University of Southampton

Faculty of Natural and Environmental Sciences

School of Chemistry

An Examination of Academic Publishing through the Lens of Disintermediation

by

Richard William Fyson

22 October2013

Supervisors: Dr Simon Coles, Prof Les Carr

Advisor: Prof Jeremy Frey

This report uses the ACM SIGCHI Conference Proceedings reference style.

# Abstract

The problems associated with academic publishing have been the subject of much debate since the arrival of the Web and awareness of its potential to revolutionise traditional communication models has grown. Numerous solutions have been presented but the extent to which these have changed the manner in which researchers communicate their findings has been limited. This research aims to reconsider many of the factors that are present within academic publishing and, by building a holistic picture of all the stakeholders and the interactions that take place among them, proposes a new approach to dissemination, which should both be compatible with stakeholder interests, but also bring the Web and its affordances more pertinently into the realm of scholarly discourse. Having gained an understanding of the contributory factors and implications of publishing, we arrive at the concept of disintermediation and think about how this might apply in the scholarly context. Whilst it becomes clear that such a change has the potential to bring a raft of benefits to the academic community, it also represents a significant paradigm shift and the re-evaluation of the role of a number of stakeholders, all of which raises many barriers to change. Disintermediation also raises many questions, principally how would this new approach to dissemination work in reality and how would it operate alongside existing models of publishing which are unlikely to fade away in the near future. Thus this report concludes with two threads of original contributions: a discussion as to how disintermediation may be implemented and the proposal of a number of tools that aim to investigate user attitudes towards the principle of disintermediation and their effectiveness of eroding the barriers.

Contents

[Abstract 2](#_Toc367889170)

[Project Aims 5](#_Toc367889171)

[Overall Aims 5](#_Toc367889172)

[2013/2014 Aims 6](#_Toc367889173)

[Introduction 8](#_Toc367889174)

[Background 9](#_Toc367889175)

[Academic Publishing 9](#_Toc367889176)

[Overview 9](#_Toc367889177)

[Academic Publishing and the Web 10](#_Toc367889178)

[Disintermediation 13](#_Toc367889179)

[Introduction to Disintermediation 13](#_Toc367889180)

[Stakeholders of Academic Publishing 14](#_Toc367889181)

[Previous Work 23](#_Toc367889182)

[Open Access 25](#_Toc367889183)

[Reward and Recognition 28](#_Toc367889184)

[Methodology and Hypothesis 31](#_Toc367889185)

[Disintermediation in Practice 33](#_Toc367889186)

[A Model for Disintermediation 40](#_Toc367889187)

[Online Academic Publishing Framework 43](#_Toc367889188)

[Conclusion 43](#_Toc367889189)

[Widgets 44](#_Toc367889190)

[Rate My Data 47](#_Toc367889191)

[ChemConnections 50](#_Toc367889192)

[Image Redactor 54](#_Toc367889193)

[Conclusion 58](#_Toc367889194)

[Discussion 58](#_Toc367889195)

[Future Work 59](#_Toc367889196)

[References 61](#_Toc367889197)

[Appendix A – 9 Month Report 66](#_Toc367889198)

[Appendix B – AltAC Executive Summary 115](#_Toc367889233)

# Project Aims

## Overall Aims

As the project has progressed, its aims have shifted and adapted as the research introduces new ideas. The overall aim of the project has remained the same however: to investigate new methods and approaches that may be employed so as to encourage open and efficient dissemination of research. Whilst it is beyond the scope of any single project to completely change the manner in which researchers go about disseminating their work, it is hoped that the outcomes of this project will form one step in not just improving accessibility to research, but also in broadening the type of research outputs that are made available. An overarching philosophy of the project is the assumption that researchers would rather spend more time and effort working on their research rather than in its dissemination to the wider community. Thus a further aim of this project is to examine new approaches to make the processes of communicating research more convenient for researchers themselves, whilst being wary of the potential impact this may have on a researcher’s attempts to secure publications and funding.

Therefore, the focus of the project is the area of scholarly discourse, loosely defined as the process through which researchers communicate their results and findings to one another so that they may in turn build upon each other’s results to arrive at new research proposals and continue the cycle. This project aims to examine how this process can be improved for the benefit of the researcher’s themselves and for the wellbeing of the subjects and disciplines being studied.

Whilst the project aims to look at how it may be possible to improve scholarly discourse to assist the progression of research, it is also important to consider that academic publishing is about more than just the communication of research outcomes. The processes involved fulfil numerous requirements; most notably quality assurance of the research communicated, the bestowing of recognition upon researchers who make valuable contributions in their field, and an indication as to the wider academic, economic and societal impacts a research project may realise. As a result of these various functions and the manner in which different stakeholders depend on one another, the area of academic publishing is a delicate one and to impose any changes requires careful consideration of the stakeholders’ many roles and desires. Thus the focus of this project is on the researcher as one of many stakeholders, all of whom add value to the research “market”, but none of whom are as essential as the researcher.

