in Economic Change – Douglass North.....	31
2.3: OA Arguments.....	32
2.3.1: The Argument from Principle.....	32
2.3.2: The Economic Argument.....	33
2.4: OA in developing countries.....	34
2.5: Stakeholders - Varying Actors and Interests.....	35
Chapter 3. Research Questions.....	38
3.1: The global static universe of peer-reviewed articles.....	38
3.2: The dynamics of change in the system of scholarly communication – share of OA in the current output of peer-reviewed articles.....	38
3.3: Change in the global static universe of articles:.....	39
Chapter 4. Metrics of global access to research – towards a method.....	40
4.1: Why periodization?.....	40
4.2: Article 50 Million – published en route to the thesis.....	44
Chapter 5. Methods.....	46
5.1: The Ulrich's universe' of titles and articles - operationally defining the universe of articles.....	46
5.2: Databases/Indexes Used:.....	48
5.2.1: Filters.....	48
5.3: Backward mapping.....	49
5.4: The altered data set	49

5.5: Other recent estimates.....	50
5.6: Conservative Adjustments.....	51
5.7: No Direct Method to Track Green OA Growth Trends.....	51
Chapter 6. Results and Analysis.....	53
6.1: The global static universe of articles:	53
6.1.1: Global total of articles: How many journal articles exist?.....	53
6.1.2: Rate of increase of universe of articles: What is the rate at which the universe of articles is growing?.....	53
6.2: Growth of the Static Universe of Articles	55
6.3: Shares of OA.....	59
6.3.1: Green OA.....	59
6.3.2: Gold OA growth.....	61
6.3.3: Adjustment of OA totals to reflect Ulrich's universe.....	61
6.4: Dynamics of current system	63
6.5: Rate of increase of OA Shares:.....	64
6.6: Implications of Data and Assumptions for Forecasting.....	66
6.7: Titles.....	67
6.8: OA Share of All Literature.....	68
6.9: Backdated OA?.....	70
8. Development and Globalization - Africa in focus	72
9. Conclusion – Global Open Knowledge.....	76
Chapter 7. Bibliography.....	78

List of Tables

Table 1: Definitions of development and globalization.....	9
Table 2: Comparison of toll and open and open access.....	13
Table 3: Definitions of different access types.....	18
Table 4: Comparison of various terms.....	20
Table 5: Metrics of the global scholarly system.....	68
Table 6: Share of researchers 2002 and 2007- Africa.....	70
Table 7: Researchers per million 2002 and 2007 - Africa.....	70
Table 8: Publications 2002 and 2008 Africa	70
Table 9: Summary of development-oriented journal access programs.....	72
Table 10: Directories of open access sources.....	72

List of Figures

Figure 1: Estimated annual global research article output at 3% annual growth.....	43
Figure 2: Ulrich's report of the rate of publication from 1665 to summer 2001 (Ulrich 2001).....	45
Figure 3: 4.5% increase in articles 2000 to present - Period of Globalization and Digital Availability:	54
Figure 4: Annual output.....	56
Figure 5: The total article universe (grey)plus annual output (black).....	57
Figure 6: The growth of OA mandates – from http://roarmap.eprints.org/	58
Figure 7: Possible trends for OA Growth.....	63
Figure 8: Entire article universe - projected share of Toll, OA and public domain articles in 2010.....	67
Figure 9: % Gold OA AJOL vs. Web of Knowledge.....	72

Chapter 1. Introduction

This thesis deals with the phenomenon of Open Access (OA) to Scholarship and its significance to global studies and international development. It is focused on the quantifiable globalization of scholarship and its impact on developing country research, particularly in Africa. The research questions attempt to quantify the size of the scholarly system and its growth, and the growth of Open Access within it. The results demonstrate that the trend in scholarly communication is towards globalization and openness, and this is mediated by rapid change in communications technology.

Development	Open Access as Development and Open Access for Development
Globalization	Of Scholarly Communication – toward a new period of global, open scholarship.

Table 1: Definitions of development and globalization.

Access to research is a development issue because (a) peer-reviewed articles communicate quality-controlled knowledge and information; b) the benefits of scholarship are unequally distributed according to the wealth of nations; d) access to research is development but not development aid, rather it serves self-determination, capability and reciprocal benefit. This last can be contrasted with many critiques of development aid – that there are 'strings attached', collaboration is asymmetrical, relationships form dependance, donors and patronizing and paternalistic, recipient institutions are fragile, interventionists do not understand local culture and create harm through misguided good intentions, and partners on either side are prone to corruption and incompetence. Yet, we suspect that development aid is more effective when it is met with capacity and strength from within. A major leap forward in access to research literature is but one spoke in that wheel, and in the tradition of scholarship our expectations should nevertheless be humble and open to further discussion.

The research proposal for this thesis began with an elaborate conceptual and empirical model to discuss university development in the South in the 21st century, with access to research literature at centre. This was refined to the primary focus here on the scope, growth, accessibility and change in scholarly communication. The results of my more focused thesis point to an elaboration of the periodization of science/scholarship and its communication. The current period of digital and globalizing scholarship that trends toward open scholarship represents a significant change for local research environments throughout the South, and hence for self-determination in development. From the beginning the region in focus has been Africa, and some evidence particular to Africa is presented.

An Open Access article is globally available with access to the internet. Markets for journal subscriptions behave very differently across regions of the world, but these are becoming more global as well. Development-oriented bargaining and concession programs have been launched in the last decade, creating new access routes in the global South to top-rated journals from the world's largest and most prestigious publishers. These are the same publishers whose market power has been cited to have pushed up subscription prices in the North leading to the 'serials crisis'. The 'serials crisis' together with the opportunity of digital publishing provoked Open Access to begin with, independent of a development rationale. The mutual North-South interest in open scholarship and the end of exclusion by price barrier creates a 'win-win' possibility for moving research communication forward collaboratively. The pace of change in globalization is rapid because of global real-time communication (Scholte, 2005).

Despite steady advance toward open scholarly communication, the presence of price barriers slows research communication, leaving scholars behind the curve at a time when the sharing of good information is critical. Humanity needs researchers *to know about the world* and to share that knowledge, and all are disadvantaged by limits placed on the source of talent, perspective, local knowledge and ideas. 'Have-nots' are at a greater disadvantage and taken as a sum, globally everyone loses and problems grow while our development ideals and efforts lead to disappointment and disillusionment. We commonly understand ourselves to be living in an information/knowledge economy, in a global information/knowledge society. Libraries and literature have long been understood to be pivotal in enlightened development, from Alexandria to our emerging global digital library. We can take stock in order to 'know ourselves first', to be reflexive the role scholarship in global development and in knowledge, and to understand how scholarly communication is changing.

This brings us to the empirical study here and the research questions are concerned with measuring the scope, growth and change of scholarly communication by its molecular unit – articles (Björk, 2008). Further, we are interested in the change towards greater, more global and more open participation in that system, as a development issue. The lineage of work behind the empirical research presented here (apart from its beginning with the first journal published in 1665) begins with Derek de Solla Price in the 1960s (Price 1963 in Tenopir and King, 2009), pioneering the field of scientometrics and studies of the system of scholarship. Price identified the periods of Little Science before WWII and Big Science he was part of in the post-War years. Noting some decline in the year to year output of publications after the mid-1970s Mabe and Amin (2003) added the period of 'disillusionment' to the periodization.

There now a direct, specific and small line of empirical research describing the self-organizing nature of the system of scholarly communication (Mabe, 2001) started by Price and continued by efforts by others from that point to the present. The work is based on studied relationships between the numbers and growth of active journal titles, the annual output of journal articles and the numbers of researchers. Recently, this has

been extended by Bo-Christer Björk (Björk et. Al 2008, 2010, 2011) and his team to look at the share of research that is Open Access. My results are based on this line of thinking that has come a long way in getting a more precise understanding of the universe of scholarship. The results best delivered by coming full circle to Price and elaborating on the periodization as follows:

1. Early Literature - 1665 – 1860
2. Little Science1 - 1860 - 1940 (Mabe, Price)
3. Big Science - 1945 - 1974 (Mabe, Price)
4. Disillusionment – 1975 to 2000 (Mabe)
5. Global Digital Science/scholarship– 2000 – current
6. Open Science/scholarship – current to future

Periods 1 to 4 form the historical context, and periods five to six are the focus of the thesis. A focus on scholarly communication, taken either from an interest in access to research from the lens of development, or taken directly forward from this line of research would naturally arrive at the themes of digital communication, globality and openness as key features of change.

1.1: In the making - a short history of OA

The first journal is reported to have been Journal des sçavans created in France in 1665 by Denys de Sallo. It was based on a project by François de Mézeray to archive and disseminate invention and scientific research. Mézeray and De Sallo lived in the same house and corresponded with Henry Oldenburg of the Royal Society in London, Oldenburg contributed to the first volume and founded Philosophical Transactions, which is frequently cited to be the first journal later the same year, but is actually the oldest journal still active today. Style, convention and content has certainly changed, but the basic model was there particularly in the purpose of sharing and preserving original research, Mézeray's motivation (Brown, 1972). The transformation of the model created by Mézeray and Oldenburg into a system of communication occurred over the next century, with Mabe (2000) reporting steady growth of article publishing for more than three centuries by the end of the twentieth. The advent of the journal follows the invention of the printing press and its development from the 14th century onwards.

Before the advent of the first online journals in the early 1990s, the only route to accessing journal articles was through print periodicals and fair use copying of these. The advent of the world wide web in 1994 positioned the world for a broader shift to digital communication of all major forms of media, led particularly by the easiest to transport – print. The original internet (ARPANET) was developed to serve communication needs in an information and research intensive military context, and the world wide web occurred as it broadened into universities (Hafner and Lyon, 1996).

The earliest adopters naturally were computer scientists and physicists. In 1995, Stevan Harnad posted his famous 'subversive proposal', which predicted the implications for open access through the shift to online journals (Harnad, 1995). The move to digital format was not subversive in and of itself, but combined with some insight into digitally expanding the traditional sharing of re-prints, it was.

Digital open access journals existed since 1990 and before the advent of the world wide web, as did open access archives such as Arxiv established in 1991 but on a small scale. Since 1995 and the www, two alternatives emerged for the free availability of articles independent of purchasing power. One possibility was that the journal itself could limitlessly disseminate copies of a digital article at no extra cost. This paved the way for Open Access journals, or the Gold route to OA. A second possibility based on the tradition of author's sharing data and article reprints was for authors to archive their work in a publicly accessible location; the green route to OA.

Open Access is situated closer to the centre in the 'copyleft' movement, and is related to Open Educational Resources (OER), Free and Libre Open Source Software (FLOSS) and Creative Commons, which challenge the traditional business model but do not infringe copyright. As such, it is far less subversive than hacking, piracy and counterfeiting, and some internet activists favor a more radical approach to openness. This had rarely touched journal publishing, but recently Harvard researcher Aaron Swartz was arrested for downloaded large volumes from JSTOR for philosophical reasons (Murphy, 2011). Internet law deals with the old problems of intellectual property, commerce, privacy, free speech and the commons in a completely novel and rapidly changing communicative environment.

1.2: Development as Change - Scholarship in Transition

This study deals with scholarly communication in a transitional period. The primary driver of the transition is the shift in the medium from print to digital. The dissemination model in the print era was the sale of copyrighted works by publishers and that model exists in the digital age alongside open access (OA). OA is based on free access to digital articles on the world wide web. Early on in the short history of OA advocacy, OA leader Steven Harnad put forward a prescient scenario of the global system scholarly communication as follows:

- “The entire full-text refereed corpus online*
- *On every researcher's desktop, everywhere*
 - *24 hours a day*
 - *All papers citation-interlinked*
 - *Fully searchable, navigable, retrievable*
 - *For free, for all, forever”*

(Harnad, 2003)

This vision is made more bold by Harnad's opinion that it is inevitable, if we agree with the opinion then OA debates are more about the pace of change and how we adapt to it. This work further argues and presents evidence that we are currently on the path of fulfilling the 'Harnad prophecy' of universal OA, suggesting that dramatic statements about change accelerated by technology cannot easily be dismissed.

What we have today is a mixed model where Open Access and subscription-based articles co-exist, and with self-archiving they co-exist simultaneously in different versions of the same article. Increasingly, this co-existence is mandated by the patron of the research. Open Access is becoming the normal way for authors to disseminate their work.

Finally, I present evidence here for globalization of scholarship. Digital mediums and open access policies potentiate a shift from Western dominance to global participation in scholarship starting at the dawn of the 21st century. Taken together we are moving in the direction of 'global open knowledge', and certainly openness suggests that it extends beyond the borders of scholarship. It is timely to propose 'global open scholarship' and premature to define too precisely what that will look like.

Access type	Toll	Open
Content	Print	Digital
Dissemination	Western knowledge	Global knowledge

Table 2: Comparison of toll and open and open access.

The sections below introduce Open Access to development and globalization studies by way of basic definitions, and vice-versa. The definitions also serve to clarify the scope and approach of this work.

1.3: Definitions – the peer-reviewed scholarly article

The primary focus of analysis for this thesis is the peer-reviewed scholarly article. Disseminated as widely as possible, journal articles transmit knowledge rooted in basic, original research across social and geographic space within and across disciplines, and to the public. The knowledge is transformed and transmitted over time, validated, challenged and built upon. The works are cited, creating a lineage of written, scholarly knowledge. In the digital era, the article more so than the journal volume becomes the 'molecular unit of scholarly communication' (Björk, 2008).

The *peer-reviewed scholarly article* thus far can be equated to the *journal article* – an article published in a peer-reviewed scholarly journal. For this study, journals are useful for the measurement of the universe of articles. However, this does not preclude the possibility of the peer-reviewed scholarly article existing independently of the journal in the digital era.

In sum, the definition of the article here is that it is peer-reviewed, academic, scholarly and is in the format of an article rather than a thesis, book or other. The form of the article has changed and will continue to change without losing essential characteristics of peer-review and the lineage of citation. Articles therefore can be discussed coherently as the building blocks of a 'self-organizing system of scholarly communication' (Mabe and Amin, 2001, 2000), a concept important to the discussion here.

1.4: Open Access Definitions:

Open Access refers to the elimination of price barriers to the digital article and the reduction of permission barriers ideally to the widest range of legitimate academic use. It has been defined most specifically by the “BBB definition” – referring to the consensus developed at through the Budapest, Berlin and Bethesda declarations. The BBB definition reflects the greatest consensus among OA advocates and refers to the “*removal of all price barriers and all permission barriers relevant to legitimate academic use*” (Suber, 2004).

However, by attempting to capture OA ideals, such a definition can be too exclusive. Further elaboration has been made, though the matter is not yet settled (Suber, 2004). Suber (2008) suggests the terms 'gratis and libre', borrowed these terms from the Open Source Software movement. Gratis OA permissions are equivalent to those of copyrighted works. Libre OA is OA that is licensed in a way that extends re-use permission barriers beyond what is allowed by copyright. Libre and Gratis OA describe the outcome of efforts at reducing price and permission barriers to articles. BBB is a subset of libre OA that offers more specific guidance on creating policies that protect author's interests in attribution and impact and create the most ideal situation for access and preservation of the articles.

The minimum of gratis OA captures the widest range of articles that are accessible without price barriers. Toll Access (TA) is now a commonly-used term in OA discussions. TA that applies to all articles where access depends on a fee, at the individual or institutional level and it still applies if that article is free as part of a donor program, but is not accessible free for all on the web. Taken together, gratis OA and TA capture all articles in a binary fashion. TA requires a transaction to get access whereas gratis OA requires only access to the web.

Gratis OA: There are no price barriers, the article is available free of charge to download and use but permissions are not extended beyond copyright.

- Libre OA: The article is available for free and is licensed to reduce permission barriers beyond what is allowed by copyright. Attribution rights to the author are protected. .
- BBB: Budapest¹ - "By "open access" to this literature, we mean its free availability on the public internet, permitting any users to read, download, copy, distribute, print, search, or link to the full texts of these articles, crawl them for indexing, pass them as data to software, or use them for any other lawful purpose, without financial, legal, or technical barriers other than those inseparable from gaining access to the internet itself. The only constraint on reproduction and distribution, and the only role for copyright in this domain, should be to give authors control over the integrity of their work and the right to be properly acknowledged and cited."
- Bethesda² and Berlin³ - a work is OA where there is no price barrier and users may "copy, use, distribute, transmit and display the work publicly and to make and distribute derivative works, in any digital medium for any responsible purpose, subject to proper attribution of authorship".
- Toll Access: The traditional model for distribution peer-reviewed articles. Toll Access includes all pay for access arrangements. Journal content access can be purchased by institutional or individual subscription, purchase of a journal volume, or purchase of a single article. Copyright may be retained by the author or transferred to a publisher.

In all cases, the scope of Open Access refers to peer-reviewed scholarly articles. By definition, OA does not imply an difference in content between an OA and a TA article.

Authors may also facilitate OA by archiving versions of their articles in Open Access repositories, including their final refereed drafts. 95% of TA journals and 60% of publishers allow some form of author self-archiving (Open Oasis, 2011), but it is not universal and publishers vary on the type of draft they allow as well as whether and how long of an embargo period. The table below summarizes the basic routes to accessing journal articles and every version of an article should be captured by the distinction:

¹ Budapest Open Access Initiative - <http://www.soros.org/openaccess/index.shtml%20>

² Bethesda Statement on Open Access - <http://www.earlham.edu/~peters/fos/bethesda.htm>

³ Berlin Declaration on Open Access - http://www.zim.mpg.de/openaccess-berlin/berlin_declaration.pdf

Green OA: Article becomes available for free on the internet as a result of the author archiving it in an open-access repository.

Gold OA: A Journal publishes the article available free of charge on the internet.

1.4.1: OA publication charges – a distinction left out of the definition

Some Gold OA publishers have shifted their revenue model from charging readers to charging authors, creating a price barrier at the author end. This barrier appears quite significant for many authors (Schroter, 2004), and particularly those where price barriers are the key problem. For this reason, I introduce the term *OA 360°* to describe the situation where there is no price barrier at the author end or the reader end. Publication charges are for many authors a reason not to publish in a journal that charges them (Swan and Brown, 2005, pg. 12, Rowlands, Nicholson and Huntingdon, 2004, pg. 29). Thus, they represent a price barrier not included in OA definitions which are concerned with barriers to users. The price barrier can be an indirect barrier to the reader however, as the author may publish the article in a TA journal without archiving it or withdraw the article. As a final note, *OA 360°* already applies to 70% of OA journals (Suber, 2004) and since there is no charge to deposit in OA archives, it applies to all of green OA. We will return to this in the discussion following the empirical questions and results.