With these different possible perspectives on scholarly discourse and academic publishing taken into consideration, the overall aim of the project is to improve scholarly communication from the perspective of the researcher, using theories and studies to ascertain how the system works for them presently and how it might be improved upon. One of the main ideas to have emerged from the project thus far is the concept of disintermediation, a process of removing intermediary figures that has arisen in many sectors that have been disrupted by the influence of the Web. In the publishing sector, the intermediaries can be identified as the journal publishers, stakeholders who, whilst playing an important role in helping research reach its target audience and bestowing recognition upon researchers, play a role which has the potential to be superseded by the Web, leaving the journal publishers in a position of adding little value that cannot be found elsewhere. This project will therefore examine what impact the Web has already had on academic publishing and explore the opportunity of making scholarly communication more open, usable and convenient for researchers through the lens of disintermediation.

## 2013/2014 Aims

The aims of this project throughout 2013 are to build upon the foundation that has been established by the research conducted in 2011 and 2012 to evaluate some of the ideas that have been proposed. Initial research used ideas from game theory, collective behaviour theory, knowledge management and an original model of how knowledge is communicated in a discipline, supplemented by some initial studies that asked early career researchers of their opinions on different approaches to publishing and how technology facilitates the scholarly communication process, to devise the “Dissemination through Disintermediation” [15] idea (for more information, see the 9-month report found in Appendix A). Further research then looked at recognition and impact and how the manner in which these operate may influence the direction taken by the disintermediation approach.

If the “Dissemination through Disintermediation” approach to publishing, dubbed AltOA in a paper submission to the WebSci 2013 conference [16], is to be successful then a fuller understanding of the current research environment is required, to identify how it serves researchers at present and where there are opportunities to improve upon it, and it is this task that occupies the future research agenda. A number of case studies will be drawn in the field of chemistry to explore different aspects of scholarly dissemination, allowing for an investigation in to the role of researchers and other stakeholders and how by understanding stakeholder motivations, it may be possible to alter or introduce certain processes to produce desirable outcomes. The research will aim to understand how academic publishing may change and adapt to better suit the needs of the researcher, providing an indication as to not only how the researcher may make their day-to-day research easier and help to better their career, but also what a new approach may require from them, with respect to time and effort they need to put into the system and potential costs and risks they may need to bear.

Having thought about the problem of the perspective of the researcher, yet also realising that in reality academic publishing is a game of many players each with their own wants and needs from the system, the research will go on to look at how academic communication can be improved upon as a whole, liberating the work conducted by researchers so that it can be used as an academic currency that can be utilised by any stakeholder as they see fit. Taking the disintermediation approach further and thinking about academic publishing from the perspective of open data, further research will look into what aspects of the process can be made open, how data may be sourced and what benefits may be expected to arise if the entire process of research is made open and transparent. It is hoped that the benefits of such an approach to organising research will allow for a more flexible research environment, that is able to more easily organise research projects to satisfy emerging research areas and smaller scale projects that produce impactful results in a short timescale.

# Introduction

Initially conceived as a tool for facilitating academic collaboration, the Web has since gone on to be shaped by both social and technical forces into a number of unforeseen directions. It continues to impact upon academic collaboration and communication yet due to the complex network of stakeholders that has formed around this industry; the Web’s impact may be considered to be unsatisfactory, with much research still inaccessible behind paywalls or simply completely unavailable. The Web provides the capacity to not only communicate ideas to an audience of any size, but to allow audiences to form around the content they search for, and as such it is a powerful tool for dissemination. Despite these affordances however, it could be argued that a majority of the academic community have failed to grasp the advantages on offer, with academic dissemination plagued by many of the problems it has faced for decades, with the Web often showing its potential to revolutionise communication in other sectors rather than academia. The purpose of this project is to examine why the Web has thus far had little disruptive impact in academic publishing, merely allowing the processes that happened previously to continue albeit at a faster pace, and to ultimately propose new approaches to scholarly discourse that aim to complement the existing systems in place whilst taking full advantage of the potential the Web permits.

Having devised new approaches based upon theory and qualitative data collected in interviews about the publishing process, the project will then proceed to evaluate the ideas. An opportunity to test such ideas has been made available in the field of chemistry, where it is hoped new methods may be employed to streamline the publishing process, making the tasks involved more convenient and useful for researchers. Chemistry provides an ideal setting to test new approaches to scholarly discourse, being a subject with data in a myriad of formats, created in various scenarios with different numbers of stakeholders and degrees of confidentiality, all of which provide a number of variables with which to examine and test new ideas. To this end a number of case studies are proposed that will examine the different types of data and knowledge that are communicated in chemistry and the reasons behind the interactions that occur, in a bid to illustrate how the new ideas proposed may provide a better and complementary service to the systems already in place.

# Background

## Academic Publishing

### Overview

Before engaging in a deeper discussion of the issues pervading academic publishing and possible steps on a pathway to resolving those issues, it is important to define precisely what is meant by academic publishing in the context of this project. Taken in its simplest form, it can be assumed to mean the release of research outcomes for access to a wider community of scholars, who can then take inspiration and knowledge previously established by these authors, to create something new, which will in turn be made available to repeat the cycle anew.

This practice as it is conducted today, was established with the publication of one of the first academic journals, *Philosophical Transactions of the Royal Society*, published in 1665 [29]. It allows authors to disclose their findings in a concise and effective format, a format that has ultimately allowed research to flourish in a pre-Web environment. Having been quite successful at disseminating research, journals and their publishers have, over the years come to define the academic community’s perception of publishing and now hold considerable influence in scholarly discourse. As a result other aspects of research have become tightly entangled with publishing, such that publishing is about more than just access to research, but concerns other issues such as quality assurance and ranking mechanisms. The *Philosophical Transactions of the Royal Society*, also happened to be the first journal to introduce peer review as a method for ensuring its content was scientifically sound, and whilst this approach tended towards being “the exception rather than the rule”, more recently peer review has become a fundamental aspect of nearly all scientific publications [35]. Similarly, journal publishers have also become a fundamental stakeholder in the process of how the community ranks the significance of its members’ contributions to their field, with big brand names such as *Nature* or *Science* giving authors who publish within a prominent standing amongst their peers. Journal impact factors are yet another facet, that extends the consequences of publishing in a manner that influences the direction of research on an institutional level [54]. However it should be noted that these extra, value adding services - useful for the research community as a whole - which would need to find an answer to many of the problems that publishing solves if publishing restrained itself to just dissemination of research - do come at a cost, with expensive subscription fees permissible as a result of increasingly prestigious journals.

Thus, with the development of these facets and the importance they now hold for publishing academics, it is possible to come to the conclusion that publishing is about more than the mere dissemination of knowledge amongst colleagues. Publishing concerns itself with reviewing work, recognising researchers and identifying impact, which in turn brings in yet more aspects of academia such as how researchers are rewarded for their efforts, how funding is allocated and how institutions are ranked. Publishing as referred to in this project encompasses all of these issues and as such becomes an increasingly difficult performance to interpret, predict and ultimately direct. It should be noted that the facets considered here are very much grounded in the social interactions that take place amongst the stakeholders, rather than concerning some of the more technical elements of publishing. For example, the issue of ensuring the preservation of articles has already been thoroughly discussed [25,56].

Figure 1 - The many facets of publishing. The roles of journal publications have become increasingly influential over the years and now have implications for other aspects of research. The monopolistic role of journal publications however means each of these facets does not work at optimum efficiency or effectiveness.

### Academic Publishing and the Web

As a result of the aforementioned, far-reaching consequences of publishing, the model that was first instantiated in 1665 exists to this day more or less unchanged; yet whilst it permitted a flourishing of research and research communities before the Web, the same model could be argued to be constricting the flow of research and preventing the research community from fully embracing the affordances offered by the Web. The Web has become the primary communication medium for scholarly discourse, allowing its users to communicate with one another instantly, at any scale, at a very low cost, but these advantages are typically only being realised with respect to informal communication among scholars. With the underlying business model of the journal publisher persisting alongside the Web, large swathes of knowledge are being closed off to all except those who can afford the subscription fees, with those who can afford access suffering an inevitable trade-off – a problem afflicting even the most affluent of academic institutions [24]. Not only is the persistence of this form of publishing costly, it perpetuates an outmoded style of communication that is unsuitable for modern research methods, having been devised at a time when science was conducted in a radically different manner which would seem significantly out of date to most modern scholars.

As the disruptive nature of the Web makes its presence felt in other sectors and markets, the anachronous nature of scientific publishing becomes clearer. This in turn has led to some academics taking action – most notably those behind the Open Access (OA) movement and others such as those who have signed the Elsevier boycott [18]. Yet ultimately the overall impact of these attempts may have been a little muted as a result of the entrenched nature of academic publishing in the wider academic community. It is only by understanding publishing as a multi-faceted process that possesses different consequences for different stakeholders that it may be possible to engineer an approach that changes publishing in a way that is compatible with individual stakeholders’ motivations and ambitions.

Having come to a definition as to what publishing actually involves and how these various consequences may be exerting a problematic influence, it is worth considering what publishing might mean to scholars in the future, providing a long term goal that some of the ideas presented in this project aim to be a small step on the path towards. It is clear that everything that publishing encompasses at present is still required by the research community, but the means to which quality control and peer assessment is conducted need not necessarily be an intrinsic part of the publishing process. But perhaps the most pressing issue that needs to be addressed and that the Web has the means to resolve is that of cost. The future of publishing should represent a much more open process that is free of most financial barriers. Where research is funded by public funds for the benefit of society, such research should be open for any who wish to use it or build upon it to further the public good that it enables.

Ideally, with the Web allowing fast, scalable and inexpensive communication, publishing should extend to allow the recognisable dissemination of all research outputs, rather than just that which is publishable by journals. Rather than editors deciding which research is of significant interest to be made more widely available, instead the wider research community should be able to provide an indication as to which research projects are of most relevance or interest. The publication of more research should also allow for greater reproducibility in science and efficient allocation of resources, with fewer research projects being unintentionally duplicated. On a similar note, in an attempt to remove elements of arbitrariness from publishing, peer review also has an opportunity to be parted from the publishing process, by allowing the community to verify scientific soundness rather than reviewers who are expected to assess for both correctness and interest. The diminishing influence of a journal publisher should also extend to allowing a greater range of publication formats. With much research being simply disseminated as articles that have followed the same format for decades, if not longer, publishing has become simply a case of pushing PDFs around the Web, in much the same way paper journals were delivered to their audience in the past. Publishing should exploit the digital medium offered by the Web to allow research outputs of any format to be not just accessible but also recognisable as a potentially useful contribution to the community. Such resources should also be complemented with rich metadata that enhances their value for the user.