1.4.2: Operational Definitions versus Discussion and Concepts

The gratis/libre distinction creates problems for the empirical focus of this thesis, which is based on determining the trend in global OA shares. If we are chiefly interested in basic availability of price-free articles, libre OA is too high a requirement and ends up excluding some of these articles. *OA 360°* is also too narrow if we are chiefly interested in the outcome of free access for users. In conducting the studies most similar to this work, Björk et. Al (2008 and 2010) reported it was impractical to distinguish libre from gratis OA for the purpose of measuring OA shares. In addition, we would lose the significance of the key distinction of gratis versus toll (the primary OA-TA distinction for consumers). An operational definition based on the gratis OA definition can more easily result in a measurement of the cost-free availability of articles. Therefore, we can distinguish between the operational definition of OA here as gratis, and the discussion of other issues in defining Open Access.

Relevant to this work, there are also development-oriented journal access programs specific to developing countries. Over the last decade, a significant level of publishers have participated in access programs where participating institutions gain free access to subscription content (chiefly the HINARI, AGORA and OARE programs to be discussed). These are not without controversy but appear to be significant

efforts at addressing price barriers. However, they are not considered OA and fall outside the most inclusive definition. This donated content is not available to someone who is not a member of an institution, and the programs are not without costs, not without the possibility of future subscription costs, not without the possibility of removal of privileges.

In the exploratory phase of research, I had hopes of developing an 'Access Index' for libraries which would add an institution's access to TA journals to the current global OA share and subtract the overlap. This would tell a librarian where they stood in providing patrons' comprehensive access to research. For instance a .95 index would mean access to 95% of global literature. This could be given for annual or longer periods, or as a portion of all journal literature that exists. The 'open' version of that index applies to the condition of zero privilege to TA literature is precisely what we are studying here.

Thus, there is a continuum of access options with regard to battling the key economic problem related to development here – price barriers. Permission barriers are not unimportant to development, but one does not even encounter the problem if the article is not available in the first place. The most significant barrier in the question of access to research for developing countries has been the exclusion for basic availability to the articles. Greater inclusion of the South would be positive for future discussions on how to define and create open scholarship, the following table summarizes where we are now:

Libre OA Libre OA 360° (no author charge)	The most open, and considered by all to be OA. Can promote greater offline availability, most open in terms of permissions.
Gratis OA Gratis OA 360° (no author charge)	Available for free to all with internet access. Permission barriers may limit re-use.
Toll Access made available for free through donation programs.	Donor-recipient model. Recipients are members of eligible institutions in the South to whom free access to participating titles is granted. Access to the internet, membership to the institution and a working log-in ID and password, a working log-in page and remote access if the patron is off-campus, are all required above access to the internet. Eligibility can shift incurring a small fee or making country institutions ineligible if the nation rises into a different income tier. If everything works according to plan, the article accessed free of charge similar to gratis OA.
Toll Access at a bargain (eg EIFL ⁴).	INASP and eIFL have developed consortia and national-level bargaining that vastly reduces the price of literature, publishers are more willing to negotiate access to digital literature at prices that reflect purchasing power in poorer nations. INASP for instance, boasts a broad selection of titles (30,000) and e-books (8,000) that institutions can choose from. In addition, regional Journals Online programs such as African Journals Online (AJOL) contain OA and TA journals, but the TA journals are delivered free of charge within the region (INASP website).

Table 3: Definitions of different access types.

In summary, the primary concern here is the basic availability of articles and this is the empirical focus. OA is the most direct route to this and OA advocates agree that libre/BBB OA is the most ideal condition for an articles ongoing accessibility, impact and usefulness, but OA is not the only strategy taken to address price barriers. For the sake of the overall discussion in development, we are concerned with the continuum.

⁴EIFL (Electronic Information for Libraries) and INASP (International Network for the Availability of Scientific Publications).

With regard to the empirical questions and operational definitions, gratis is sufficient for the article to be included in 'shares of OA' – the percentage of literature that is freely accessible on the web.

Having made clear distinctions, we can measure OA globally, and we can also look at concession and bargaining programs introduced as development programs and made possible by the internet. The continuum reflects the need to adapt our picture of research access to a development context.

1.5: Definitions and Distinctions - Globalization? Development?

It is necessary to articulate what the words globalization and development refer to here, and what they don't. The specifics of Open Access terminology may be unfamiliar to those involved in the study of globalization and international development and vice-versa.

1.5.1: Development

This thesis posits Open Access as a question for development. A study of the word development can be reduced simply and semantically as observing change in a system. Development is change, but change can refer to a discrete event, whereas development refers to a process. Here we are interested in human social systems – this applies to definitions 'human' and 'international' development (ID) (AND NOT 'child development' etc.).

<p>Human development</p> <p>The concern is with trends over time for all of humanity, and distinctions across nations. A human development approach would reveal various disparities, though the North-South divide is not necessarily the intentional focus. Human development is more focused on observing change in various indicators on the whole, than in analysing the impact of specific programs and policies.</p>	<p>International development –</p> <p>The concern is with poverty alleviation and advancing a more equitable world. ID is interested issues arising from North-South income divide and the relationship of cooperation or perhaps conflict between rich and poor regions and nations. A new term 'South-South', referring to South-South cooperation may also fit the definition. ID is interested in studying the impact of programs and policies and changes in these.</p>
<p>Global view</p> <p>concerned with trends across time for all of humanity on key development indicators.</p>	<p>International view</p> <p>traditionally ID is interested in North-South problems, and the disparity between rich and poor regions. The global North is a term that captures the advanced developed countries of Europe, North America and Australia New Zealand. The global South encapsulates the rest of the world.</p>
<p>National and Regional</p> <p>concerned with trends across time for key development indicators within geographic boundaries.</p>	<p>ID</p> <p>approaches within regions/nations – concerned with disparity between groups, focused on poverty alleviation and problems of marginalized groups.</p>

Table 4: Comparison of various terms.

An approach from human development may be to study changes in the patterns of research output across nations in the same way we track child mortality or income changes. An approach from international development moves towards action -policy and programs- and may study the impact of OA policies on development outcomes, and to focus on North-South or within-country, within-region disparities, on equity issues across identifiable groups. As with most research, both views are important to the actual work and it a question of emphasis. This particular work can be taken both ways, but the focus is on very broad global changes and this is a very challenging task already. Thus, many of the thorny issues of development and its language cannot be dealt with properly here. The North-South distinction clearly is an over-simplification, for instance.

Equity of participation in development is a common theme of development research. A positive development effect would be achieved if OA provides a greater space of participation among those previously excluded or limited in the power and capacity to determine development for themselves and their communities and nations, and particularly when the root cause is lack of access to information. The research here is focused on describing change in scholarly communication quantitatively and offering some tentative suggestions about implications for development, global studies, implications on equity, social justice. These can inform OA policies and programs, and those of development, and especially those that link the two.

1.5.2: Globalization

The changes occurring with scholarly communication intersect with the trends of globalization, which has a long and short history. The results of the empirical research are a demonstration of globalization in a specific sphere. There is accelerated change in the last decade coinciding with the rise of digital communication -theorized to be a key driver of contemporary globalization.

Globalization is a useful concept in interpreting the data here and building on previous periodizations by Mabe (2003) and Price (1963). The hypothesis is that we have entered into the start of a 'global' period of science/scholarly communication that is moving toward an 'open' period. However, the world globalization has unhelpful political connotations that are not intended by its use here. Given various popular uses of globalization and the anti-globalization movement, a further distinction is made. The use of term globalization here will refer to the 'academic study of', and does not refer to its other uses. A good summary of this use and the distinctions is found in *Globalization: a Critical Introduction* by Scholte (2005), who distinguishes globalization by the characteristic "supra-territorial". However, I would argue that it would be premature to set the definition if we anticipate broader participation going forward.

To provide a starting point for analysis, let us suggest that globalization describes two worlds along a historical continuum. Globalization is a continuum and we are more global today than in the past. Unglobalized, or less globalized social worlds once existed with little or no relation to one another. For instance, 15th century Europe was unaware of civilization in North America. Many still thought the world to be flat, so there was not broad awareness of a globe (let alone globalization). Yet, trade and expansion had already expanded the social world of Europeans through contact with the South and East, so it was still in the process of globalization. Classical science in Egypt, Greece and Middle East had made strides towards the discovery of the Earth as a sphere orbiting the sun long before the Renaissance as a result of knowledge shared across trade routes, the discovery of the globe itself being a step towards globalization and an effect of it.

With globalization, our social worlds are increasingly in contact, in communication and economic and socio-cultural relation with one another. This forces inevitable systematic integration and negotiation of boundaries. Globalizing forces are cyclical in a relative sense, and linear in an absolute sense. For instance, international trade flows *per capita* are similar before WWI and after WWII, and lower in between (Scholte, 2005). Yet, the post-WWII world is far more globalized than the pre-WWI world as a result of advances in science and technology, particularly in communication and transportation, and continued expansion of civilizations.

In summary, I am defining development by its use in the terms 'human development' and 'international development'. I am using the term globalization in its academic and conceptual meanings, referring to Scholte (2005), who is among few who have consolidated an overview of globalization from a social studies perspective.

1.6: ICT4D and K4D not included

Before we go on, some notes on the history of the research proposal are necessary. The original proposal grew out of discussions about literature access in Africa. The initial proposal was based on a model where an entire set of indicators and conceptual development were devoted to 'Information and Communication Technologies for Development' which was concerned with the state of the 'pre-requisite' to OA in internet access, and another towards 'Knowledge for Development' concerned with OA's development impact. However, the original proposal was too ambitious and had to be narrowed. The front and back was dropped and access to research became the focus. The original model developed aside from the thesis project in collaboration with Dr. Diack, as the Development Equation (DEq) and is a work-in-progress online on WikiEducator under 'Learning and Knowledge in the 21st Century'. Still ICT bears mentioning as more relevant to Open Access debates in developing regions than in the North. K4D remains as a link to more general development impact of knowledge access.

There remains a disparity among rich and poor regions in access to computers, personnel, infrastructure and resources, but one that is narrowing. There remains a digital divide, but one that less likely to be absolute. Beyond the university community, mobile access is growing much faster than internet penetration (ITU, 2011). Mobile Web for Development (MW4D) (W3C, 2011). is another significant topic for research. The expectation is that for university communities, information and knowledge are the core resources and the web is indispensable to them. For researchers interested in ICT4D and MW4D, the results here demonstrating a rapid increase in digital availability of scholars represent change in the nature of content delivered as a result of successful work against the digital divide. Such research in turn informs open access policy-making by answering the question of whether network access can be assumed in South and in rural/remote regions – impact lost or gained.

The International Telecommunications Union (ITU) reports annually on the state of ICT Development (ITU, 2011). Colle makes compelling arguments for innovative approaches to ICT development in university contexts in Africa (Colle and Roman, 2003, Colle, 2005). ICT diffusion has levelled off in the North and is growing rapidly in the South (Clapshaw, 2007). Clapshaw's extensive dissertation provides evidence that ICT diffusion is highly correlated with three main variables that account for 85% of the variance – teledensity, corruption, and higher education enrolment. Only one of these correlates with characteristics of the user, demonstrating the strength of demand at universities. At the national level, the data confirms suspicion that cumbersome telecom policy and government corruption remain obstacles. However, this is where the university itself can play a healthy role in providing evidence and stimulus for change. In an IDRC discussion paper Open ICT4D, Smith et. Al have linked network access opportunity to openness in content (Smith, 2008).

I can report with some optimism, good experience with internet access in the city of Dakar and at the Universite Cheikh Anta Diop in the summer of 2008, and humbling encounters with students developing computer applications there. Still Dakar and UCAD are both in a good position comparatively in Africa. Dakar is a port city has better access to undersea cables than landlocked countries, Senegal is a country without violent conflict and UCAD is a top university in Africa.

The link between OA and ICT development, and the links between the digital revolution and the openness paradigm are critical to development. The focus here will be on the implications of the growing share of Open Access, and it is anticipated that conditions for network access will not be ignored by others and will continue to improve across the South. As of now, we can assume that though it may be easier and cheaper for some, the vast majority people in the world that are part of a university community are able to gain access to an OA publication, and that is a simple enough starting point for the work ahead.

Chapter 2. Rationale, Context and Theory – Literature Review

Price barriers began to constrain largely publicly-funded library budgets in the 'North' (advanced developed countries) in the 1990s. According to the European Commission (2006) study on the technical evolution of scientific publishing markets, the crisis developed towards the 1990, by which time journal prices had risen 300% beyond the rate of inflation. Libraries had formed into consortia to increase their bargaining power, but many researchers remained dissatisfied with access lost. Funders were concerned about loss of impact on investments, “and that scientific progress is retarded through inadequate access to related research conducted globally.” (European Commission, 2006).

The E.U. study's first recommendation was that government funding agencies should mandate that the research they fund be made available without charge on the internet. The recommendation became a campaign and is supported by a 27,000 strong petition from the global research community including many high-profile scientists and several Nobel Prize winners. This recommendation is accompanied by several declarations that have defined the open access movement. The Budapest Open Access Declaration of 2001, became the formative document for the OA movement and the Budapest Open Access Initiative, which continues to be a leader in OA advocacy. The BOAI was followed by the Bethesda (2003) and Berlin (2003) Declarations, and several others. For a complete list see the document "Declaration in Support of OA"⁵.

To give focus to the OA-development link, we give focus to Subsaharan African and specifically development as it relates to academia in the region.

2.1: Development Context - Universities in Subsaharan Africa

During the 1960s investment in education at all levels was a main focus of national and international development throughout the world. Canada greatly expanded post-secondary education between 1964 and 1971 (Banting, 1995). As early as 1945, Harvard president James Conant argued that a strong system of higher education was essential to American democracy, social mobility and egalitarianism (Conant, 1945). Thomas Schultz championed a 'human capital revolution' in the approach to development economics through the 1960s (Schultz, 1960). During the independence era in 1966, Nigerian author Dowuona discussed a cycle beginning with the weakness of higher education translating into a lack of trained post-primary teachers culminating in a vacuum of skilled personnel to support the few highly educated Nigerians running the country (Dowuona, 1966). Dowuona saw a critical role for universities in self-

⁵http://oad.simmons.edu/oadwiki/Declarations_in_support_of_OA.

determination stating that “until the developing countries build up the right foundations and tradition for training to the highest level, they will not be able to be masters of their own destiny.” In 1968, U.K. statesman Sir Andrew Cohen placed education as the highest priority in African development, and universities as the first concern. He saw a critical role for institutional capacity-building through some 50 arrangements underway between British and mostly African universities – giving the example of the cooperation of medical schools in Nairobi and Glasgow (Cohen, 1968).

Responding to this push for higher education in Africa, international aid and national efforts pushed allocations for education to over 1/5 of national budgets in some developing countries (Lulat, 1982). However, the scale and speed of these initial efforts outpaced development in general, and the expected returns in development outcomes did not materialize. Human capital theory was sidelined in the 1970s when ‘educational reform’ became the development buzz-word. Reforms tried to shift the focus towards basic needs for the rural majority, strategic investment in quality, introduction of technology and nonformal education. However, reforms were not carried through extensively owing to lack of resources and political will (Lulat, 1982).

Following the initial push, the development of tertiary education and research in Africa has been neglected (Birdsall 1996, Bloom, Canning and Chan, 2005). Spending on higher education has sometimes been seen by donors and policy-makers to favour elites and exacerbate inequality (Lulat 1982, Birdsall 1996, Kapur and Crowley 2008). However, some have seen tertiary education not as a distant goal to come after primary education has been dealt with, but as the root for effective national development including the effectiveness of primary education (Schulz, 1960, Dowuona 1966, Cohen 1968, Lulat, 1982, Stiglitz, 2000, Colle, 2003, Kapur and Crowley, 2008) and at least one author proposes to focus on the quaternary level with graduate training and research (Birdsall, 1996).

By the turn of this century, SubSaharan African nations had not by and large realized the vision for higher education set out by Cohen, Dowuona and others. As an example, Zambia followed the normative pattern of ambitious investment in the sixties and ambitious reform in the seventies. Zambia’s copper crisis in 1975 led to austerity measures and a focus on more urgent needs. Zambia’s failure in reform was also linked to the mismatch of ideology and political will (Lulat, 1982). More dramatically, at the Juba Health Institute in Southern Sudan, progress in medical and nursing education was halted by a two decade long civil war where the entire academic staff was in exile, emerging only recently with the peace accord signed in 2005. Various patterns of conflict, fiscal crises, ideological constraints, Cold War interference and other circumstances have held back the development of universities in the South and the focus on higher education.

A handful of influential authors have noted the significance of the infrastructure of knowledge creation and dissemination itself to the process of development. A leading voice on knowledge for development is Nobel Prize winner in Economics and former World Bank president Joseph Stiglitz, who stated the following in a speech from 2000:

“If the developing countries are really to be “in the drivers’ seat” they have to have the capacity to analyze the often difficult economic issues which they face. Local researchers, combining the knowledge of local conditions—including knowledge of local political and social structures—with the learning derived from global experiences, provides the best prospects for deriving policies which are both effective and engender broad-based support. That is why locally-based research institutions are so important.”

Joseph Stiglitz, 2000

This brings us full circle to Dowuona and Cohen, who were both concerned with capability and self-determination through higher education and research, and the same sentiment is reflected in Stiglitz’s approach to knowledge in general. Increased access to literature over the past ten years provides new opportunities for realizing these goals, occurring through concession programs (chiefly the UN's HINARI, AGORA, OARE), new organizations and a better bargaining position (eIFL and Perii-INASP), regional free access (AJOL free to African institutions) and Open Access – all of which have arisen recently as a result of the shift to digital.

The OA-development link is made stronger by demonstrating how it is situated in theories of economic change and development.

2.2: Review of Key Theoretical Literature - Development as Economic Change

“History bears witness,” writes a sociologist, “to the cataclysmic effect on society of inventions of new media for the transmission of information among persons.”

N. St. John (1967)

Three seminal works and concepts from development economists are useful to provide a framework for open access advocates and librarians to make a link to development theory, and for development people to work from to understand open access as a problem related to it. Despite the difficulty in fitting this in in the theoretical section, the legacy of Canadians Harold Innis and Marshall McLuhan must be mentioned for their contributions in communications, in this case 'the medium is the message' is central, concise and obvious enough to simply state. The effect is so

ubiquitous now, particularly in the North, that simple reflection on our own lives over the last 20 years is a clear demonstration of the social and economic impact of technological progress in communications and change in medium. From speech to writing, the printing press, wired communications, telephony, radio, television, digital – all are economic revolutions – yet for scholarship the printing press and the internet are the most significant. With that in mind, I will focus on three contemporary Nobel laureates whose work in economics and development as poverty alleviation is relevant:

4. The concept of Development as Freedom from the work of the same title by Amartya Sen.
5. The concept of Knowledge as a Global Public Good from work of the same title by Joseph Stiglitz.
6. The concepts of path dependence, from evolutionary approach to economic change (related to the N.I.E. School - New Institutionalism) as referred to in the work Understanding the Process of Economic Change by Douglass North.

The concepts these works are based upon are reduced here to the most basic argument relevant to this thesis.