Figure 2 - Facets of publishing that can be improved or expanded upon by leveraging the affordances of the Web. Many of these areas are left either under-served or highlight flaws in the current approaches to publishing. By fully taking advantage of the digital medium of the Web, it should be possible to remedy many of these issues to create a publishing environment that better serves researchers and future knowledge generation.

The image of publishing in the future is one of openness and flexibility that removes restrictions and allows researchers and the wider community to use the outputs of research as they see fit. The functions of quality control and ranking of peers are not ones which are built into the publishing system, but instead can sit on top, using publishing as a foundation without exerting an influence on how work is itself disseminated, whilst simultaneously operating in more efficient and desirable ways. Not only should it be possible to implement the traditional, value adding services independently from publishing, but an open and flexible platform, should allow for further value-adding services that make scholarly artefacts more useful and informative, acting to accumulate more information as they pass on from one researcher to another, spinning off in different ways to create new strands of research. However, an important property to consider when imagining the future of academic publishing is the nature of this publishing environment, which comprises a multitude of different disciplines, each of which has its own nuances and approaches to dissemination. As with approaches to publishing that predate the Web, it is likely that one size does not fit all and new systems will need to be wary of the context in which they operate.

## Disintermediation

### Introduction to Disintermediation

In many instances when individuals need to interact with one another, a third party is required who acts as an intermediary, facilitating the interaction between the other two actors. The demand for an intermediary typically arises when one individual lacks the necessary skills, knowledge or resources to successfully take part in the transaction and thus another actor is required to lend their expertise to a certain extent. In doing so, the intermediary allows the transaction to run smoothly and is also presented with an opportunity to add value to whatever good or service is being provided, contributing to the overall system. This relationship typically exists whenever an individual needs someone to act on their behalf; travel agents or insurance brokers making for typical examples.

However the Web has presided over a shakeup of many of these traditional relationships which has the seen the roles of intermediaries redefined or, in some cases, removed completely. This process can be defined as disintermediation: “the displacement or elimination of market intermediaries, enabling direct trade with buyers and consumers without agents” [58]. The travel agents example nicely illustrates how the Web has affected intermediaries, with disintermediation in this instance taking place in two phases, an “initial” phase in which the Web assisted travel agents at their job, and a “terminal” phase in which the need for travel agents has been surpassed [55]. The Web is renowned for its ability to break down barriers to communication and by doing so, lessens the demand for certain intermediaries, resulting in pressure on intermediaries to demonstrate that they can still add value. As a result intermediaries often need to find new roles in the market, falling back on sub-functions they may offer, to prevent themselves from becoming irrelevant, with the only route to a sustainable business model being able to understand the other stakeholders’ perceptions of value-adding services [53].

Disintermediation is not the only Web enabled process that threatens to shake up the traditional roles and interactions that take place in many markets. Apomediation, may result in increased prominence given to existing players, with “influential peers and opinion leaders” becoming the “primary conveyors of trust and credibility” [14]. Apomediaries operate in a similar fashion to intermediaries, but direct potential consumers “through the onslaught of information”, giving “additional credibility cues” and providing “metainformation”, rather than the approach of traditional intermediaries of carefully selecting information to be presented to consumers [14], a role made feasible by the rise of social networks such as Facebook [37]. As discussed by Eysenbach, whilst apomediaries are seen as more “equitable and democratic” approaches to providing value added services, conversely influential “hubs” may emerge develop increasingly powerful positions, and the traditional intermediaries may often find themselves in an optimum position to establish themselves as “credibility hubs” [14].

### Stakeholders of Academic Publishing

Having established the manner in which the Web can affect traditional relationships and interactions among stakeholders it is possible to see these disruptive effects in a number of markets. However it can be argued that these disruptive effects have not yet been actualised in academic publishing, as previously discussed, with traditional business models continuing to dominate the market. Scholarly discourse is a very different environment to more traditional markets, with the relationships between the conventional roles of consumers, producers and intermediaries differing from those found in most markets. To understand the potential of disintermediation in academic publishing, it is first necessary to build a network of the stakeholders and understand the role of each actor. Table 1 attempts to list the major stakeholders, and whilst it may not necessarily be exhaustive, endeavours to highlight all the relevant actors.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Stakeholder** | **Example** | **Goal** | **Provide to the System** | **Expect from the System** |
| Researcher | PhD students, Professors | * Do research * Enhance status in community | * Research findings | * Research findings * Reputation/Status |
| Institution | Universities | * Maximise researcher output * Increase ranking | * Organise research * Provide facilities and services | * Researchers to secure research grants * Funding from government * Students |
| Discipline | Chemistry | * Understanding | * Epistemologies | * Disciplinary corpus * Quality assurance |
| Research Council | HEFCE, EPSRC | * Report academic, societal and economic impacts | * Allocate funds to projects * Report back impact | * Demonstrable academic, economic and/or societal impact |
| Society | Parliament, public | * Achieve academic, societal and economic impacts | * Funding to research councils | * Academic, economic and/or societal impact * Growth |
| Journal Publisher | Nature | * Profit | * Research findings * Organise peer review * Editing, typesetting * Impact factors * Source of recognition | * Subscribers/readers * Profits * Publishable findings |
| Institutional Library | University libraries | * Provide researchers with access to research | * Manage journal subscriptions and access | * Funding from university. |
| Learned Society | Royal Society of Chemistry | * Earn sustainable income to support services for members | * Publishing outlets * Conference organisation * Accreditation * Source of recognition and rewards * Non-profit operations | * Funding via membership. * Funding via subscriptions |
| Third party database | ChemSpider | * Earn sustainable income/profit | * Scholarly resource * Outlet for research data * Value adding services | * Non-open access databases require subscriptions * Free, crowd-sourced databases require user input. |
| Third party metadata | Web of Knowledge, Scopus | * Profit | * Metadata aggregators | * Subscriptions |
| Student | Undergraduates | * Education * Employment | * Source of funding * Future generation of researchers | * Teaching * Research led learning environment |
| Industry R&D | GlaxoSmithKline | * Develop new products and services | * Collaborations with universities * Research findings from R&D projects | * New products and services * Profits |
| Research Consumer | Hospitals, SMEs, etc. | * Improve products and services | * Resources to conduct research | * Access to research to achieve better services/growth |

Table 1 - The actors present in scholarly discourse, illustrating their ultimate goals or motivation, what they provide to the system and what they may expect in return.

As demonstrated in Table 1 the stakeholders in scholarly discourse are a varied collective that represent a market which is different from the traditional consumer, producer and occasionally intermediary set-up that is more common, with many stakeholders not easily placed on any one side of the transaction. Researchers are the agents responsible for both producing and consuming research and thus the main transactions of scholarly discourse occur within in this group. The majority of the other stakeholders (except perhaps for industry) are all present to enable or facilitate the work of the researchers, some of whom are more vital than others. For example, stakeholders that enable researchers to their work include disciplines which are required to provide structure and epistemologies for researchers to follow, researchers need public funds to do their work, and institutions provide a venue for research to take place (although it could be argued this last stakeholder is not vital to the process of research, but simply makes it much more economical and feasible). Some stakeholders however provide useful services, but are ultimately wholly dependent on what other stakeholders provide or expect from the system: research councils allocate the funds from one stakeholder to another, journal publishers and learned societies distribute research among the researchers, and the third party organisations facilitate the work of researchers, but in some instances provide little original value – and it is these stakeholders that represent the intermediaries in this market.

Whilst it is clear that each stakeholder has their own individual role to play in the process of scholarly discourse, the interactions between the stakeholders themselves are not so clear. Figure 3 shows the interactions between stakeholders and what services or products they provide to one another, which in turn reveals two different overall goals that academic publishing ultimately contributes towards.

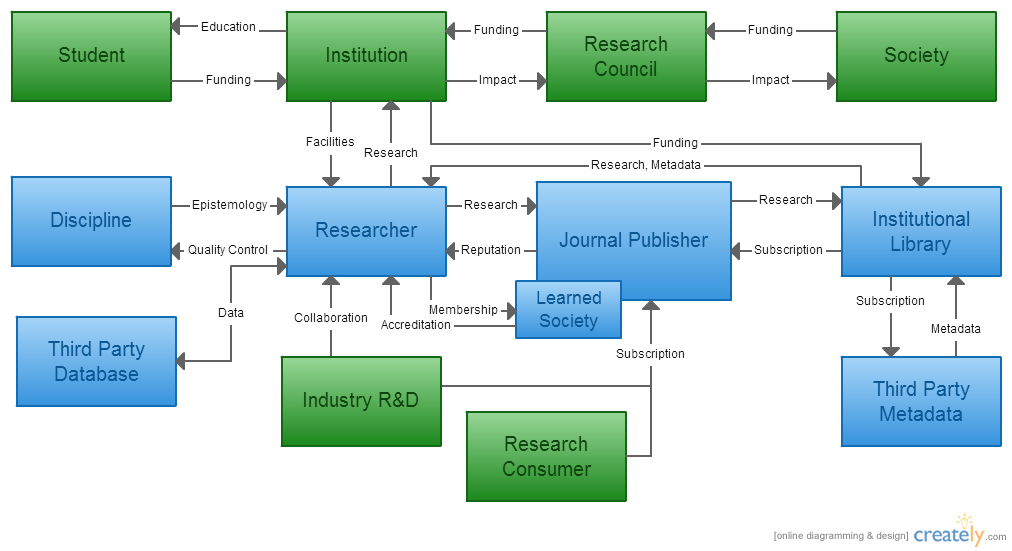


Figure 3 - The interactions between the main stakeholders in academic publishing. The arrows connecting each stakeholder indicate what service or resource one stakeholder provides to the other. The stakeholders depicted in green represent those who take part in the market in which society demands research. The stakeholders depicted in blue represent those who take part in the market in which researchers demand access to research so they can further expand the boundaries of their discipline.