2.2.1: Development as Freedom – Amartya Sen

Development as Freedom is a seminal work of economist Amartya Sen who received the Nobel prize in Economics for his work on social choice. The core idea in Sen's work is that development should be evaluated in terms of “the expansion of the ‘capabilities’ of people to lead the kind of lives they value—and have reason to value”, this is Sen's definition of freedom.

Access to research can be analyzed with regard to the capability concept in two ways. The first is with regard to all of the people for whom accessing research is essential as a capability that allows them this freedom. Development can be evaluated then, by the extent to which there is disparity in this particular capability among people in developed and developing nations. Insofar as there is deprivation of this capability, there is 'unfreedom' and insofar as it is expanded there is greater freedom and hence greater development, in this case for those who value access to research.

The second way in which capability can be analyzed is with regard to the role in society of individuals who work with knowledge ('knowledge workers'), and the way in which they are engaged in *development* as it is understood in this model, in the expansion of capabilities for the people they serve. Sen's logic requires us to understand development as a process of expanding capabilities, and as outcome where we can evaluate it's success in this. Insofar as we gain freedoms on the basis of relations with workers who serve to empower our capability, their freedom relates to our freedom. This makes it possible for intentional development programs, as well as the consequences for development of any policy to be evaluated against a criteria with a universal ethical dimension of social choice. This is one that is particularly visceral

when confronted with its denial, and can be understood by people who are meant to benefit from development. It is general enough to apply to many situations, including OA.

The significance of Sen's contribution is in contrast with conventional and dominant ways of evaluating development, focused on particular results such as economic growth as measured by GDP, without any external indicator that would give such results meaning, and the GDP measure itself does not reveal root causes of underdevelopment. Capability, on the other hand, is an inherent characteristic of people and is expanded or limited by conditions they live under, conditions related to what we understand as development. Assuming published research is valued by knowledge workers, the freedom to access that research enhance the ability for those workers to express their occupation role, which contribute to the ability of the general population to lead the kind of lives they value and have reason to value.

2.2.2: Knowledge as a Global Public Good – Joseph Stiglitz

Knowledge as a Global Public Good (Stiglitz, 1999) is a chapter in an Oxford/UNDP publication 'Global Public Goods' edited by leading development social scientists. Stiglitz won the Nobel Prize for his work market distortions brought about by asymmetric information. In his words, the problem can be summarized as 'some people know more than others.'

A public good is something that people cannot acquire through market mechanisms but nevertheless is required or benefits them. An example is a piece of infrastructure such as a traffic light or a bridge. People benefit from these in various ways, but a traffic light or bridge is a good that is available through non-market mechanisms and is best defined by how it is consumed, as non-excludable and non-rivalrous:

“if no one can be barred from consuming the good, then it is nonexcludable. If it can be consumed by many without becoming depleted, then it is nonrival in consumption.”

When articles become digital, they increasingly come closer to the qualities described here. Further, something is a global public good when this logic applies everywhere, that is the good is non-rivalrous and non-excludable regardless of location and nationality, they are global and open.

The relevance to policy and development is that certain goods are best understood as public goods, and certain of these as global public goods, and that many goods have qualities of public goods. These qualities are inherent in the goods and these goods are foundational to functioning economies. Policies can result in the under-production of such goods, and under-production would impact negatively on the diffuse externalities that these goods create. Such policies would be generally bad for

economies. The application of this is not esoteric, one can easily see how under-production of a public good such as a bridge would impede all manner of economic, social and health activity. This is quite fundamental, though clearly not easily achieved in all cases and quite possibly not achieved as a result of negligence to the production of public goods in policy. The authors of this volume are putting forward the concept of global public goods so that policy-makers can attend to this fundamental and the concept of national public goods can be understood in a global context. The state has a key role in the provision of public goods. National public goods are defined with reference to national economies, but they translate well and are informed by an understanding of global public goods.

Stiglitz argues that knowledge is a global public good. He begins with a quote from Thomas Jefferson:

“he who receives an idea from me, receives instruction himself without lessening mine; as he who lights his taper at mine, receives light without darkening me.”

This quote describes a key feature of knowledge, that it is non-rivalrous, the consumption of one individual does not detract from another. He goes on to establish that knowledge is non-excludable in general (only through intellectual property regulations can knowledge be made partially excludable). Anyone who can obtain knowledge, in general cannot be excluded from the benefits of it. His analysis provides the key economic problem by which the access to research dilemma arises. Knowledge itself is a good that is naturally free for all to use despite any charges associated with the work of transmission and limits on commercial application with regard to patents. Since these aspects of knowledge can be appropriated through intellectual property, knowledge is often thought of as an *impure* public good.

Knowledge is a public good, but it is also a determinant of how public goods are manifested. The assessment of whether a bridge should be prioritized, how to obtain resources, how to pay for the bridge, what materials to use and how to build and maintain a bridge, and the implications of the bridge for social and economic activity are all types of knowledge and information necessary for the evaluation and production of the bridge. Those who have the knowledge are in the 'driver's seat'.

The role of the state in the provision of knowledge as a public good has two expressions. One is the extent to which the state establishes a regime for protecting appropriations of knowledge, an intellectual property regime. The second is direct support. Stiglitz introduces two important concepts with regard to balancing such a regime:

Dynamic efficiency: how well an incentive works to stimulate the development of knowledge and knowledge products.

Static efficiency: the ease by which people can access the knowledge and its products.

Strengthening intellectual property grants greater monopoly to the owner, a strong incentive to produce, but it comes at a trade-off for static efficiency. Stiglitz states that most economists agree that providing the incentive through an intellectual property regime is a better approach for inventions, since it provides a better selection mechanism than a bureaucracy choosing what to support directly. This is a clear problem for the trade-off in the incentive to research new drugs, and the cost associated with the monopoly that can deprive sick people of the benefits. Stiglitz points out that the situation is entirely different with regard to basic research.

“But there are some important situations where the costs of the improved appropriability strategy are high. This is particularly the case for basic research because its benefits are widespread and diffuse, and because attempts to appropriate its returns may significantly slow the overall pace of innovation.”

Stiglitz, 2000

Stiglitz's statement is that basic research ought not to be appropriated. He does not address copyright in this discussion, but his analysis has implications for differentiating basic research that is copyrighted for the sake of the incentive to publish, and research that is patented for the sake of discovery. The world wide web itself, and much of the digital innovation has been based on 'open' knowledge for software and systems that has been necessary because of the static efficiency required for the www to function, and 'open source' has developed many alternatives to patented software products. With regard to the print model of publication, the charge to consumers (usually paid on their behalf by libraries) and the support from the state in the form of copyright is justified by dynamic efficiency providing the incentive for the work of publishing but it creates an impediment for static inefficiency.

With regard to the electronic model, the retention of the same balance is not a foregone conclusion. The question is whether a higher pay-off exists for the funders of basic research in ensuring it is disseminated without charge, than in granting appropriability to the results of research to publishers. The incentive model for funded basic research is direct support, whereas in the print era the incentive for dissemination was copyright. The maximum static efficiency exists when there is no fee for use, as in the case of the traffic light and the product is expressed as a pure public good. Where an article is freely available globally to access and copy, this form of knowledge is available as a pure global public good. Policy-makers are not unaware of this changing incentive structure. The direct result has been funder and institutional mandates for open access archiving and support for open access publishing resulting in

the Green and Gold routes to OA. As funders and institutions begin to mandate open access, a key problem is whether the trade-off in dynamic efficiency in copyright protection for publishers will threaten journals viability. The question is critical for libraries, whose concerns over inflation of subscription prices – known as the 'serials crisis' – was a key driver for open access. The inflationary problem is an example of the balance shifted too far in favour of dynamic efficiency. The shift to electronic publishing provides new possibilities. The theory of global public goods provides strong economic rationale for open access. However, transplanting the copyright model for print to electronic remains the model for the majority of journal publishing today. There remains 'path dependance' – and overcoming it requires adaptation.

2.2.3: Path Dependence in Economic Change – Douglass North.

Path dependence is a key concept in “Understanding the Process of Economic Change” (North, 2005). Nobel laureate North analyzes societal economic successes and failures through history to identify what went right or wrong in terms of the structure. Path dependence is a critical concern for North. Our current structures are based conditions that existed when they were erected. Since conditions change with novel situations, they may not necessarily apply the same way to new conditions. Adaptive efficiency is the degree to institutions can adapt to change by allowing for trial and error and risk and being open to change based on new knowledge. Institutions discriminate in their responsiveness to various actors, and some political and economic entrepreneurs may have more influence on institutions. They can be held rigid by vested interests created under incentives meant to deal with past situations.

It is important to note that North's perspective on economic change is one which views classical economics within a distinct evolutionary analysis on how humans erect such structures. The rationality assumption central to classical economics is seen to be limited by the function of consciousness. Beliefs, whether rational or non-rational, shape institutions which shape organizations' policies which shape outcomes. Institutions are defined as the set of formal and informal rules that humans erect based on evolutionary cultural experience (this use of institutions is distinct from 'organizations' – in North's verbiage, University of Ottawa would be an 'organization'.)

Where institutions are failing to adapt due to path dependence, human's can intentionally produce change through political and economic entrepreneurship. However, imperfect information, erroneous beliefs, and interests that derive from the power structure can limit change and produce unsuccessful outcomes. North identifies three important factors governing how humans shape institutions:

1. the demographic characteristics of society
2. the stock of knowledge that a society can rely
3. the deliberate incentive structure as defined by the nature of institutions.

These interact dynamically, however, institutional structure is seen as central, and can shape the demography and the stock of knowledge.

In terms of adapting to changes brought about by the communications revolution, with regard to 'the academy' as the relevant institution we can analyze the evolution of scholarly publishing as the case in point. Open Access is a form of political and economic entrepreneurship.

1) **Demographic characteristics:**

The demographic characteristics are significant and undergoing change. There is a generation entering adulthood who has grown up with digital technologies, the digital divide is falling and the audience for internet publications is globalizing, with the younger generation dominating outside the West. In the West, where scholarly literature has been centralized since its inception, those making decisions over the structure of publishing are part of the baby boom, a large demographic that grew up in the print era.

2) **Stock of knowledge:**

Open Access in fact deals with the stock of knowledge and its globalization in general, creating the self-reflexivity problem in the study of it. But within that stock of knowledge is what we know about publishing, which is based on our longer experience with print and our recent but quickly developing experience with internet.

3) **The deliberate incentive structure:** Here we have the incentives provided to do research, including the writing up of research articles which are provided by research funders, largely public. Then there is the structure of copyright and purchase of the scholarship between researchers, an incentive structure to provide the funds for dissemination of the research.

As the publishing process changes from print to internet, the question is whether the incentive structure continues to make sense. However, even if a changed structure makes more sense, the adaptation is hampered by path dependence on the old model. For many advocates of open access, the evidence and logic for an open system appear overwhelming. Within Douglass North's framework for analysis, resistance to open access may be viewed as a case of path dependence on the print model.

2.3: OA Arguments

2.3.1: The Argument from Principle

The primary argument for OA is a principle that is rooted in the values of the scholarly tradition. This principle is best expressed by John Willinsky (2006):

“A commitment to the value and quality of research carries with it a responsibility to extend the circulation of such work as far as possible and ideally to all who are interested in it and all who might profit”.
(Willinsky, 2006)

It is notable that the Access principle is a perennial one, not a result of digital technologies. However, in the context of the digital revolution and the world wide web, the principle becomes far more expansive and the ideal attainable. Interestingly, in a study by Swan and Brown (2005), the key reason authors who make work their work available through OA type licenses was the access principle. In fact, this principle ranked higher than all other rationale, including more direct self-interest as expressed by citation advantage (Swan and Brown, 2005).

2.3.2: The Economic Argument

The economic arguments for OA came forward in the North independently of development concerns. However, given the key development issue of North-South income disparity the problems OA addresses in equity and efficiency are considerably more acute.

Key features of the economic arguments are:

- 1) The taxpayer argument - As the main patron of research, on the behalf of the public the government is paying to have research done for impact. Impact is highest with widest dissemination (Houghton, 2009), . The argument is also expressed as paying twice, once to have research done, and a second time to buy it back though this expression is more controversial (Harnad, 2007).
- 2) The efficiency and equity of global public goods argument. Basic research is a global public good where price barrier-free conditions to its use is optimal and direct stimulus is better than stimulus through extending intellectual property privileges (static efficiency versus dynamic efficiency) (Stiglitz, 1999). The impact of research investment is improved when the results are made freely available on the internet (EU Commission, 2006). The publishers enjoy market power which allows them to increase prices above cost and inflation and open access is an alternative to a better balance, and mandating open access is recommended for optimal value in impact for research investment (EU Commission, 2006). The dropping of price barriers creates greater equity with regards to the benefits of public investment in research. The case of entire regions like Africa operating in a literature vacuum with consequences that limit economic development can be avoided (Willinsky, 2006).
- 3) Unknown loss of scientific discovery. We do not know the capabilities of all people who face a barrier to research. What benefits are lost when genius is deprived of opportunity because of geographic misfortune? From where will the

next Einstein emerge? What are the economic costs of such deprivation, and what benefits lie ahead from reduction of barriers? (no reference, author's contribution)

2.4: OA in developing countries

In the Access Principle, John Willinsky argues that developing countries have the opportunity to emerge from a 'literature vacuum', having seen decades of cuts to journal subscriptions owing to price barriers. Willinsky cites impacts such as Kenya's KEMRI (medical research institute), and the boon for training and research brought about by programs for free access to participating publishers through HINARI. Two trends provoked the open access movement; a) the escalation of journal prices 4 times faster than inflation (Suger, 2004 in Willinsky, 2006); and b) ICT/digital publishing. Civil society, the public and democracy stand to benefit, and Willinsky cites the use of research material by journalists to critique government programs in education. Arguing that research authors require the body of literature they contribute in order to do so, Willinsky proposes the Access Principle to advocate a principled approach to publishing. The book will cover both practical matters – digital journals, copyright law, publishing economics, journal design and indexing; as well as expansive themes such as the research capacity of developing nations, public rights to knowledge, policy and political contributions of research (Willinsky, 2006). In many ways, Willinsky's rationale parallels the DeQ model focusing on ICT4D, OA and K4D, alluding to the openness paradigm and networked partnerships.

A few key authors globally have made frequent reference to OA as a development issue, and there have been some studies analyzing OA issues, with surveys and data. In a 2005 article Arunachalam and Suber make key arguments for open access in the developing world. It achieves two goals, widening dissemination of author's works and making more research available to them. They discuss the digital divide as a possible impediment, but argue that network access is steadily increasing and that the content divide should be addressed to take full advantage of each step towards better internet access. They argue that developing repositories for archiving has even greater advantages in cost and impact than OA journals (Arunachalam and Suber, 2005). Arunachalam, Kirsop and Chan (2005) argue that "there are both North to South and South to North knowledge gaps, leading to incomplete pictures of global science in general, and global health, environmental and development issues in particular." This article emphasizes self-archiving as the fast track to building research capacity. Willinsky conducted a survey-based study of universities in Cameroon to determine the state of access to research through the lens of students and researchers' perspectives. Willinsky (2005) concludes similarity that digital divide and knowledge gap issues go hand in hand, but reveals that despite weaknesses in network connectivity and computing resources, all universities in Cameroon had internet access and it was improving, while access to research gaps had improved they remained more severe. Computer use often occurred outside home and university, at public and commercial

facilities. Respondents were relatively enthusiastic about the advantages of open access. Willinsky concludes with an emphasis on the *global* circulation of knowledge and with optimism about the potential.

Co-authored by several African authors, Smith et. Al (2007) conducted a study of five African universities on access to biomedical literature and internet use. As with other studies, respondents accessed the internet more typically outside the university, making frequent use of internet cafes. Awareness of free access resources was variable, HINARI was helpful but not unproblematic to access. Ouya and Smart (2006) conducted a survey of OA awareness among editors of African-published journals, receiving 48 replies from 230 questionnaires sent to journals from the African Journals Online (AJOL) database. The authors focus on a global picture and identify a role for North-South partnership as well as successes to be adopted from OA publishing in Brazil and India. The authors conclude that while behind the curve, editors are catching on to OA as an alternative to subscription-based models. They recommend training and resources for electronic OA publishing, practical demonstrations and for a continuing a role for bargaining programs such as INASP-Perii and services such as AJOL which editors report to have been helpful.

Through conferences and presentations and developing the DeQ conceptual model on the web, Moustapha Diack and I have been making the case for 'Open Solutions – Open Source, Open Access and Open Educational Resources (OER) and North-South 'networked partnerships' amongst universities. Development has not been central to either the OA debate, and OA has not been central to the development discourse. It is hoped that this work contributes to the discussion.

2.5: Stakeholders – Varying Actors and Interests

There are somewhat different concerns for patrons (researchers as authors and users, students, the public), libraries, publishers, research funders, institutional administrators (higher education and research) and people interested in the study of scholarship as a system (information scientists, systems theorists, historians, international development etc.). Interests and viewpoints around digital literature and open access vary between these groups and within them, the roles frequently overlap, and the situation varies in different parts of the world. This section is an independent analysis, though influenced by a similar analysis by Willinsky in the Access Principle (2006).

For patrons, the ideal is the most comprehensive availability of quality literature and an effective way to retrieve it without stopping to transact a fee in the process. Who are these patrons? Traditionally, they are members of academic communities who obtained through memberships to institutional libraries. Potentially, a patron is any person who can benefit from directly accessing scholarly literature. It remains true that a research article need not have broad public appeal, and the public and even non-specialist scholars may not even have the capacity to understand or apply the research.

However, the free availability allows for the participation to be dictated by choice and capability rather than membership to an institutional library that is able to afford the subscription.

Librarians act on patron's behalf to provide a broad source of literature. Whereas the concept of a university, is an institution that potentially covers any type of inquiry, a university librarian truly pursues an ideal of comprehensive access for patrons. In the context of toll-access and scarce resources, there are some limitations. Despite these, universities in the North tend to include more resources than they exclude, University of Ottawa for instance carries more than 23,000 online titles (University of Ottawa, 2011). Libraries are positioned in a zero-sum game bargaining position with publishers to acquire journal copies or acquiring licensed authorized access to online journals, acquiring other services enabling patron's to retrieve articles. Libraries are adjusting to Open Access. Librarians have no need to select these for purchase, but need to make choices about what resources to index, and how to assist patrons in navigating and retrieving all resources possibly available to them whether purchased or open. Libraries are also engaged in developing open access repositories for research authors to archive work produced at the institution.

Publishers have an interest in gaining an audience of readers, recruiting authors submissions and peer-review volunteers, and maintaining quality and prestige. Publishers are interested essentially in the bottom-lines of their operations, and in providing an intermediary service between authors and readers that enhance the quality of the communication between them. King (2004) notes that though every publisher aims for 'quality', there is a range which is necessary for literature. Journals with greater prestige for quality have a much higher rejection rate. Journals with lower prestige are necessary otherwise quality would serve as a bottle-neck for the dissemination of works. Quality concerns emerge if the standard falls too low or ethics are at stake. Publishers therefore have an interest in maintaining their own minimum standards and ethics, an interest in editorial independence and academic freedom. The rest of the system then has a mutual stake in the interests of publishers, and a counter-argument to OA is based on its potential threat to publishers ability to continue this role (Willinsky, 2006).

Research funders are interested in accountability to the public, the impact of research as a public good, and research policy (on the economy, health, etc.). Funders have an interest in the ethics, in equity and efficiency, in demanding accountability from recipients and being accountable to the public. This can be tricky where impact of research is a function of academic freedom and accountability is expressed through control.

Institutions are interested in producing research impact through high-quality work and publications, and attracting research personnel and funds, as well as the quality of education for students. These are concerns of administrators of higher education and research institutions.

People interested in the study of scholarship as a system are interested in how the pieces work together, what it looks like as a whole, and how the system evolves. The changing picture of research access has relevance to all of these audiences. In particular this is relevant to international development and global studies. The universe of literature archives knowledge relevant to development issues, open access will make this literature more widely available. Changes in the system of how literature is accessed means changes in people's capabilities of applying knowledge to global and developmental concerns. Without OA, the internet is a medium with a vast quantity of information but where some of the best, most useful and most powerful quality-controlled information is available only to a few.

There appears little standing in the way of Open Access, yet adaptation is never easy and interests conflict. It may be difficult to achieve objectivity on an assessment of debates about OA, because there is no distance between the literature and the subject – the self-reflexivity problem. It is suggested to the reader that when analyzing OA arguments, take into account how the individual proposing them is situated and what their interests are. Take into account your own interests with the understanding that they are not independent of the interests of each actor in the system of scholarship. Then take a leap of faith that your perspective will add benefit to the mix. My own efforts are more towards transparency than objectivity, it is often easy to spot an OA advocate by noting that most or all of their published material has been made available OA in some way, and one can 'Google Scholar' my name to verify this. Openness for some, including myself, extends to sharing beyond publication – in the form of drafts, preprints, blogs, online forum discussion etc. It is not easy to present neutrality in such a context, and yet it appears very important to consider the range of arguments and interests.

Chapter 3. Research Questions

The research questions originated as indicators in the proposed model of ICT4D, Access to Knowledge (A2K including OA) and Knowledge for Development (K4D) (synthesized as ICTs *and* Access to Knowledge *for* Development (ICTA2K4D). This is a nexus frequently referred to in various ways when OA is analyzed in a development context. After receiving very positive comments from readers and discussions with the supervisor, the key problem with the proposal was that it was too expansive for an M.A. thesis. The author elected to limit the study to key indicators around research publication availability and open access, which was nevertheless the central topic of the original proposal with ICT as a necessary and sufficient condition and K4D as a way to relate outcomes to more traditional economic and social development concerns. These resulted in selecting appropriate empirical research questions based on extrapolation of existing data with some original research for the sake of reliability, and discussion and interpretation of the results for the implications on the question of OA for development, and in terms of globalization of research.

The research questions decided upon for the thesis are as follows:

3.1: The global static universe of peer-reviewed articles

- Global total of articles: How many journal articles exist?
- Rate of increase of universe of articles: What is the rate at which the universe of articles is growing?
- What will the size of the system be in 2035? (close to the doubling time)

3.2: The dynamics of change in the system of scholarly communication – share of OA in the current output of peer-reviewed articles

- Annual OA share: What is the annual share of article output that is Open Access (defined here most broadly as 'freely accessible on the web')?
- Rate of increase of annual OA share? What is the rate of increase of the annual OA share share of article output?
- What is the combined share of annual output available to developing country research institutions as a result of OA and non-OA access?
- What will the share be in 2035?

3.3: Change in the global static universe of articles:

- OA absolute share total 2010: What is the share articles that exist that are OA?
- Rate of increase OA share absolute total: What is the rate of increase of the share of articles that exist?
- What will the share of OA articles be in 2035?

Chapter 4. Metrics of global access to research – towards a method

In 2003, the WHO (Aronson, 2003) published a study reflecting the dire literature vacuum that had come to pass in low-income countries, one that had worsened in the fourth quarter of the twentieth century. Few countries in the world were publishing cited literature. As of 2004, 31 countries accounted for 98% of the top 1% most-cited literature worldwide. 162 countries accounted for less than 2% (King, 2004). The UN then established its three voluntary donations programs (HINARI, AGORA and OARE), and eIFL and INASP-Perii developed negotiating programs aimed at advancing literature availability in the South, while OA has grown. Currently, there is a better, albeit more complex system that can be regarded as transitional.

The first set of metrics describe the entire system of scholarship from the beginning in 1664, a process that provides a denominator for an operationally defined universe of articles, for identifying absolute shares of all research – a static system where once published and preserved an article takes its place in an architecture of knowledge - and for tracking the dynamics of annual research output. All metrics require more study, and the effort is to make 'best available', article totals (as denominators for shares), OA share estimates and predictions are based on a universe of articles operationally defined as those contained in journals indexed by Ulrich's. Thus, the totals are a minimum, and the shares reflect adjustments to fit within that universe.

4.1: Why periodization?

Periodization was not initially considered in the research questions, the relevance was uncovered with the literature review and collection of data. Thus, this discussion section is included before the results. Working in a period of high investment in science, Derek de Solla Price pioneered the field of scientometrics, estimations of journal growth and identified a periodization of scholarly output in reference to the rise of science after WW2. Writing in the 1990s, Michael Mabe validated the relations between growth in researchers, articles and journals, made more accurate estimations of the size and growth of journal titles and updated the periodization, noting a transition in the 1970s as public support and growth of scholarly output slowed. Björk followed up in 2008 and 2010 publications, noting that the digital age allowed for a closer analysis of the article as the primary unit of scholarly communication rather than the journal titles, and applied the methods employed to make estimations of annual article output and putting the focus on the OA share in this output (Björk, 2 refs and unpublished).

Björk did not refer to the periodization, and initially, it did not appear centrally in the focus of the research, however this changed with an analysis of the data. The data revealed quite significant changes since Mabe's last mention of the periodization in

2000, in UN figures on the number of researchers and the number of published articles (based on the Web of Knowledge database) and the regional distribution of research output. Price broke the ground on modern scientometrics and established the relationships of researchers to journal titles and article output, and articulated the relationships between articles and analysis of citations. These demonstrated scholarship to be a self-organizing system that could be studied and organized by citations, from which the first comprehensive citation databases were later developed starting with Thomson's ISI (which became Web of Knowledge).

Price's gross overestimation in predicting growth in journal titles does not detract from these contributions, which is encouraging to current research. While estimates improve they remain uncertain, and predictions remain even more uncertain. Yet, the value of scholarship is that the process becomes transparent to others who can improve it, who can correct the estimates and move the general understanding of how things work forward. Mabe was able to build upon the work of Price and the understanding of the dynamics of the system and by the 1990s had access to improved digital search technology in the world's most comprehensive and long-standing database of periodicals. The nature of the search technology (then by CD-ROM and now available on the web) lent itself to operational definitions of journals and articles as those classed as Active, Refereed and Scholarly/Academic in Ulrich's. Mabe was also able to use the 2001 CD-Rom provided by Ulrich's to validate hypotheses regarding the stability of growth rates and periods of increased and decreased activity.

From this, Mabe revisited the Little and Big Science periods, using UN data to observe more recent growth rates in the numbers of researchers and the CD-ROM to track changes in the number of periodicals over three centuries. Mabe reported an average 3.4% growth rate over three centuries. He observed some changes related to the periods, though reporting that despite external events, growth remained remarkably stable and changes could be viewed as less dramatic than expected. Thus, the Little Science period could be measured as the period before WWII, where the year over year growth rate of active journal titles was around 3%, the Big Science period where the rate was 4.35% from the mid 1940s to the mid-1970s. Mabe observes a decline to roughly 3% again from the mid-1970s to the time of writing in 2000, which he calls the Disillusionment period, attributing this to a relative decline in public patronage of science after the high post-WWII investment in science saturated. The Disillusionment period coheres very well with studies about global changes in social policy across the West, and an understanding of major political change from a 'Keynesian' post-War era to a 'neo-liberal' post-recession era in the last quarter of the century (Banting, 1995). Referring again to the relationship and data on the number of active researchers and average output per researcher, Tenopir and King (2000) estimated a global annual output of 1 million articles at the turn of the twentieth century.

In a review of scholarly publishing and relying on the same line of research, Ware estimated 1.4 million articles writing around 2007 (Ware, 2008). These estimates vastly improve upon those of Price, and are more precise than estimates for numbers of journal titles made King et al (1977) at 57,400 and in 1995 by Meadows and Singleton at 70,000 to 80,000 (in Tenopir and King, 2009).

In the each of the estimates based on the relationships between number of researchers, number of journals and number of articles (Mabe, Mabe and Amin, 2001, Tenopir and King, 2000, 2009), cohere well with the relationship to the reported growth rates (Mabe, 2003, Ware, 2008), as well as the methods of using key databases with journal to article ratios, the use of Bradford's law to estimate the number of non-ISI journals from the number of ISI-indexed journals (Mabe, 2003, Björk, 2006), and the sampling methods employed by Björk. The following graph locates these estimates along the growth curve for scholarly output over more than three centuries.

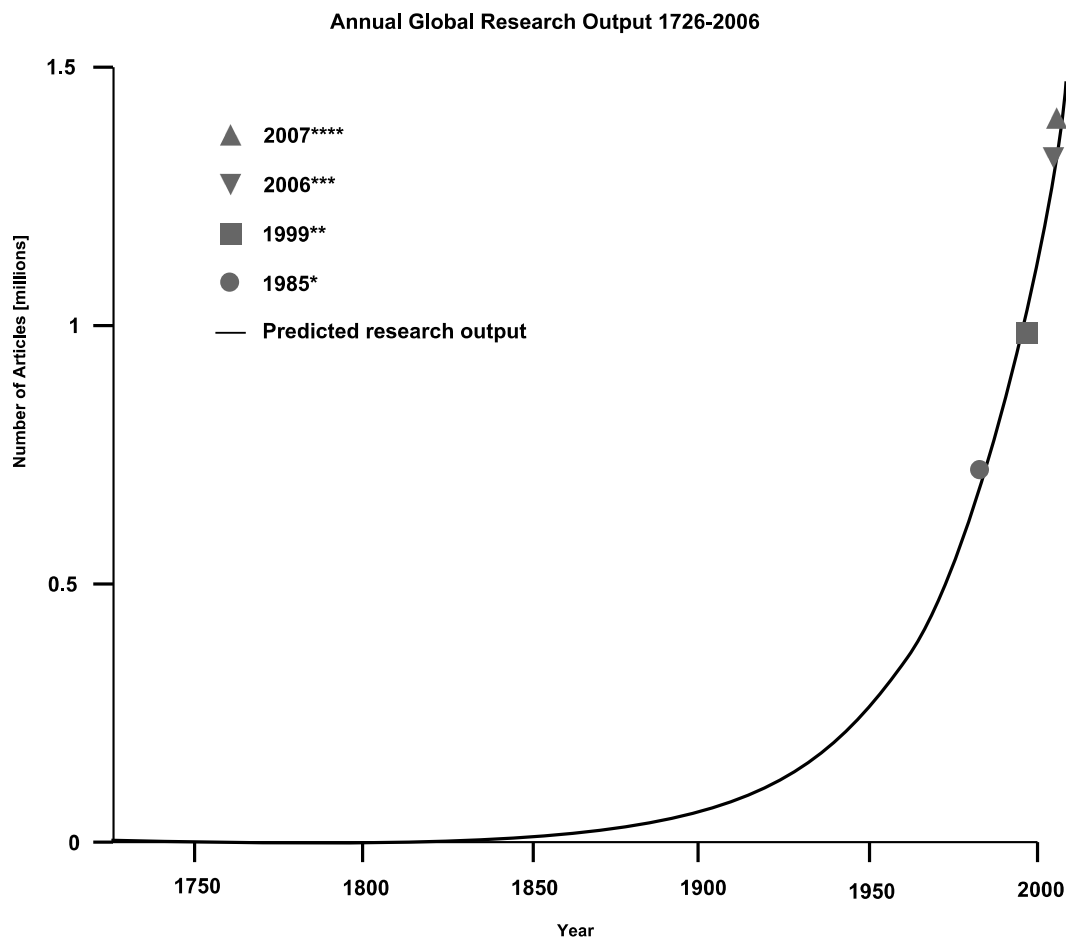


Figure 1: Estimated annual global research article output at 3% annual growth

- * Year 1985 – doubling time of just under 24 years.
- ** 1999 – Corresponds to estimates by Tenopir and King[6], for research output in the late 1990s, of 1 million articles per year.
- *** 2006 – Björk et. al's estimated 1.35 million articles in 2006.
- **** 2007 – Writing in 2008 - Ware reported 1.4 million as the current annual output.

3 tentative conclusions can be made: a) the 'best estimates' of the quantitative scope and growth of the global scholarly system have improved and we can have greater confidence in the 'ballpark' and b) it is possible to estimate its absolute size by creating a curve based on growth rates and summing annual estimates; c) the current methodologies suffer some weaknesses that can be addressed by future research. The development of further methods and technologies, the internet and openness (OA and Open Source) can have a key role in terms of the possibilities for future search, navigation and filtering technologies which will face fewer barriers (such as digital

locks) while software developers may enjoy rapid acceleration in an open source environment naturally situated in the academic environment (This last will be explored a little in the discussion looking towards the future). There is more work to do.

4.2: Article 50 Million – published en route to the thesis

Two interesting discoveries were made as data was collected in the early phases of the thesis. The first was that the first empirical question of 'how many articles exist?' could now be estimated. This was done by the author and published in 2010 in *Learned Publishing*, the title revealing the result – Article 50 million. Secondly, as will be discussed, the rates of growth for journals and articles were discovered to have increased over and above the Big Science period since 2000, reversing the disillusionment trend. Further, the shares of research output coming from and the number of researchers operating in the global 'North' (Europe, North America, Australia/New Zealand) declined significantly, though still accounting for a majority. Thus, the increase is attributed to globalization. The globalization is not distributed equally globally, not suprisingly it is China, India and the Asian Tigers that account for most.

Still, Africa's share of this output remained the same or increased slightly during the last ten years where the entire system grew. Only the share to the North declined. Thus, it became necessary and quite exciting to be forced to conclude that the disillusionment period no longer applied, and the most likely candidate for a new period would be 'global science' or 'global scholarship'. Finally, the purpose of addressing these questions was to arrive at the question of interest – the OA share and its growth, both annually and as a function of the universe of articles. The evidence suggests some early support for a prediction made by Steven Harnad, that universal OA is inevitable. Growth rates are advancing dramatically, both in the numbers of Gold OA titles and articles and in the number of Green OA mandates. Advocacy continues to be strong and there is no evidence of a reversal of the trend. Already 20% of articles are freely available on the web after 20 years of OA with the fastest growth in the last five years. Thus, a final addition to the periodization is a predictive one which we offer with the caution provided by Price's early work. It is offered from the best available evidence with the hopes that others take the research further. The prediction leads to suggesting a period of Open Science emerging from within Global Science.

As a final note, it is commonly known that journals began in 1665 and Mabe suggests from his research into the Ulrich's database, that steady growth at 3% in the Little Science period can be observed from 1860 on. A final addition to the periodization is the period of Early Literature, similar to Little Science but one where less can be observed. Given the dwindling share of total accounted for by earlier and earlier literature, the fact that less is known about its size and stability is not very problematic for making estimations here. For the purpose of extrapolation and backward mapping, the year 1726 is chosen, corresponding to the reproduction by Mabe representing 'stable growth of 3.4% across three centuries.

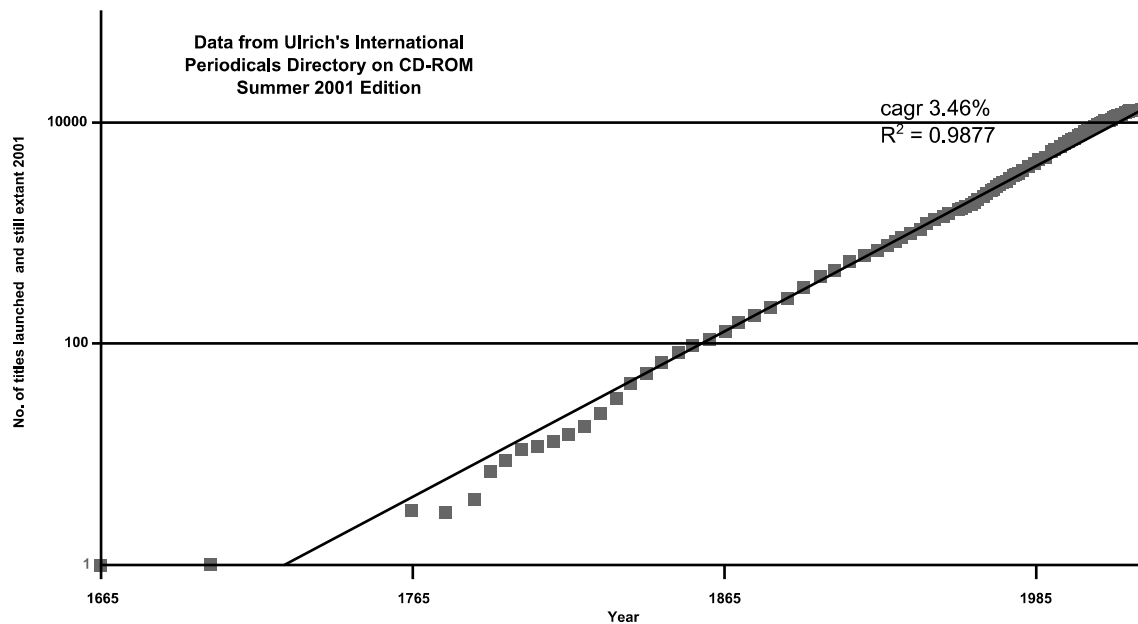


Figure 2: Ulrich's report of the rate of publication from 1665 to summer 2001 (Ulrich 2001)

Thus, the following periodization is hypothesized which brings the research full circle in the methods and analysis from Price to today.

1. Early Literature - 1665 – 1860
2. Little Science⁶ - 1860 - 1940 (Mabe, Price)
3. Big Science - 1945 - 1974 (Mabe, Price)
4. Disillusionment – 1975 to 2000 (Mabe)
5. Global Science/scholarship– 2000 – current (Jinha)
6. Open Science/scholarship – current to future

⁶ It is somewhat unfortunate that the word Science is used, rather than Scholarship which is more encompassing of peer-reviewed literature.

Chapter 5. Methods

5.1: The Ulrich's universe' of titles and articles – operationally defining the universe of articles

The attention to progress in estimation methods and the extrapolation from previous data from which the estimation is made has advantages and disadvantages, but is regarded to be an improvement over previous attempts.