When considering academic publishing from the perspectives of the different stakeholders, two separate markets start to emerge. First there is the “macro” level in which society demands research to be conducted in order to achieve technological and social progress, improving overall quality of life and wellbeing. To supply this demand for research, there is also the “micro” level market in which transactions take place among researchers. In this market, the good that is being traded is research outputs, necessary for the continued production of yet more research, putting researchers in the rather unusual position of both being the suppliers and consumers of the good that they are exchanging (a relationship which is identified by the loop in Figure 3 where research stems from the researcher, through to a journal publisher or learned society, on further towards the institutional library before it is then passed back to the researcher). At first inspection there should be no reason why researchers cannot conduct these transactions amongst themselves, but the sharing of knowledge is a complex issue. Not only do the typical market mechanisms fail when exchanging knowledge, due to information asymmetries [7], but this market also happens to serve other purposes such as quality control and allocation of reputation and status upon individuals, roles which journal publishers have taken it upon themselves to fulfil.

Looking at Figure 3 it becomes apparent that the network of stakeholders and the way in which they interact is quite complicated and so any changes that are made may have widespread and possibly unforeseen consequences. The entangled mesh of transactions may also help serve to explain why it has been difficult to disrupt the traditional processes in academic publishing, with different stakeholders being pulled in various directions, limiting their potential to try new approaches to certain tasks without potentially risking their progress towards their own personal objectives.

Yet in both markets opportunities for disintermediation, and thus efficiencies are possible, with the extent to which these might be pursued depends how sufficiently the affordances of the Web allow other stakeholders to adequately fulfil the roles originally conducted by the intermediaries and how such changes impact on a stakeholder’s relationships elsewhere. Figure 4 shows a new configuration of the relationships between stakeholders once disintermediation and apomediation has taken place, representing the two markets where these intermediary altering processes have had a significant effect (an effect that is quite ambitious and indicates significant changes which would be impossible in the short term scope of this project).

As illustrated by Figure 4 there are numerous opportunities for disintermediation and apomediation in the micro level market, where the researcher now plays a much more central role in academic publishing. The transactions that occur at this level facilitate the exchange of knowledge from one research or group of researchers to another in return for credit or recognition and the important role of the researcher is now clear as they become responsible for providing research to one another, generating their own metadata and allocating reputation amongst themselves (as symbolised by the looping arrow that no longer passes through any of the other stakeholders).

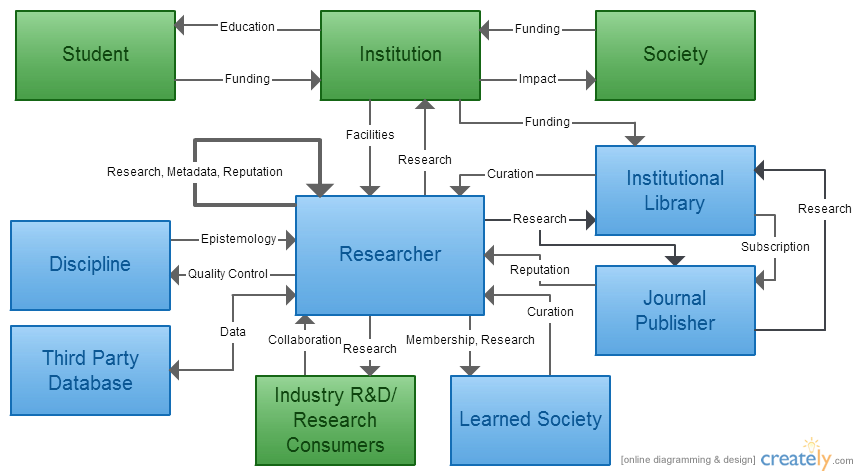


Figure 4 – The interactions between the main stakeholders once disintermediation has taken place. The researcher now plays a much more central role, whereas other stakeholders have been removed. It should be noted that this figure illustrates an exaggerated view of the effects of disintermediation rather than a pragmatic one. The green boxes represent the stakeholders in the macro level market for research, and the blue those who take part in the micro level market.

It is clear that the Web is capable of allowing people to freely exchange information (although its ability to easily exchange knowledge may be a more ambiguous notion) and this has already been observed in the publishing industry as researchers access papers online. What is perhaps less clear however is how efficiently and effectively the Web is able to facilitate the allocation of recognition and the process of establishing a reputation, an unavoidable concern when working in academia (discussed in more detail later when looking at the reward and recognition structure of academia). This function of academic publishing is not so easily disintermediated; it should be safe to assume that a researcher cannot be responsible for determining the level of one’s own abilities and reputation, but the wider community is capable of evaluating the contribution of an individual researcher. At present this process is achieved via combination of the peer review process and the power of journal brands, however these approaches have known flaws [41,42], and the scope for improvement is available with all the information required to evaluate a researcher’s value ultimately being provided by the research community itself.