First, it leads to an operational definition of the universe of articles that reflects the contents of the most comprehensive database available, Ulrich's. The advantage of this is journals listed are classified in way that allows filtering out of journals that are not 'active', 'academic/scholarly' and 'refereed'. A second advantage is that the reported growth rates of 3.4 % for journals and 3% of articles are based on the study of Ulrich's data, hence a universe of Ulrich's articles. Thirdly, the growth rates and the annual figures obtained with these methods have internal consistency with studies based on what is known about the relationship between the numbers of researchers and the average # of publications by researchers, the number of journals and the number of articles. Despite external upheavals in the world, there is good evidence to suggest that the number of researchers has grown at a steady and linear pace from year to year, and thus the number of journals and articles has grown steadily, and because of the nature of citation, the volume of previous work steadily stimulates current work as a general rule (Ware, 2008, Mabe and Amin, 2001, 2001, Mabe, 2003). It has thus been described as a self-organizing system. The average rates experienced some variation in the 20th century, leading to a periodization by Mabe of little science, big science and disillusionment periods (Mabe, 2003).

More recent evidence presented here suggests change in the system provoked by digitization and globalization, and an increase in the number of researchers and publications since 2000 driven primarily by transition countries like India and China (also referred to as developing, but not least-developed) but with the Americas and SSA also gaining in the share with the developed North (N.A., Europe, Australia and New Zealand) still producing the majority but with a declining share (UNESCO, 2010). This suggests a new period of globalization and digital availability.

The work by Björk to identify a number for annual output for the year 2006 follows similar methods to the work of Mabe and Tenopir and King in the late 1990s and early 2000s, which builds on work from the 1960s which demonstrated the relationships in the growth of research funding, researchers, authors and journals in a self-organizing system, but over-estimated future journal growth (Mabe, 2003, Tenopir and King, 2000, 2009). This work and that of Björk is based primarily on an analysis of

the world's largest and longest-standing database of periodicals, Ulrich's. The work therefore tends to operationalize the universe of journals and articles – which I will term here 'an Ulrich's universe', with the caveat that explains the limitations of this, mainly that it is not comprehensive and the information is chiefly based on self-identification (Tenopir and King, 2009).

The work of Mabe and Amin (Mabe and Amin 2001, Mabe, 2003) sought to correct the overestimation of the numbers of journals by referring to new filtering capabilities in the Ulrich's system (then accessed on CD-ROM and currently accessible on the web platform) that allowed one to filter by the designations of 'active' – 'academic/scholarly' and 'refereed'. 'The Ulrich's universe' can be assumed to represent an under-inclusion of the actual universe of journals that is greater than a possible over-inclusion by mis-identification of journals as 'active, academic/scholarly, and refereed'. Moreover, the subset of 'the Ulrich's universe' that can be defined as ISI-indexed journals represents the highest volume and most-cited journals, accounting for 75% of the publications (Björk, 2006) and 95% of the citations (Mabe, 2003) of journals in an Ulrich's universe.

What is outside the Ulrich's universe would be composed of even smaller, lesser-known journals. Björk compiled a master journal list as part of the investigation of journals and articles for global OA shares, that included some 30,000 titles that fit the criteria. This list may not be entirely comprehensive, but closer to the goal. The list adds some 4,000 or so titles to the universe, but given that the journals are not ISI-indexed, the article average for them would be expected to be the same or lower as the subset in Ulrich's that is not ISI-indexed. Thus the the Ulrich's universe may represent 80-85% of titles, but upwards of 95% of articles.

Given a chance of error in any calculations that would lead to an overestimate in the number of articles that is lower than the rate of under-inclusion, we can safely say that the Ulrich's universe is an accurate description of the minimum number of articles. This is further assisted by the application of Bradford's law to the ISI-index, a mathematical technique based on the logarithmic relationship of citations and the data that shows 95% of citations are based on the ISI journal set. Application of this to the size of the ISI set typically gives a figure for annual journal titles that is slightly larger than the search in Ulrich's produces for a total of titles.

On this basis, we should regard figures in the most recent studies to represent best available estimates. We can regard the 'Ulrich's universe' as a minimum that is close to the actual universe, and we have evidence of internal consistency of the estimates by studying the relationships between numbers of researchers, journals and articles. Other estimations tend to produce errors that owe to a lack of definition (active, academic/scholarly, refereed) and do not cohere with the known relationships.

Some databases produce a higher number of records suggesting estimates based on these methods to be somewhat conservative, but these may be due to duplications or inclusions of grey literature, or failure to limit by year, or by type of publication.

5.2: Databases/Indexes Used:

Acronyms are provided for reference within the paper, or for future reference.

SCI	Science Citation Index
SSCI	Social Science Citation Index
AHCI	Arts and Humanities Citation Index
DOAJ	titles indexed in the DOAJ
U	Ulrich's
ISI	The Web of Knowledge is originally the ISI, the original citation database. The Web of Knowledge contains the SCI, SSCI and AHCI. Articles indexed in the Web of Knowledge are still commonly referred to as 'ISI-indexed'.
JCR	Represents a slightly smaller subset than ISI-indexed. The difference occurs because one can search Ulrich's for JCR and Web of Knowledge for SCI, SSCI, and AHCI.

5.2.1: Filters

The following is a list of filters used when using the databases for this research. Notably, these databases are designed for ordinary research to locate periodical titles or articles, which requires us to make adjustments.

ASRA	academic/scholarly, refereed and active (a filter which operationally defines articles within an Ulrich's universe and applies to 'ISI-indexed' as a subset).
------	---

The following all include the ASRA filter:

U-ASRA	titles in Ulrich's (and article totals based on this) filtered by ASRA.
U-ASRA-OA	as above, ASRA titles listed as Open Access in Ulrich's.
U-ISI	titles in Ulrich's that are also ISI-indexed. There are a few to approach this. For the sake of operationalizing this set, it is defined by the totals of a search in Ulrich's filtered by 'Abstracting and Indexing Services', that includes SCI OR, SSCI OR AHCI.

- U-ISI-OA. A search using the filters above for U-ISI and Open Access.
- U-JCR. A search used by Björk to get article averages for ISI indexed journals. Björk also notes a slightly smaller average for the AHCI compared to SCI and SSHI, and incorporates the AHCI average into their definition of ISI-indexed by combining it with U-JCR. U-JCR produces smaller numbers than U-ISI-OA, the method I employed to get results that are consistent with Björk 's methods.
- U-JCR-OA. A search I performed to get titles for ISI-indexed journals for article totals for OA and make another comparison of averages with Björk 's figures. The use of different methods adds reliability to the estimates.

5.3: Backward mapping

The operationalization by an 'Ulrich's universe' of active, scholarly and peer-reviewed journals (U-ASRA) allows the possibility of mapping backwards to generate a grand total for all articles with a data set that gives year by year totals since 1726, a year that corresponds to the identification of the start of steady growth. 1726 is not the initial year of journal publishing, so the number of articles for that year would be greater than zero. The number chosen was that which leads to correlation with Björk 's report for annual output for the year 2006, a figure that was achieved independently of the growth line, so it is a method of 'backward mapping'. The number is arbitrary and is not meant to be historically precise, but the total for 1726, as well as the hypothetical sum of articles from before 1726 are small enough to be negligible to grand sum. The further back one goes, the more the figures for each year should be treated in terms of the relationship to the method of backwards mapping and averages over periods of time.

5.4: The altered data set

The data set used in Article 50 million does not include altered growth rates though one was produced in the course of the research. The altered set did not affect the totals significantly so was not presented for the sake of the focus of that question, but year to year totals would be less accurate to the date. The data set here does accommodate these altered rates in the different periods, including new research showing a new period since 2000. Thus, the year to year totals are more accurate and this is important as they are used as the denominator for the share of OA indicators to calculate growth rates of the share. This resulted in the use of a different figure for a starting point, also negligible to the total. To summarize,

1. Despite the altered growth rates, the overall growth rate of 3.4% for journals and 3% for articles holds when calculating the sum for all articles.

2. The year to year figures from the second data set will be more precise because we have taken into account altered growth rates in four periods.
3. All of the figures are based on defining the universe of articles by the nature of the search for the 2006 figure, Ulrich's universe of 'active': 'academic/scholarly': and 'refereed' (U-ASRA). The earlier figures stand in relation to the growth of a system of such articles. As with other authors, Björk and myself consider this methodology to produce 'best available' results understanding the limitations and encouraging further research.

The disadvantages of using 'the Ulrich's universe' are as follows:

- a. It does not include all journals that are active, refereed and academic/scholarly. There are likely some journals that are not in Ulrich's universe and this may apply more frequently to new journals, with an unknown average time lag for entry into the database;
- b. if one drops one of the filters, the numbers can increase drastically; and
- c. it relies on self-identification, and thus there is the possibility that some journals listed as active, refereed and academic/scholarly. Some journals that may fit all criteria may not be listed appropriately and would not be counted when filtered.

The disadvantages are buffered by the cancelling out between over-inclusion and under-inclusion effects.

5.5: Other recent estimates

Compared to other estimates recently and in the past, the results are more conservative and more precise. The estimate 2.5 million articles per year is cited in many places (Gargouri et. al, 2010), but this is based on the number of titles in Ulrich's and an article average based on the Web of Knowledge (ISI-indexed) (Harnad, personal communication, 2011). However, as noted by Björk et al (2008), the average articles per title in the smaller number of titles in Web of Knowledge is much higher. The 33% of titles in WoK (hitherto referred to as ISI-indexed articles) compared to Ulrich's produce 75% of the articles published. Calculating the remaining 2/3 of titles that only account for 25% of output at the same average produces a large over-estimation for the number of articles. Though this rough estimate was published in a refereed paper, it can be regarded as an overestimate.

The methodological challenges are complicated further by the attempt to estimate the global OA share. In Björk's study for the year 2006, the internal consistency of 'the Ulrich's universe' is maintained, since OA journal totals from Ulrich's are used, rather than the DOAJ which lists a higher number of titles. In the Björk study of 2009 (Björk, 2010), the universe is extended by the compilation of a master database of 30,000 titles and the DOAJ is used rather than a search in Ulrich's

filtering for OA to identify the sample of journals. The results are based on a user's potential of finding a gold OA article or a green OA version of any article randomly sampled from the journal database.

5.6: Conservative Adjustments

The effort here has been to make adjustments that result in the most conservative estimates for current shares and growth rates. Thus, more current estimates of OA totals are expected to err on the side lower OA share for a more conservative snapshot, and older estimates to err on the side of higher share (and thus more conservative growth). This is a trade-off that allows us to use simple arithmetic approaches and report a current OA share 'of at least xx%', and growth of 'at least xx%/year', rather than attempt to achieve more precision by using esoteric statistical methods.

5.7: No Direct Method to Track Green OA Growth Trends

Finally, despite efforts to find assistance with the question through contacting the most relevant experts, it appears no one is able to determine a good method of calculating growth of the Green OA share. The prediction of this share and its impact on the total can only be made by assuming 2 possible scenarios for calculation, and a third possibility through hypothesizing the pattern of Green OA growth to Gold OA growth. The most conservative estimation is based on simply taking the two studies by Björk (2008 to 2010) for years 2006 to 2009 and assuming a linear pattern. This however, results in a pattern where Green growth is outpaced significantly by Gold. There are good reasons to predict that Green will maintain either a pace close to that of Gold, or faster. First, the difference between the two studies may be under-estimated due to methodological differences and as explained below, under-estimation seems more likely than over-estimation. There has been intense advocacy for Green OA based on the premise that Green is the fastest route to universal OA, a strong logical argument based on years of analysis and experience with self-archiving and OA mandates, emphasized by many OA leaders.

This advocacy cannot has resulted in an exponential growth of OA mandates from 2006 to 2008. OA mandates, as defined by green OA advocates, differ from voluntary policies which have also grown dramatically, in being the only effective way to move an institution from the reported stable, but low rate of voluntary self-archiving at 15% to near 100% archiving. Institutional mandates are accompanied by departmental mandates (within-institution), funder mandates and the beginning of multi-institutional mandates. Though there is great difficulty measuring this precisely because mandates vary so much in their design and enforcement, a mandate is expected to have this effect on moving the archiving rate towards 100% after two years. The Björk study published in 2010 based on the year 2009 would have tracked the impact of mandates in 2006 as they hypothetically reached fulfilment by end of 2008. As one

can see below, 2006 marks the beginning of the rise. Studies done thus far have not tracked the impact of a dramatic rise in mandates. Despite the rise, the number of mandates that exist to date are still a small portion of the number possible (Harnad, 2011) giving room for growth while advocacy has not slowed. Harnad frequently points out that mandates give rise to close to 100% OA, whereas one Gold OA journal at a time seems a long route. However, this is 100% at a time for what is covered by a department, institution or funder, and achieving these is a political process quite different and often more challenging than a decision by a small journal to go OA. It remains impossible to force institutions to act or predict how quickly they will. Does the general policy and advocacy environment affect decisions of how to launch a journal? A quick test on Ulrich's using the filters -U-ASRA and U-OA together with Start Year 2009-2011 reveals that 29% of Ulrich's journals launched (or quickly converted) Gold OA in the past 2 years.

Thus one can describe a situation in the past 5 years where the publishing of Gold OA articles has shown a dramatic rise which has an immediate effect, and the number of mandates has had a dramatic rise which has a delayed effect not yet observed. The fact that Green mandating and Gold publishing have both risen together suggests a common phenomenon – an OA trend rooted in its economic and ethical rationale and supported by ongoing advocacy. We have a correlation that suggests interaction, and a good hypothesis for further research.

This leads to a second possibility for predicting the trend based on a relationship with Gold OA growth. Green OA has been responsible for a larger portion of the OA share than Gold over all of the studies. If we assume that the Green portion maintains at least the same level at a minimum of 3% higher than Gold, we can calculate the trend of based on adding 3% to the results for extrapolating Gold. By this method, we can make a prediction based on expectation that OA continues to rise at least at the pace of the last 5 years.

A third possibility, not calculated, is that the advocacy of Harnad and others is effective, mandates continue to rise and their effect is observed in the future to show Green growing faster than Gold. Though this is not calculated, it suggests that overall OA growth even faster. Another possibility is that Gold OA is outpacing Green OA, and this would be surprising given the fast and continuing rise of mandates whose effect has yet to be tracked. However, it is another hypothesis that could be tested by a researcher who can sort out how to measure green OA growth.

Chapter 6. Results and Analysis

6.1: *The global static universe of articles:*

6.1.1: Global total of articles: How many journal articles exist?

I published this result in *Learned Publishing*, currently there are roughly 50 million peer-reviewed journal articles in existence, published from 1665 to today (Jinha, 2010). The results were very close in a calculating based on average growth rate at 3.4% and one based on periods with varying growth rates. The formula is the same, as follows: $n_{j+1} = n_j * (1+i)^p$, where n_j is the number of articles in the current year; n_{i+j} is the number of articles in the next year; i is the percent growth rate between zero and 1; and p is the number of periods.

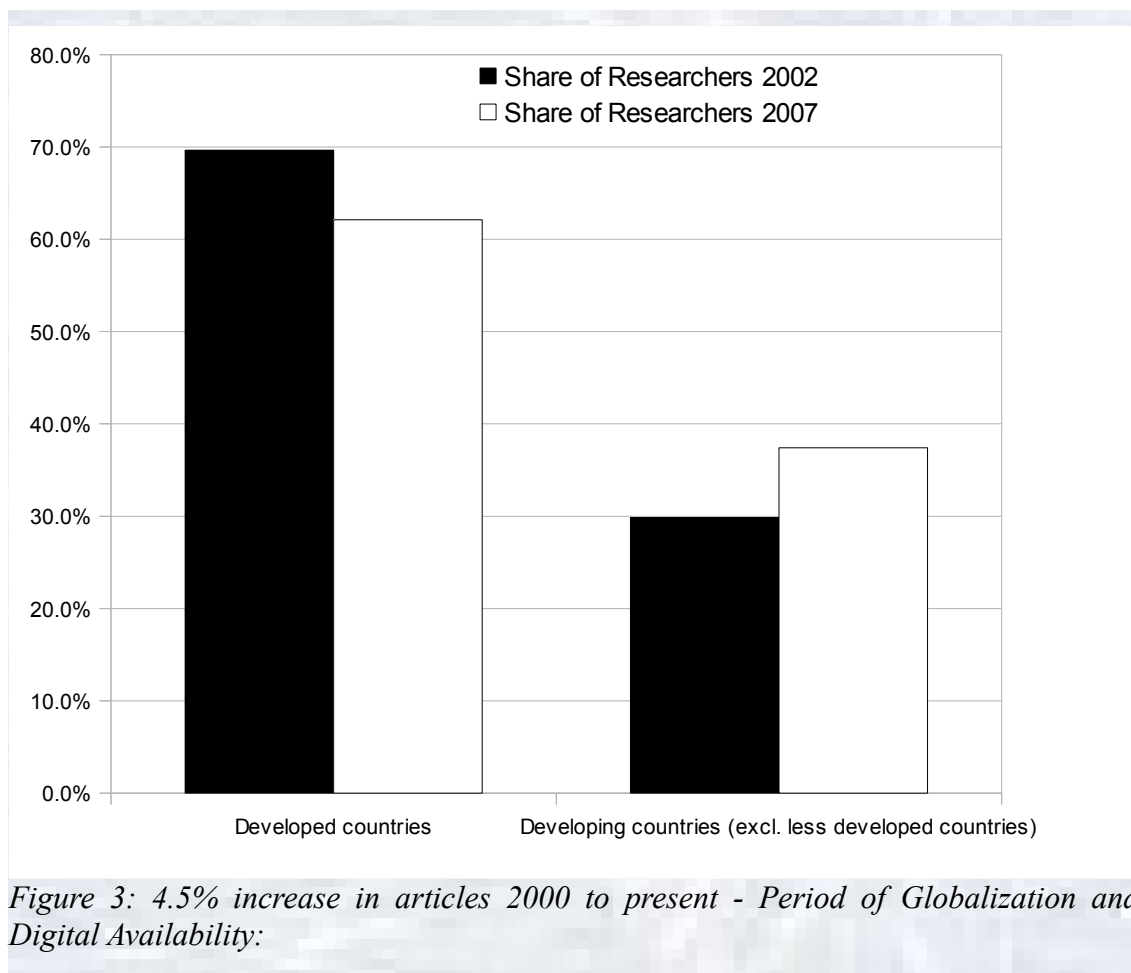
6.1.2: Rate of increase of universe of articles: What is the rate at which the universe of articles is growing?

The rate of 3% per year over three centuries has been reported (Mabe, 2003, Ware, 2008) based on Ulrich's data and studies relating numbers of researchers, journal titles and article output.

The rate of active journals increased 3.46% per year for three centuries on average. It grew slower from 1860 to 1940 at 3.25% and faster between 1945 and 1975 at 4.35%, slower again from 1977 to 2000 at 3.26%. The article rates in the adjusted data set produces only a slightly different result than Article 50 million for the timing of when we reach 50 million articles, and provides more precise annual totals for use as denominators of OA shares.