For example, it would be possible to implement a Web service which allows the research community to organise the peer review process; one could imagine an online platform whereby researchers join, state their area of expertise and affiliations and upload papers and articles that need reviewing. The site could then match papers to potential reviewers and the process of peer review could be conducted in an open, reliable and hopefully efficient manner. Yet such an approach has not emerged – the research community highly values the peer review aspect of publishing [57] and it is perhaps this value that makes it too precious to experiment with thus engendering an element of conservatism in a community whose sole aim is to be innovative.

Should the research community be unable to be self-organising the roles played by publishers and learned societies may ultimately be retained, albeit as apomediaries rather than intermediaries, assisting researchers in making these decisions, whilst hopefully not holding so much influence such that it is not possible to achieve the free and open exchange of knowledge and the further opportunities for disintermediation and value-adding services it may bring. Approaches to help elicit this researcher-centric method of assigning reputation within the community should be the subject of future research, with a disintermediated peer review platform being an intriguing research area.

The potential for disintermediation extends beyond just the distribution of research outcomes and recognition however. At present, institutional libraries pay not just for access to journal publications, but also to metadata provided by third party organisations such as ISI Web of Knowledge or Scopus. If research outputs were made openly available, then the tools that provide this aggregated information on research could also be made openly either by the community or via open source business models. This encourages existing third party metadata services to either be more competitive with their pricing models or add new value to the data they provide, with either route taken advantageous for research institutions and the researchers they cater for.

Not only can existing tertiary services be provided via the use of openly published research and the data that becomes available as a result, but there would also be new opportunities for adding value; a virtue that has been embraced by the Open Data community which sees great value in data being made available wherever possible (data from government departments being made openly available to improve public services being one example) [5]. For example, whilst it is unlikely (and undesirable) for all publishers to be disintermediated, many of the administrative roles that are conducted by intermediaries, such as by administrators at research institutions for example, may be streamlined by tools built on open approaches to research publications. One such use case may be for a researcher to easily pull together their year’s research outputs to submit a report to their institution which wishes to monitor the output of its academics for a performance review. Ultimately, a more open scholarly environment should permit a more agile approach to research, in which researchers can more accurately assess each other’s progress and easily identify opportunities for collaboration. A more rigorous case for openness in academia is made by the *Science as an open enterprise* report commissioned by the Royal Society, which makes a number of recommendations for the future of scientific publishing to “improve the conduct of science, respond to changing public expectations and political culture and enable researchers to maximise the impact of their research” [51]. The report emphasises the important role that should be played by numerous stakeholders, such as universities or learned societies, in supporting researchers being more open with their research outputs; in effect making a case for disintermediation not taken to its full extent where researchers are left to communicate findings by themselves, but to a degree whereby the role of potentially inhibitive stakeholders is lessened and those stakeholders who are present to serve researchers rather than to profit, have their roles expanded upon [51]. An executive summary of many of the ideas discussed here, which aims to present a vision of how an institution may progress to a more agile research environment and the rewards it may bring, can be found in Appendix B.

With disintermediation and apomediation being possible in a number of the transactions that permeate the micro level market in academic publishing, there is then further potential for these changes to initiate a trickle up effect to the macro level with the openness of research making the functions of the intermediaries in this market more easily replicated by other means. For example, in the macro-level market the roles of research councils and funders may be able to undergo disintermediation. They take funds from one stakeholder and allocate it to another. Their role is necessary for two immediately clear reasons: firstly the source of much of the funds, the taxpayer, does not possess the knowledge to know how to most efficiently allocate the funds available (in a market where the market forces approach fails to operate) and it can be assumed that researchers would more often than not desire their project receives funding over another. Thus research councils provide a useful role in overcoming these problems, however the process by which funds are allocated is typically administered by other researchers, which in the UK takes the form of the REF exercise. If a form of peer review which operates via disintermediation could be expanded to work on a national level, it may be possible to reduce the roles of research funders too, eliminating the problems associated with the REF procedure [54]. Avoiding institutional biases or gaming of the system might be difficult however and would require reviewing that was both independent and carefully regulated. However, whilst the vision of a research community which could allocate funds efficiently amongst its own community is an ambitious one, it is nevertheless one which should be considered when thinking about the potential for disintermediation and its ultimate scope. The Web is a powerful tool for communicating with others both at scale and across vast distances; perhaps it would be possible to organise researchers on an international level to judge the efforts of their peers in different countries, thus allowing a fair reviewing system whilst also hoping to remove any preferential treatment from local ties.

Such a solution would certainly not materialise in the short term and, as is the case with many of the ideas suggested here, to adopt these radically different approaches to scholarly communication would require a significant cultural shift from researchers and many of the other stakeholders. It is clear that the possibilities of disintermediation and apomediation pervade both of the two main markets in scholarly discourse, and, as seen in other industries and the new markets that have emerged online, the technology is available to make much of this disintermediation possible and thus introduce efficiencies into the market. Yet in scholarly publishing this desire for change is lacking, a deficiency of progressivism from the community may be a key factor in the lack of disruption in this market that has led to scholarly publishing becoming stuck in the “initial” phase of disintermediation [55]. The Web is asserting its functionality in academic publishing, but only to facilitate many of the existing processes. Despite the possibilities made available by the technology, an impetus to embrace change is missing from some of the major stakeholders, namely the researchers who are largely content with the workings of the current system and the publishers who continue to profit from the present system and have no external stimulus to change their business model. As a result the “termination” phase [55] is never reached and thus many of the problems of scholarly publishing are yet to be “terminated”. Naturally, some stakeholders will not wish the termination phase to begin, with it not being in their best interests, but if the market is to move ahead and take advantage of the Web, it may be an inevitability (however, as discussed previously this does not necessarily mean the end of their participation in the market, just a shifting of roles).