For the historical rates, the method is to adjust article rates by the same arithmetic ratio as the overall rates for journals to articles - 3% to 3.46%. Primary data for these periods is harder to access and has already been analyzed to show a consistent 1:1 ratio for growth of researchers and articles. For the recent period, the article rate is adjusted according to change in the increase in researchers.

My results show evidence for a minimum 4.5% growth rate for article from 2000 to the present day, based on recent UN data showing a 4.5% increase in researchers from 2002 to 2008.



The argument can be made that globalization may alter the ratio, that more researchers may not translate into more articles, the evidence from the increase in journal titles would suggest an even higher growth rate in articles. UN data shows that as the numbers of researchers increased 4.5% from 2002 to 2007, the number of ISI-indexed articles increased 5.8%. Researchers reported 8000 journals in the ISI at that time and 15,000 in Ulrich's. Björk also estimates 8,000 titles in ISI in 2008 (Björk et al 2010), but closer analysis reveals a higher number is possible.

A rigorous understanding of how to interpret journal title totals from various sources requires reporting further detail of the use of filters and databases in this study. If Ulrich's is used to identify journals in the Journal Citation Reports (the U-JCR filter) as ISI-indexed, the number is just under 8,000. However, if searches for each of the indexes for the Web of Knowledge are used, the number is 10,000. If the master journal lists for Arts & Humanities Citation Index (AHCI), Social Science Citation Index, and Science Citation Index are used, the number is just over 12,000. Thomson's reports indexing 23,000 journals and over 11,000 high-impact journals. Each of these figures are associated with the designation ISI-indexed, and it is a matter of using

corresponding figures to get article averages, and the appropriate figure to calculate a sum. The approach here was to calculate averages in more than one way to get a more reliable average and check for consistency with the literature, and use the U-ISI filter as described above to calculate a sum consistent with the Ulrich's universe.

The UN data calculates the total number of articles from the Web of Knowledge. The number of journals in Ulrich's rose from 14,694 (Mabe) in 2001 to 23,750 in 2006 (Björk), representing a 10% increase, and rose another 5% between 2006 and the current total in spring 2011 of just over 28000. The 10% could be due to a methodological factor, Mabe used a boolean search and not only included 'academic', 'active' and 'refereed', but used ANDNOT to exclude a number of factors. The search was done on a CD-ROM and difficult to replicate. I conducted a test of some of these terms, inclusion of the ANDNOT can reduce the number but not substantially. The brief test I did revealed that using ANDNOT tends to exclude journal titles that should be included. The test excluded up to 100 titles when it had an effect, and some terms did not. Therefore Mabe's estimate can be regarded as too low by hundreds, but not thousands. A rounded figure of 15,000 titles in Ulrich's is a good estimate for the year 2001 and therefore the rise is still between 9 and 10%, which is more than expected based on the 5.8% article growth. One possibility is that this indicates that following the move to digital, Ulrich's has become more comprehensive. The Ulrich's universe grew faster than the actual universe of journal titles, and has come closer to mirroring it. Despite this, we can still predict that what fell outside Ulrich's represents smaller journals with lower article averages, and that the Ulrich's universe growth patterns and variations would be a mirror of the actual universe of titles and articles.

However, the increase of 5-6% in ISI indexed titles, the 5.8% rise in ISI-indexed articles, the 5-10% increase in Ulrich's titles and the the 4.5% increase in the number of researchers and all suggest an increase in the rate of growth of publications since the start of the century. If the same ratio of researchers to unique authors who publish an average of one article per year (Tenopir and King, 2000, Mabe, 2003) holds, articles should increase at the same rate as researchers – 4.5% per year. The adjusted data set based on the historical variations in growth reported by Mabe could then be adjusted for the recent years of the 21st century, yielding more precise denominators for calculating OA shares.

6.2: Growth of the Static Universe of Articles

What will the size of the system be in 2035? 2035 is close to the doubling time for the overall growth rate for articles, which is the basis on which it was chosen. Taking into account the faster growth rate recently, the doubling time would be earlier. However, it is a round figure representing a quarter century from now and the choice is to retain the marker and how the research question was constructed before discovering the change.

Scenario 1 – assuming the reported growth rate of articles

If we assume the 3% growth rate of the Ulrich's universe of peer-reviewed articles, the number should pass 100 million in 2034, with a 24 year doubling time. If under-inclusion is the assumption (which appears more likely than over-inclusion), the estimate based on 'the Ulrich's universe' would be regarded as conservative, or a minimum.

So, at least 50 million articles in 2010 and at least 100 million articles in 2035.

Scenario 2 – adjusting for recent period of globalization of research

If we adjust for adjusted growth rates in the 20th century, and the the 4.5% increase since 2000, and hold constant for the researched figure for 2006 (by adjusting the initial figure for 1726 from 344 to 275), the number of articles will pass 50 million in 2012 and 100 million in 2031.

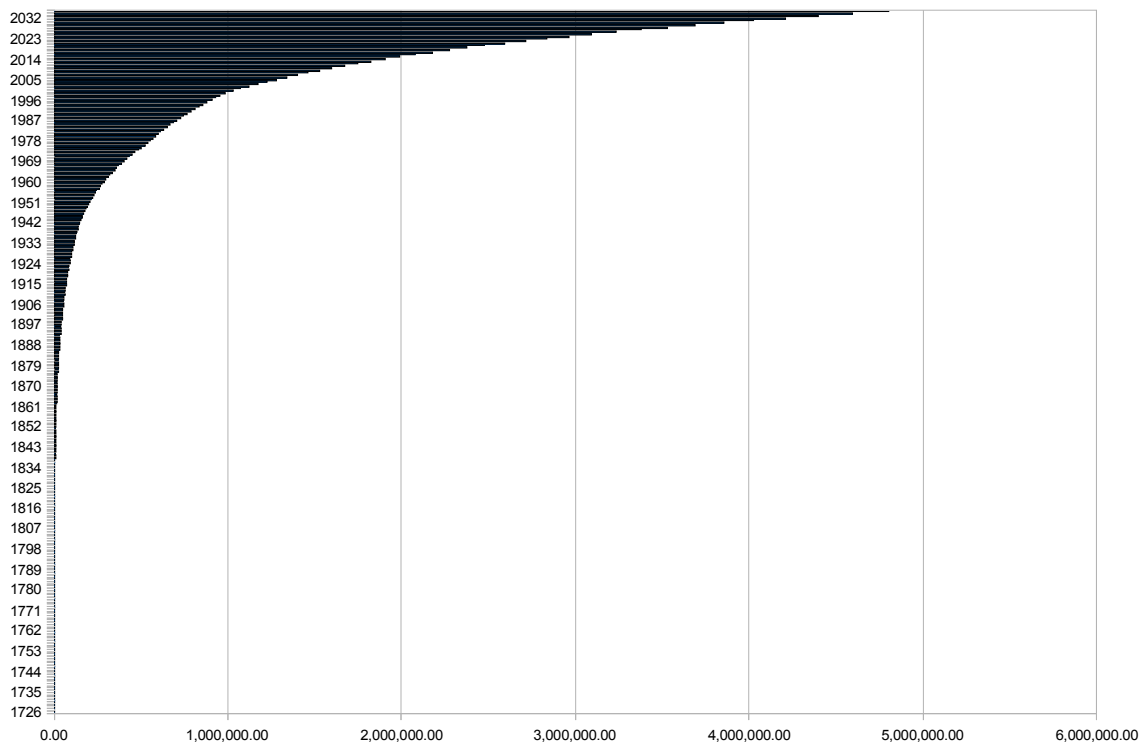


Figure 4: Annual output

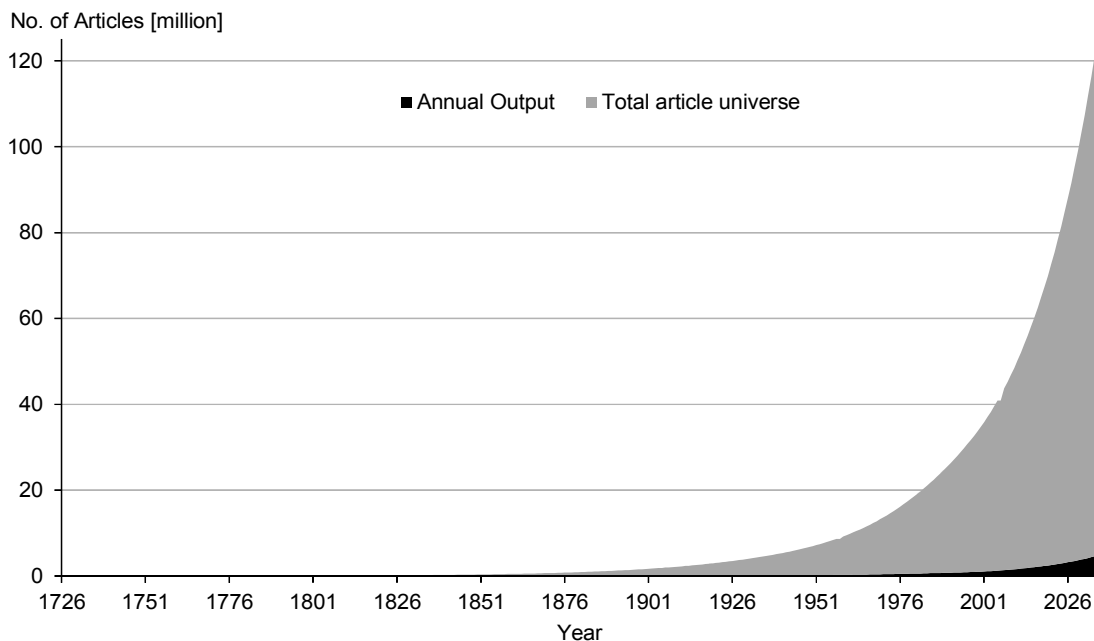


Figure 5: The total article universe (grey) plus annual output (black).

6.3: Shares of OA

Annual OA share: What is the annual share of article output that is Open Access (defined here most broadly as 'freely accessible on the web')?

6.3.1: Green OA

Archiving of green OA articles can be traced to the establishment of Arxiv in 1991(Ref). Institutional and funder mandates have risen dramatically since 2005-06, now tracked by Enabling Open Scholarship – openscholarship.org

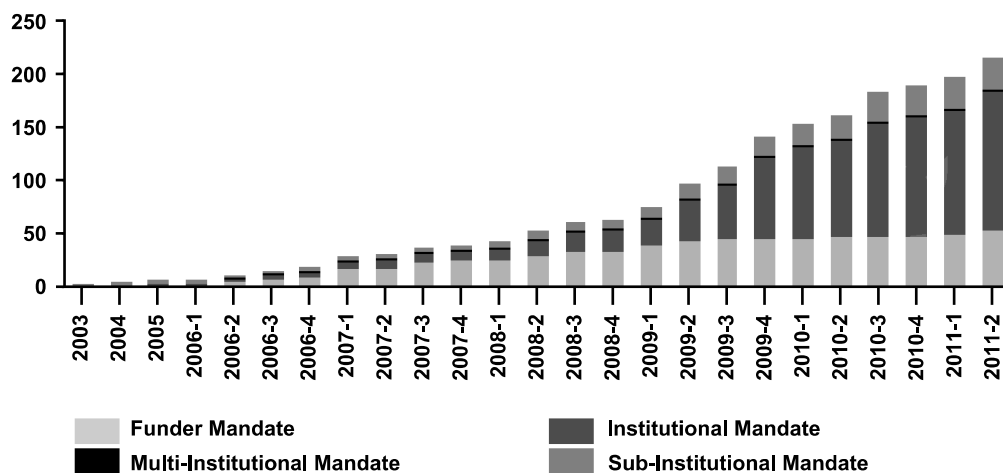


Figure 6: The growth of OA mandates – from <http://roarmap.eprints.org/>

Advocates of mandating point out the vast difference in success of voluntary policies compared to mandates in the percentage of deposits. An excellent study was performed by Sale (2006) that tracked the difference between voluntary and mandatory policies, and some variation between institutions as well as the time it takes to reach 100% compliance. Swan and Brown (2005) conducted a study of author's attitudes towards mandates. Less than 5% stated they would not comply, 13% would comply reluctantly and 81% would comply willingly. On the basis of the evidence, Harnad (2010) proposes that within 2 years, an institutional mandate should approach 80-90% effectiveness. Funder mandates such as the NIH follow the same pattern in low rates of deposit for voluntary policies, and rapid increase towards 100% within a few years. Still, with the range of mandates, it is not clear we can define this as a useable average. For Green OA, there is less direct data, so we can make only indirect hypotheses and suggest different scenarios. The data we have comes from two studies by Björk et al. (2008, 2010).

Green OA reported by Björk

2006 – 11.3%

2008 – 11.9%

Both of these figures were obtained based on a sample of articles from Ulrich's journals, excluding gold OA articles, and the rate at which a useable green copy could be found. This represents an increase of .3% per year. Mapping backwards to 1991, one would have a 6% rate of Green OA for the year Arxiv was established, which appears impossibly high unless we consider that some form of self-archiving has existed pre-internet. For the sake of argument, we will use this as one hypothetical calculation going forward, though we can safely assume a faster growth rate over the last 20 years. The best approximation would be based on finding a true base rate for unmandated self-archiving and adding the overall effect of new mandates to this, but I am unable to find a way to do this not for lack of effort, none of the experts consulted were able to advise on this. It is difficult to use the .3% rate of increase meaningfully to go forwards because of the rapid increase in mandates between 2006 to present, particularly in the last 18 months. Hypothetically, there is some lag to these mandates achieving an effect and the rate would be expected to increase. Further, the smaller manual sample of Björk 's initial study may be much less reliable than the sample using a software program in the second study, so the size of the gap may be an artifact of methods.

This is not quite satisfactory and moves toward the kind of efforts people make at betting on baseball, where people may actually improve their odds through a thorough knowledge of the game. In the absence of a method, there are some other interesting possibilities left to suggest from what is known, from four years of the author's attention given to the game of OA. Given the pace of mandates and the continued emphasis by many advocates on self-archiving over Gold OA (Harnad and

others, several), *a prediction can be made that Green OA will at minimum stay ahead of Gold OA*. Since the Green OA is linked to broad policy efforts rather than independent efforts by publishers, this means of prediction is also a challenge to Green OA advocates to at least 'keep up with Gold' and maintain a pace of increase that does not fall behind Gold. Falling behind Gold would contradict the logic of many leading OA advocates who see self-archiving as the faster path to open scholarship (Harnad, 2011, Chan, Kirsop, Arunachalam, 2005). This second hypothesis for forward and backward mapping would be to maintain Green OA delivering more OA availability than Gold by the current 3%. This interpretation is much better when in mapping backwards to 1991 and leads to realistic numbers at the beginning of Green OA. Because of the longer-standing tradition of self-archiving related to author's making prints available, we may even be able to assume that as Gold OA began, Green was already ahead a few percentage points and remained there, with each route growing together from the beginning. To perform this calculation, we first need to analyze Gold separately.

6.3.2: Gold OA growth

The two studies by Björk show an 8.1% to 8.5%, an increase of .2% per year. This would be the crudest and most conservative estimate, and when combined with the same treatment of Green, calculating the OA share would be based solely on the face of these two studies – an increase from 19.4% to 20.4% between 2006 and 2008, .3% per year.

This method does not reflect the trend over the last 10 years when OA became prevalent, or the last 20 since its inception. Backward mapping based on these rates results in impossibly high percentages at the dawn of OA, and unreasonably high percentages at the year 2000. The actual growth rate has been faster. What is interesting is that modelling on this pattern that is undoubtedly slower than that of the last 20 years still reveals a steady march where 1/3 of global research is freely available in 2035.

In Björk et. Al's newest study just published (Björk, 2011), the focus is on Gold OA articles totals and the rate of increase year to year and the DOAJ is used. In each of the successive studies, the samples are improved and the statistical power of the results is better. Björk's results for OA share are calculated here against denominators based on the 'Ulrich's universe' adjusted for variable historical growth rates (the adjusted data set). Adjustments were therefore made (and detailed in the next section), too restrict the article totals to reflect only OA journals listed in Ulrich's and calculating totals based on Ulrich's titles and the adjusted data set that reflects altered growth rates for periods.

Though much of this research relies on Björk 's team, my contribution has been to use my data sheet for the Ulrich's universe of articles and to use OA article totals that reflect the Ulrich's universe. Based on Björk 's figures (2011) combined with some adjustments I have made, we have an average yearly increase of 1.13% for Gold OA over five years. This method applies evidence of an average ratio of 6 to 4 for full Gold versus hybrid/delayed over the 2 Björk studies and the use of this average for 2007 and 2009.

6.3.3: Adjustment of OA totals to reflect Ulrich's universe.

Björk reports article totals for Gold OA based on a sampling of the DOAJ, and has calculated year to year percentage growth. However, the total number of journals listed in the DOAJ is larger than that discovered by using the Ulrich's-OA filter. However, we have article totals based on using Ulrich's-OA to use as a base year and use the percent increase to calculate totals for other years. We also have data on delayed/hybrid Gold OA articles from two Björk studies for years 2006 and 2009, and the ratio of delayed/hybrid to Gold OA journal. By using the ratios and percent increases we can estimate a share which includes Gold OA journals as well as hybrid/delayed, and is consistent to 'Ulrich's universe'. This can be done for years 2006 to 2009. The denominators are based on my data set that originated in Article 50 million and has been adjusted to reflect periods of altered growth.

For the years 2003 to 2005, I elected to use the total given by Björk for Gold OA journals based on DOAJ. There is a trade off between including articles from DOAJ titles that are not Ulrich's and not adding in the unknown number of delayed/hybrid articles. For the years 2002 until the beginning of OA, there is no delayed/hybrid. I elected to continue to use the totals based on DOAJ for consistency of the decrease as it approaches nil. Though imperfect, this process allows one to show an approximate pattern to the increase in Gold OA. It is assumed also that the DOAJ expanded it's comprehensiveness along the way, so the use of DOAJ totals in the early years is another trade-off. Additionally, the better figures based on more direct evidence over the past five years are used to calculate future growth. The past years are used to show the trend up to the present, and to calculate the share of the entire universe, where the smaller totals have less effect.

In summary:

- For 2006-2009, the shares are based on Gold OA journals in Ulrich's plus delayed/hybrid Gold OA.
- For 2002-2005, the shares are based on the DOAJ, where the under-inclusion of delayed/hybrid is offset by overinclusion of non-Ulrich's DOAJ titles.

- From OA inception to 2002, the figures are based on the DOAJ. Under-inclusion would include OA journals not in DOAJ (as DOAJ was becoming established), over-inclusion of journals in DOAJ and not in Ulrich's. For years earlier than the DOAJ the numbers are negligible.
- The denominators are based on the data set from Article 50 million, adjusted for altered growth rates.

When we accommodate a faster growing total universe of articles at 4.5% for the same period where OA totals and yearly increases are given, the result is a more conservative estimate of Gold OA growth than treating the total universe as expanding at 3% per year as it had previous to 2000. The denominator grows faster against the growth share. Calculating based on non-adjusted figures would give a Gold OA average of 1.25%. With an adjusted data set for denominators based on a faster growing universe of articles and totals adjusted to more conservatively reflect 'Ulrich's universe' the denominators are based on, the totals can be calculated as shares, and the increase from year to year can be calculated and then averaged. To reflect recent trends, the a five year average was taken. In recent years distortions from early growth of digitization and OA are less apparent. OA has become normalized in the scholarly system, including delayed/hybrid OA, the digital directories of DOAJ and Ulrich's have become more comprehensive in their coverage, and there is more direct data on OA shares from studies.

The average rate of increase in the Gold OA share inclusive of delayed OA is estimated to be 1.13% per year for the past five years.

Now we can present the results to the research questions at hand.

6.4: Dynamics of current system

Annual OA share: What is the annual share of article output that is Open Access (defined here most broadly as 'freely accessible on the web'? What is the rate of increase of the annual OA share of article output? What will the OA share be in 2035? (the year after the universe of articles will double again from today)

Annual OA share:

For the year 2006: The first study answering this question was published by Björk et. al in 2008, combining an analysis of articles by Green and Gold routes, the estimate is 19.4% for the year 2006.

For the year 2009: Using more precise sampling methods, the share was estimated in Björk 's 2010 article at 20.4% for 2008.

I have not been able to replicate the method for Green, but my research and replication attempts of Gold OA produces data that coheres with that of Björk's for the year 2009. The methods appear to be reliable, with some error between the 2 studies. Björk's forthcoming figures suggest a wider range between 2006 and 2008, faster than is suggested by 8.1 and 8.5%. The same may apply to Green. However, for the sake of defining current as 'towards the end of the first decade of the 21st century, a good estimate is 1 in 5 articles published in current years are now freely available from the web.

By 'best available estimates', currently, approximately 1/5 of global annual output of articles is accessible on the web free of charge.

6.5: Rate of increase of OA Shares:

A few methods can be used to calculate the rate of increase. One is to draw a line from the 2 known figures which amount to a 1% increase over three years and take a simple average. This would yield an average year to year increase of 0.5%. If the increase in overall OA keeps to a pace of .5% per year, the overall OA share in 2035 will be 1/3 of annual output, and will approach 50% near 2070. (This method does not depend on altered growth rates of the total.)

Scenario 1: Gold is increasing faster than Green

The sampling method used for Green OA in the 2 Björk studies was similar, and there is no other direct evidence of the rate of increase of the Green OA share. If Gold reflects the evidence provided by Björk (forthcoming article), and Green is increasing at the .3% suggested by the two samples, we would proceed as follows (this method is also based on the 4.5% increase in articles from 2000).

Given the same average increase in Gold OA of the past five years – 1.13%, and the increase in Green OA between 2006 and 2008 - .3% per year is maintained, the global share of OA in 2030 will be – 60% of annual output, 40% Gold and 20% Green OA. OA annual output will pass 50% around 2027. My data analysis produces figures comparable to other studies. A large study done by the Society of Open Access Publishers estimated the Gold OA share to be roughly 10% for 2008 (SOAP, 2010). My growth curve based on the adjustments discussed and Björk's year to year growth increases has the percentage for 2008 at 9.81%.

Scenario 2: Gold and Green are increasing at the same pace:

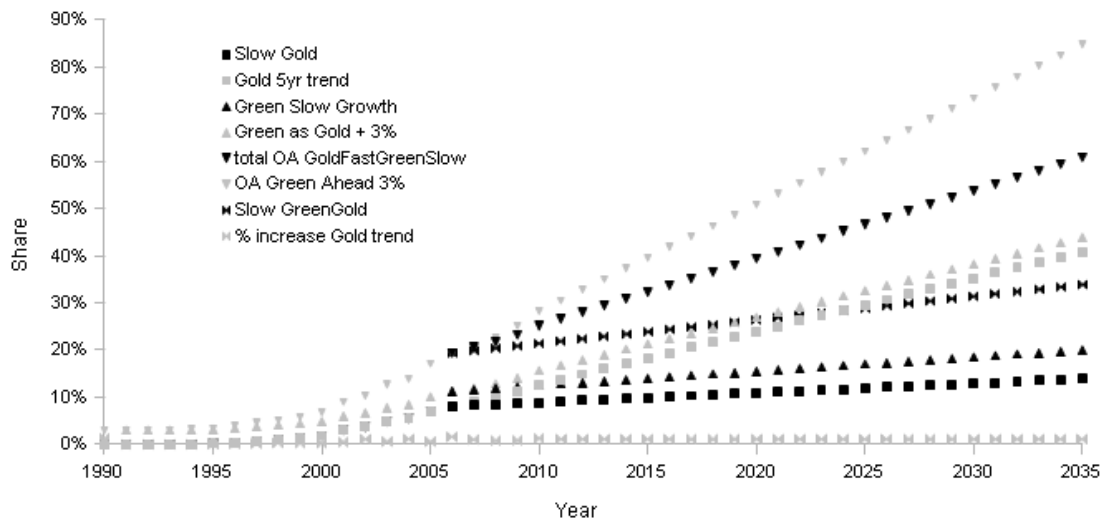


Figure 7: Possible trends for OA Growth

If Gold continues to increase at the same rate as the past five years, and Green stays 3% ahead of Gold, the annual OA output will be 85% in the year 2035. Gold OA will account for 40% and Green 43%. OA will pass 50% in 2020. There is no data as yet to show this, but if it is the case, given the pace of the increase in global OA share that Gold is responsible for, we would be moving very quickly indeed towards 50%, in less than 10 years. We would likely achieve universal OA for current output by 2050 or earlier.

The space for alternative publishing models remains contested and TA publishers may seek to retain it or to convert their model. Within OA, there is contested space for Green and Gold. Within this debate Harnad (too frequently to cite) suggests a preference for Green as complementary to the traditional publishing model and that subsidizing Gold interferes with progress towards universal OA, and Guedon argues for complementary between Gold and Green routes to OA (Guedon, 2004) The question remains as to whether OA at 20% today will soon push against growth limits while TA pushes back, or whether OA growth causes further acceleration and conversion to OA as the share grows.

'The writing on the wall' for the future of publishing may read differently, depending on how one views the environment for publishing in terms of advocacy, policy, markets and social trends, and how one views the degree of and pace of change provoked by digital communication as well as the demographic shift and the culture of the next generation of scholars.

Final calculation – Gold Share for 2011 by article averages, March 2011

As of March 11, there were 702 Gold OA journals that are in the Web of Knowledge AHCI, SCI, from a search in Ulrich's for OA, ASRA, ACHI or SCI or SSCI.

Two tests were performed on the Web of Knowledge to determine an article average of 105 per Gold OA journal in the database from titles from titles derived from an Ulrich's U-JCR-OA search. Only one title from that search was not in Web of Knowledge, a title that had only recently converted OA from Springer to Biomed Central. The first test delivered an average of 105 articles per journal by entering each title, and the second test by entering the first 25 titles delivered an average of 108 articles. From these tests, I we can assume we are in the range of the article average for ISI-indexed journals determined by Björk (106), and there is no difference for Gold OA as a subset within that class.

From the number of titles listed by the U-ISI search in Ulrichs (very closely matching what is in the Web of Knowledge) and using 105 as the average, the number of Gold OA articles that are Web of Knowledge (ISI-indexed) is about 70,000. The average for non-ISI journals in Ulrich's is 25. This average was calculated using a combined sample of 100 journals in Björk 's 2006 study, and 25 I sampled in August 2009. The number of journals listed as OA in Ulrich's is 3458. Multiplied by 25 and added to the 70,000 gives us a total of about 140,000 articles from Gold OA journals. The ratio of Gold OA articles to hybrid/delayed averages across two studies to be 6 to 4. Using this ratio, there would be about 90,000 hybrid/delayed OA articles. The total – 230000 Gold OA articles. Divide this by the number predicted on last 10 years 4.5% increase in articles - 1,677,000 gives us a share of 13.8% to Gold for 2011, fully accessible 2012.

This method is similar to the original one devised by Björk in for base year 2006 – an increase of 5.7% in five years and yearly average close to the other calculation at 1.14%. *The result is slightly higher than what is anticipated by the 5-year average growth for Gold.*

6.6: Implications of Data and Assumptions for Forecasting

The slow growth rate presented in the first scenario reflects neither the trend over the last five years, nor the rise in OA from the inception. The rise in Gold OA is easier to measure, and the predictions for Gold OA depend on whether this trend continues. The predictions for Green OA are more difficult to measure and depend on a correlation with Gold, and on the effect of the rise in mandates. Moreover, Gold OA is a greater reflection of decisions made by market players, a rise in Gold OA reflects decisions made by journal boards, editors and publishers.

As the OA share climbs it appears likely to alter the balance of the zero-sum game between libraries and researchers as purchasers and publishers as sellers. New journals are more likely to launch OA, and though some journals have converted from OA to TA, the more common conversion is towards OA (Björk et. al, 2011).

Further changes may depend on negotiated changes in funding arrangements between library consortia, governments and big publishers trying to adjust to a new reality. The assumption that Green stays ahead of Gold presents a scenario that moves quickly towards this situation. Authors tend to support Green and Gold OA, and are not enthusiastic about article charges (Nicholas, Rowlands and Huntingdon, 2004, Schroter, 2004) . The preference among authors appears to be toward an open system with no fees to share research. Perhaps the most significant result of this thesis, is that a 50/50 situation of OA and TA is likely to arise within 25 years, before the doubling of the article universe unless there is a significant reversal of the trends discovered.

6.7: Titles

Despite the focus on articles, some interesting data for titles was collected in the course of the research:

- 6% of Web of Knowledge (ISI-indexed) titles are now Gold OA.
- 20% of Gold OA titles that are in Ulrich's are ISI-indexed.
- 11% of the 6313 Gold OA titles in the DOAJ are ISI-indexed.
- 33% of titles in Ulrich's are in the Web of Knowledge, these produce 75% of the articles (Björk, 2006) and 95% of citations (Mabe, 2003).

An increase in Green OA puts pressure on non-OA to convert to Gold. An emphasis on supporting Gold OA appears to motivate Green advocates to shift the focus to mandates. The two routes collectively expand the share of publications which are freely available which should enable libraries to be in a better bargaining position with publishers. These factors signify the possibility that growth in either route leads to possible acceleration on the whole and faster growth towards a tipping point, and universal OA. There is no sign of the abating of either Green and Gold routes. It is a question of how soon we will reach a point where it no longer makes sense to drive a business model for publishing based on the old system of copyright relevant to the print era, where articles are monetized and a new system where there is an incentive for quality in the system of scholarly communication but the donated work of scholars is not monetized. In a global audience where authors are the consumers, and are critical thinkers who are compelled towards overturning models of the past, this seems the likely outcome.

6.8: OA Share of All Literature

What share of literature that exists is Open Access? (the share in the absolute static universe – articles treated without regard to date of publication)

Using either the adjusted data set.

Currently (2010-11), 5-6% of all literature that exists is Open Access. Literature that is in the public domain can be digitized and made available by anyone. If we assume all literature that is more than 110 years old (75 years after the death of the original copyright holder – the author, copyright expires for their work) 3-4% % of the literature, all of what is published before 1900 is in the public domain. Thus, if this material were digitized and posted to the web, 10% of the entire corpus of literature that exists would be freely available.

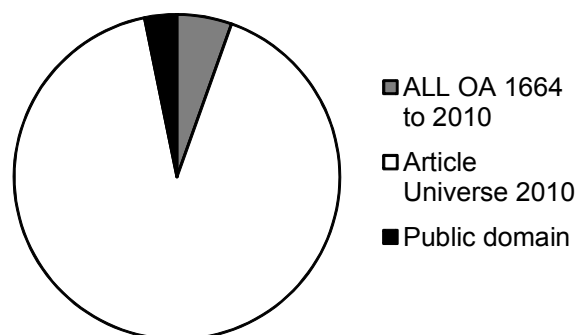


Figure 8: Entire article universe - projected share of Toll, OA and public domain articles in 2010

By 2035, up to 30 to 40% of all literature that exists will be Open Access, and all literature prior to 1925 will be in the public domain, another 2-3%.

OA global metrics – data, method and sources

Metric	Year	Figure	Data method	Source
# of Articles that exist	2010	50 million	Estimate extrapolated based on backward mapping from current output by growth rate.	Jinha
	2012			

OA global metrics – data, method and sources

Metric	Year	Figure	Data method	Source
Annual article output	2006	1.35 million	Sample/databases	Björk
	2009	1.50 million	Sample/databases	Jinha (2010)
	2009	1.48 million	Sample/databases, adjusted data set	Jinha thesis
OA annual share	2006	19.40%	Sample/databases	Björk (2008)
	2008	20.40%	Sample – web/google	Björk (2010)
Gold Share	2006	8.1% (all)	Sample with averages for journal size plus sample of hybrid/delayed	Björk et al (2008)
	2008	8.5% (all)		Björk et al (2010)
	2009	7.7% (not including hybrid/delayed)	Large sample/databases	Björk et al (2010)
	2009	11% (all)	Based on Björk et. al data, adjusted for consistency to 'Ulrich's'. plus hybrid/delayed by ratio (Gold OA journal-hybrid)	Jinha (thesis data)
Green Share	2006	11.30%	Sample/databases	Björk (2008)
	2008	11.90%	Sample/database	Björk (2010)
% of author's archiving	Several	15.00%	Expert opinion	Harnad (frequent)
	2002	23%	Sample	
% of authors that archived in any way, one time	2005	42.00%	Survey	Swan and Brown (2005)
Ulrich's titles (ASRA)	2010, August	26743	Ulrich's ASRA filter	Jinha thesis
	2011, Feb	28,230	Ulrich's ASRA filter	

OA global metrics – data, method and sources

Metric	Year	Figure	Data method	Source
U -OA	2011, Feb.	3471	U-OA filter	Jinha thesis
	2011	6205	DOAJ	
WoS titles	2010, Aug	10041	U-ISI	Jinha thesis
	2010, Aug	10131	JCR list	
	2011, Feb	12751	JCR list	
OA - ISI titles	Aug 2010	385	U-OA-ISI filter	Jinha thesis
# of OA-ISI Articles	Aug 2010	40500	Web of Knowledge exact count	Jinha thesis
average articles/title U-ISI-OA	Aug 2010	105	Web of Knowledge articles/# of OA-JCR titles used to count	Jinha thesis
Average # articles non-ISI	2006	26.2	Sample (U-ASRA excluding ISI journals from the sample)	Björk et al (2006)
	Two samples aug 2010 and 2006 (Björk et. Al, 2008)	25	Combined thesis sample with Björk (U-ASRA excluding ISI journals from the sample)	Jinha thesis plus Björk et.al (20062)
# of new journals in Ulrich's (launched since 2009)	2011, Aug	2728	U-ASRA Start Year 2009-2011	Jinha thesis
new OA (as above OA)	2011, Aug	727	U-OA Start Year 2009-2011	

Table 5: Metrics of the global scholarly system

6.9: Backdated OA?

On a final note, in my research I encountered a few converted OA journals. Most of these journals only posted articles OA since the time they converted, however at least one made their publications OA all the way back to the founding of the journal in the 1930s. From a static universe point of view (articles treated equally and ahistorically - regardless of date of publication), backdated OA would increase the size

of OA in the absolute article universe and it would push back the start date for OA. Anecdotally, backdated OA still seems too rare to factor into the massive totals so it is more of a suggestion. Backdating would also be a possibility on the Green route for living authors who would voluntarily archive older works, perhaps as a legacy activity.

Much emphasis has been put on current OA availability, but older articles are important to some researchers. Swan and Brown's 2005 survey on self-archiving asked the question. 4% were interested in articles up to 50 years, old 17% searched literature over a 20 year time span. A small minority - 5 authors representing 1% of the sample reported that most articles consulted were on average more than 50 years old. 8 % were very current, most literature consulted was less than 20 years old. The biggest groups (35% and 34%) consulted literature 5 to 10 years old.

If the goal is universality in this regard, there may be some interest in backdating OA by mandate or by publisher. The main difficulty is that most of the literature of the 20th century is licensed by copyright typically with a life span of 75 years plus the life of the holder. The literature of the Big Science period 1945 to 1977, would only enter the public domain between 2050 and 2100, at which time it will may only be of interest to that 1% of historical researchers. Meanwhile, all of the literature currently still published under the copyright model in subscription journals may have to wait longer unless some kind of settlement is reached (as with Google Books).

8. Development and Globalization - Africa in focus

This is early evidence for the globalization of scholarship occurring with two likely drivers. The first is the ongoing rise of India and China, and the second is online availability. Though South-East Asia is the key region that diverts the greatest share from the North, in the context of absolute overall growth Africa is still increasing its share of researchers and publications. This open data is from the UNESCO Institute for Statistics⁷.

Share of Researchers

Region	2002	2007
Africa	2.2%	2.2%
South Africa	0.2%	0.3%
Sub-Saharan Africa (excl S.A.)	0.5%	0.6%
Arab States in Africa	1.40%	1.40%

Table 6: Share of researchers 2002 and 2007- Africa

Researchers per million

Region	2002	2007	Increase	% inc
Africa	150.2	164.3	14.1	0.04
South Africa	311.4	392.9	81.46	0.12
Sub-Saharan Africa (excl S.A.)	49.4	57.5	8.08	0.08
Arab States in Africa	444.1	477.1	33.03	0.04

Table 7: Researchers per million 2002 and 2007 - Africa

Publications

	Articles ISI			Annual avg increase	World Share		
	2002	2008	Change		2002	2008	Change
World	733,305	986,099	34.50%	0.058	100	-	-
Developed	617,879	742,256	20.10%	0.03	84.3	75.3	-9
Developing	153,367	316,742	105.90%	0.177	20.9	32	11.1
LDC	2,069	3,766	82.00%	0.14	0.3	0.4	0.1
Africa	11,776	19,650	66.90%	0.112	1.6	2	0.4
South Africa	3,538	5,248	48.30%	0.08	0.5	0.5	0
Other SSA	3,399	6,256	84.10%	0.140	0.5	0.6	0.1
Arab in Afr	4,988	8,607	72.60%	0.12	0.7	0.9	0.2

Table 8: Publications 2002 and 2008 Africa

⁷<http://www.uis.unesco.org/>

While the North has dropped 9% of its share of global researchers, Africa has increased its share slightly. Global availability and access in Africa to African journals has improved dramatically. Using Ulrich's to search by publication country, there are 595 journals published on the continent. AJOL provides free access to African institutions for all its titles. AJOL currently carries over 400 journals. Using Ulrich's to test for OA, 110 titles, or 27% of these journals are Gold OA.

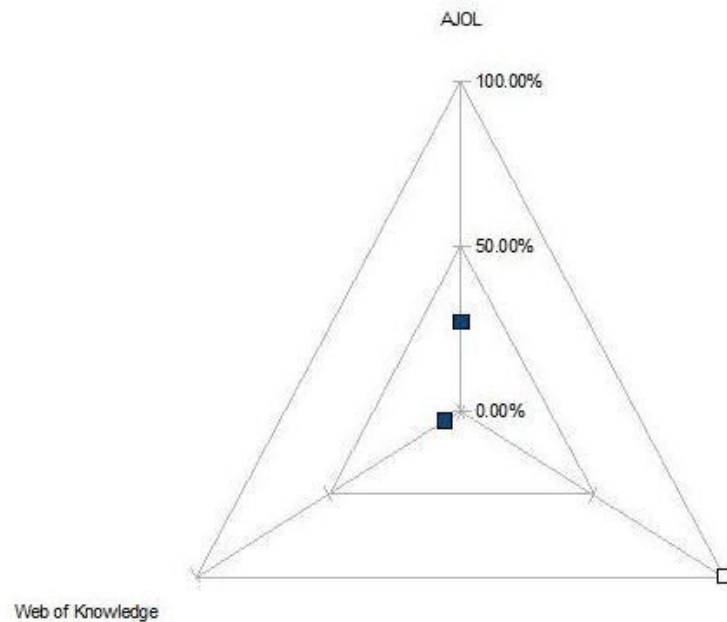


Figure 9: % Gold OA AJOL vs. Web of Knowledge

INASP-Perii, provides access to negotiated titles institutions can choose from. The access level varies by the agreements. EIFL conducts national-level bargaining, with access opportunities varying. These overlap with the UN's programs. The following is data collected on these programs.

01/03/11		
Organization	Programme	Material and services
World Health Organization - WHO	HINARI	7500 “information resources” mostly journals
Food and Agriculture Organization – FAO	AGORA	1900 journals
United Nations Environment Programme - UNEP	OARE	1300 journals
International Network for the Availability of Scientific Publications - INASP	- Programme for the Enhancement of Research Information – PERii. - African Journals Online - AJOL	23 000 journals & 7000 e-books (to be selected from by library consortia), citation & bibliographic databases, document delivery
eIFL	Country-level consortia negotiations	Per country/institution by agreements eg. http://www.eifl.net/oxford-journals-collection

Table 9: Summary of development-oriented journal access programs

Finally, the directories to search OA journals and repositories continue to grow in size and scope of resources.

Directory of Open Access Journals (DOAJ)	6882 journals (Aug, 2011)
Directory of Open Access Repositories (Open DOAR)	2000+ repositories (Aug, 2011)

Table 10: Directories of open access sources

The last 10 years shows small growth or holding steady for share of researchers and publications in Africa, while on the whole we are experiencing the fastest growth in history at 4.5%. The share of publications meanwhile dropped by a tenth in the advanced developed countries. This is not a drop in research output in the North, but reflects the diffusion of participation globally. This is a significant change for scholarly communication already. It will take time for development effects to show up, but not much longer for access gaps to close toward an open system. Björk's studies have more precise breakdowns, but demonstrate that biomedical literature is at the forefront of Gold OA with more than 25% of titles available free from the publisher, while HINARI is the most populated concessional program. Physics and similar sciences lead the way in Green OA. Humanities tends to trail last despite publishing the earliest OA

journals, and social sciences sits in the middle. Despite the relevance of OA to development, international development studies is far from leading the way in OA availability. Taken together, my best estimate for any researcher who is able to take advantage of both OA and programs mentioned, is that ***today one should be able to access at least 60% of current global literature freely at a university in the South***, and access is better for some disciplines than others.

9. Conclusion – Global Open Knowledge

Currently 20% of global literature is available for free online, that share is increasing towards 50% rapidly which suggests a trend a global, open system for scholarship. With respect to development, the university communities in the South are best positioned to interpret the significance of the changes afoot with regards to development, and the same is true through the South. In Africa, university communities that can fully take advantage of concession and bargaining programs, can already access a majority of top journals, a stark difference from 10 years ago. Despite the disjointed complexity of the transitional digital scholarship system, Africa universities are moving from near zero to potentially beyond 60% access (if all OA and other programs are taken advantage of) to current global literature in a decade.

What one decides to do with this is up to the researcher. The enlightened self-interest of the next generation would be served by both the North and South contributing more open access knowledge to the global commons. My own education of the world would be best served by the transmission of knowledge by my peers who are immersed in their local society and culture and have access to all literature. Out of mutual interest, self-interest, the interest in the survival and integrity of humanity, it is easy to recommend that the world embrace an open system of scholarly communication.

What may be more appropriate is to focus on the shift this suggests for how universities in the North, like University of Ottawa, might view partnership in development. We are moving into a new era in scholarly communication, as demonstrated by the data. The shift is towards greater global participation. This is not a decline in the activity of the North, but an overall increase in research activity, with the shares increasing in the global South and decreasing in the North. Again, the suggestion is to embrace open access journals and repositories to pave the way for transition, out of reciprocal enlightened self-interest (Tocqueville, 1899).

There is a demographic shift occurring. The Baby Boom, a phenomenon of the global North, arguably the most powerfully global society in history who shaped the late 20th century, are reaching retirement age. By contrast, many countries of the global South are composed of populations where youth are the largest cohort. These global youth more equitably access information and knowledge. It is those that can access information and knowledge that shape social thought, so social thought is more and more globally shaped. The demographic shift in can most simply be described as a shift from the values of the Northern Baby Boom, to the values of a global, digital generation. We recently witnessed the Arab Spring, but it is not difficult to see the Arab spring as a part of a whole, of a global spring.

The generation of the global spring will have available to them a massive amount of information, but the degree to which they get quality information and are able to navigate, filter and retrieve it, and the degree to which scholarship plays a role in channelling quality information will be determined by the behaviour of scholarly communities in the North and South. The pace and delays of this will be determined by the way we treat Open Access, in initiatives we take and policies we make. The data here demonstrates that a 50/50 situation, where the annual output of scholarly articles that is freely available reaches 50% is approaching rapidly.

Perhaps the legacy of the generation of scholars to retire this decade, will be to retire the price barrier to research, and allow the generation ahead to inherit the entire global heritage of knowledge. They will certainly need good information. This would be my recommendation to my professors and university administrators, to open scholarship. Engage the next generation of scholars in developing a new normal that fits their medium – open access to research, the openness paradigm and the realization of the Access Principle. The entire corpus on every researchers' desktop and immediately accessible. The result of this would be equality among seekers of knowledge rich or poor, and the creation of the *university*.

Chapter 7. Bibliography

Aronson, Barbara, 2004. "Improving Online Access to Medical Information for Low-Income Countries." *New England Journal of Medicine* 350(10): 966-968.

Arunachalam, Subbiah, 2004. "India's march towards open access." *Science and Development Network*
<http://www.scidev.net/en/opinions/indias-march-towards-open-access.html>.

Arunachalam, Subbiah, and Peter Suber, 2005). "Open Access to Science in the Developing World." *World-Information City*.
<http://www.earlham.edu/~peters/writing/wsis2.htm>.

Banting, K. 1995. "Social Policy Reform in Canada." *In Social Policy in a Global Society: Parallels and Lessons from the Canada-Latin Experience, edited by Daniel Morales-Gómez, and A. Mario Torres, 25-105. Ottawa: International Development Research Centre.*
http://web.idrc.ca/es/ev-69130-201-1-DO_TOPIC.html

Birdsall, Nancy, 1996. "Public spending on higher education in developing countries: Too much or too little?" *Economics of Education Review* 15(4): 407-19.
http://upi-yptk.ac.id/Ekonomi/Birdsall_Public.pdf.

Björk Bo-Christer, Welling Patrik, Laakso Mikael, Majlender Peter, Hedlund Turid, and Gudni

Gudnason, 2010. "Open Access to the Scientific Journal Literature: Situation 2009." *PLoS ONE* 5(6):

Björk, B Bo-Christer., Roos, Annikki. and Mari Lauri, 2008. "Global annual volume of peer reviewed scholarly articles and the share available via different Open Access options." *Proceedings of the ELPUB2008 Conference on Electronic Publishing.* Toronto.
<http://oacs.shh.fi/publications/elpub-2008.pdf>

Brown Patrick O., Cabell Diane, Chakravarti Aravinda, Cohen Barbara, Delamothe Tony, Eisen Michael, Grivell Les, et.al., 2003. Bethesda Statement on Open Access Publishing.
<http://www.earlham.edu/~peters/fos/bethesda.htm>.

Brown, Harcourt, 1972. "History and the Learned Journal." *Journal of the History of Ideas* 33: 365-78.
<http://www.jstor.org/stable/2709041>.

Chan, Leslie, and Sely Costa, 2005. "Participation in the Global Knowledge Commons: Challenges and Opportunities for Research Dissemination in Developing Countries." *New Library World* 106(1210/1211):141-163.

Chan, Leslie, Kirsop, Barbara and Subbiah Arunachalam, 2006. "Open Access Archiving: The Fast Track to Building Research Capacity in Developing Countries, Working Papers. id:527, esocialsciences.com.

Clapshaw, N.C., 2007. The quality of higher education, internet and computing technologies: Exacerbating or lessening the divide across countries? An analysis at three levels: National, institutional and classroom. Doctoral Dissertation Leadership and Policy Studies, Graduate School of Vanderbilt University, Nashville TN.

Cohen, A. 1968. "Development in Africa: The problems of today." *African Affairs* 67(266): 44-54.

Colle, Royal D. 2005. "Building ICT4D capacity in and by african universities." *International Journal of Education and Development using ICT* 1(1):101-107.

Colle, Royal D., and Raul Roman. 2003. "ICT4D: A frontier for higher education in developing nations." *African and Asian Studies* 2(4): 381-420.

Dowuona, M. 1966. "Training for development in Africa." *African Affairs* 65, (259): 135-147. e11273. doi:10.1371/journal.pone.0011273.

European Commission, 2006. Study on the Economic and Technical Evolution of the Scientific Publications Market in Europe. Science and Society.
http://ec.europa.eu/research/science-society/pdf/scientific-publication-study_en.pdf

Gargouri Y., Hajjem C., Larivière V., Gingras Y., Carr L., Brody T., and S. Harnad. 2010. "Self-Selected or Mandated, Open Access Increases Citation Impact for Higher Quality Research." *PLoS One* 5(10):e13636.
<http://www.plosone.org/article/info%3Adoi%2F10.1371%2Fjournal.pone.0013636>.

Guédon, J. 2004. "The 'green' and 'gold' roads to open access: the case for mixing and matching." *Serials Review*. 30(4):315-328. doi:10.1016/j.serrev.2004.09.005.

Hafner, Katie and Matthew Lyon, 1996. Where wizards stay up late: The origins of the internet. Simon & Schuster, New York.

Harle, Jonathan, 2010. "Growing Knowledge: Access to research in east and southern African universities." London: Association of Commonwealth Universities. Arcadia.
<http://www.acu.ac.uk/publication/download?id=291>

Harnad, S. "I. Overture: The Subversive Proposal." In *Scholarly Journals at the Crossroads: A Subversive Proposal for Electronic Publishing: An Internet Discussion about Scientific and Scholarly Journals and Their Future*, edited by Ann Shumelda Okerson and James J. O'Donnell, 1-2. Washington, DC: Office of Scientific & Academic Publishing, 1995. <http://eprints.ecs.soton.ac.uk/3351/1/toc.html>.

Harnad, S. 2007. "The "Double-Pay"/"Buy-Back" Argument is Invalid." *Open Access Archivangelism*, September 9.
<http://openaccess.eprints.org/index.php?/archives/290-guid.html>.

Harvard University Committee on the Objectives of a General Education in a Free Society. 1945. *General education in a free society: Report of the Harvard Committee*. Cambridge: Harvard University Press.

Houghton J.W. 2009. *Open Access – What are the economic benefits? A comparison of the United Kingdom, Netherlands and Denmark*. Brussels: Knowledge Exchange.

International Telecommunication Union. 2011. "New ITU report shows global uptake of ICTs increasing, prices falling."
http://www.itu.int/newsroom/press_releases/2010/08.html.

Jinha, Arif E., 2010. Article 50 million. The number of scholarly articles in existence. *Learned Publishing*. 23(3): 258-263. doi:10.1087/2010030.

Kapur, D., and M. Crowley. 2008. *Beyond the ABC's: Higher Education in Developing Countries*. Center for Global Development Working Paper No. 139. Available at SSRN:<http://ssrn.com/abstract=1099934>.

King, D.A. 2004. "The scientific impact of nations." *Nature* 430(15 July 2004), 311-316.

King, David, A., 2004. "The scientific impact of nations." *Nature*. 430, 311-316.

Kossinets, Gueorgi, and Duncan J. Watts. 2009. "Origins of Homophily in an Evolving Social Network." *American Journal of Sociology* 115:405–50. Accessed February 28, 2010. doi:10.1086/599247.

Laakso, M., Welling, P., Bukvova, H., Nyman, L., Björk B.-C., and Turid H. 2011. "The development of open access publishing from 1993 to 2009." *PLoS One*. 6(6):e20961.
<http://www.plosone.org/article/info%3Adoi%2F10.1371%2Fjournal.pone.0020961>.

Lulat, Y. G. M. 1982. "Political constraints on educational reform for development: Lessons from an African experience." *Comparative Education Review* 26(2): 235-53.

- Mabe, Michael, and Mayur Amin, 2001. "Growth dynamics of scholarly and scientific journals." *Scientometrics*. 51: 147–62. doi:10.1023/A:1010520913124.
- Mabe, Michael. 2003. "The growth and number of journals." *Serials*. 16: 191-197.
- Morris, Sally, 2007. "Mapping the journal publishing landscape, how much do we know?" *Learned Publishing* 20(4): 299-310. doi: 10.1087/095315107X23965.
- Murphy, S. 2011. "Guerilla Activist' Releases 18,000 Scientific Papers." *Technology Review*, July 22. <http://www.technologyreview.com/computing/38112/>.
- North, D. 2005. *Understanding the Process of Economic Change*. Princeton: Princeton University Press.
- Open Access at the Max Plank Society, 2003. "Berlin Declaration on Open Access to Knowledge in the Sciences and the Humanities." Last modified October 22,2003. http://oa.mpg.de/files/2010/04/berlin_declaration.pdf.
- Open Access Scholarly Information Sourcebook (Open OASIS), 2010. http://www.openoasis.org/index.php?option=com_content&view=article&id=552&Itemid=375.
- Ouya, Daisy, and Pippa Smart, 2005. "Open Access survey of African published journals." http://www.codesria.org/IMG/pdf/Daisy_Ouya_Paper.pdf
- Rowlands I., Nicholas D. and P. Huntingdon. 2004. "Scholarly Communication in the Digital Environment: What do Authors Want? Findings of an International Survey of Author Opinion: Project Report." London: Centre for Information Behaviour and the Evaluation of Research .
- Schroter, S, Tite, L., and R. Smith. 2004. *British Medical Journal* 330:756. doi: 10.1136/bmj.38359.695220.82.
- Sen, A. 1999. *Development as Freedom*. New York: Alfred A. Knopf.
- Smith, Helen, Bukirwa, Hasifa, Mukasa, Oscar, Snell, Paul, Adeh-Nsoh, Sylvester, Mbuyita, Selemani, Honorati, Masanja, Orji, Bright, and Paul Garner, 2007. "Access to electronic health knowledge in five countries in Africa: a descriptive study." *BMC Health Services Research*. 7(1):72. doi:10.1186/1472-6963-7-72.
- Smith, M., Engler, N.J., Christian, G., Diga, K., Rashid, A. and K. Flynn-Dapaah. 2008. "Open ICT4D [WorkingDraft]." Ottawa: International Research Development Centre. http://web.idrc.ca/uploads/user-S/12271304441Open_ICT4D_Draft.pdf.

St. John, N. Review of *The New Media and Education: Their Impact on Society*. by Rossi, P.H. and B.J.

Biddle. *American Journal of Sociology*. 73(2):255-256. September, 1967.

Stiglitz, Joseph, E. 1999. "Knowledge as a Global Public Good". In: *Global Public Goods* edited by Kaul, Inge, Grunberg, Isabelle and Marc Stern. New York: Oxford University Press.

Stiglitz, Joseph, E. 2000. "Scan globally, reinvent locally knowledge infrastructure and the localisation of knowledge." In: *Banking on Knowledge: The Genesis of the Global Development Network* edited by Diane Stone, 23-43. London: Routledge.

Suber, Peter, 2003. "The Taxpayer Argument for Open Access." *SPARC Open Access Newsletter*. September 4.

<http://www.earlham.edu/~peters/fos/newsletter/09-04-03.htm>

Suber, Peter, 2004. *SPARC Open Access Newsletter*. September 2.

<http://www.earlham.edu/~peters/fos/newsletter/09-02-04.htm>.

Suber, Peter, 2008. "Gratis Libre and Open Access." *SPARC Open Access Newsletter*. August 8.

<http://www.earlham.edu/~peters/fos/newsletter/08-02-08.htm#gratis-libre>.

Swan, A. and S. Brown. 2005. *Open access self-archiving: An author study*. Cornwall: Key Perspectives. <http://eprints.ecs.soton.ac.uk/10999/>.

Tenopir, Carol W., and Donald W. King, 2000. *Towards Electronic Journals*. Washington DC: Special Libraries Association..

Tenopir, Carol W., and Donald W. King, 2009. "The growth of journals publishing". In *The Future of the Academic Journal*, edited by Bill Cope, and Angus Phillips. Oxford: Chandos Publishing/Woodhead Publishing Ltd.

Tocqueville, Alexis de. 1899. *Democracy in America* (in the public domain). Translated by Henry Reeve. Charlottesville, NC: University of Virginia.

http://xroads.virginia.edu/~HYPER/DETOC/toc_indx.html

University of Ottawa (2009). Collections – uOttawa Library. Retrieved July 6th 2011 from

<http://www.library.uottawa.ca/content-page.php?g=en&s=biblio&c=abt-collections>

W3C. 2011. "Mobile Web for Social Development". <http://www.w3.org/2008/MW4D/>.

Ware, Mark, 2006. *Scientific publishing in transition: an overview of current developments*. Bristol: Mark Ware Consulting.

Willinsky, John, 2006. *The Access Principle: The Case for Open Access to Research and Scholarship*. Cambridge: MIT Press.

http://mitpress.mit.edu/books/willinsky/TheAccessPrinciple_TheMITPress_0262232421.pdf

Willinsky, John, Jonas, Randall and Kiven Charles Wirsiy (2005). "Access to Research in Cameroonian Universities." *Electronic Journal of Information Systems in Developing Countries* 21(2):1-15.

<http://www.ejisd.org/ojs2/index.php/ejisd/article/viewFile/137/133>.