The lack of readiness to change the nature of scholarly communication can be more easily identified when looking at Open Access as a case study as an on-going attempt to revolutionise academic publishing, serving to highlight how despite sufficient technology and infrastructure available to make research more openly available, a lack of both consensus and interest from parts of the academic community can result in a lack of what many would regard as progress.

### Previous Work

The Web’s potential to revolutionise academic publishing by breaking down the barriers that may exist between researchers has not gone unnoticed. Previous work in this area has also looked at the multi-stakeholder situation that occurs in academic publishing, with Line identifying authors, publishers, librarians and consumers as the principle stakeholders that all have dependencies on the publishing medium and highlighting an unfortunate trade off that permeates all work on academic publishing: that “none performs very well for all parties” [31]. Harnad phrases the potential impact new communication technologies may have in a different context, bringing to attention the “continuum” of publishing that highlights that the various stages of research have an opportunity to be revolutionised, placing particular priority on the “prepublication phase” where “most of the cognitive work is done” [22].

Perhaps the most notable proposition of a new view on academic publishing is that described in Smith’s “The Deconstructed Journal”, which arose from a number of familiar observations: that “e-journals attempt to maintain the style and organisation of paper journals”; “the ‘paper’ model constrains and inhibits the inherent flexibility of the network medium” and that the “scientific journal has existed almost unchanged for 200 years” indicating its fundamental importance in scientific discourse and the need to maintain the status quo in some form [49].

The analysis of academic publishing in the proposal of the Deconstructed Journal pursues a similar course to that taken by this project, by analysing the roles that a journal plays; with several functions being identified, some more explicit than others. These numerous functions once again suggest that the role of publishing and the medium which facilitates it extends beyond the mere dissemination of work, drawing attention to other roles publishers perform. Some of these functions have been discussed already, such as the need for quality control (both with respect to “content” and “form”) and “conferring recognition of work done”, but Smith identifies a few more yet to be considered in this project such as an “editorial” and “marketing” role [49]. The classification of some roles as being hidden is also a unique insight and whilst the model of publishing discussed by Smith does not take into consideration the wider network of stakeholders and the different levels at which the publishing market may operate, by highlighting select publisher roles as being hidden, Smith draws attention to the fact that some members of the research community are likely not to possess a thorough understanding of the system they are part of. Such an insight is important to consider when disintermediation requires the same such stakeholders to adapt the nature of the roles as they currently perform them.

Having “deconstructed” traditional journal publications to gain a greater understanding of their place in scholarly discourse, Smith goes on to “reconstruct” a “New Scientific Journal” (NSJ), one which is designed to much more ably exploit the affordances of the Web and bears some likeness to the disintermediation model proposed above. The NSJ proposed would be a web service in which the operator does not own the rights to the items which it points towards; can act as filter between the contents and the user (rather than as a repository of the material); and may or may not be responsible for the quality control stages of publishing [49]. Thus the paper presents a feasible route to a publishing system that is not entirely unlike the description of disintermediation given previously. Yet actual examples of the NSJ are very rare if not non-existent. The paper touches on some of the difficulties of introducing a new approach to publishing, focusing on the need for “acceptance by the user community” and reaches the conclusion that acceptance will come when “professional and funding bodies accept publishing in this model” and their ability to treat it on a par with conventional publishing [49]. Whilst acceptance from the community is a recognised requisite for a new publishing model, the emphasis here is placed on authoritative bodies. However the paper may be underestimating the greater need for researchers to recognise a need for change, for without this the authoritative bodies will have little impetus to act – the absence of any form of NSJ may thus be further indication of researchers’ insufficient desire for change.

The lack of success for the NSJ can provide a number of useful insights in to future attempts of introducing new publishing models. It highlights the difficulties inherent with any paradigm shift, principally the use of certain words and terms that have different meanings in the old context compared to the new [49]. The outcomes of previous attempts in this field do not indicate necessary future failure for similar such projects however. The nature of the Web has changed significantly over the last decade; with most researchers being much more Web savvy individuals who are familiar using the Web for both personal and professional activities. The evolution from Web 1.0, read-only Web to the read-write Web 2.0 has resulted in a change in the manner in which individuals, including researchers, feel able to communicate their ideas with peers and wider audiences [36]. It has previously been assumed that by simply providing a platform on to which researchers can upload their work for all to benefit from, the rewards of opening up scholarly discourse would materialise, but it is now becoming clearer, with the rise of influential social media, that it is a social layer that has been missing from previous attempts to affect change in scholarly discourse. Thus not only is an understanding of the bigger picture of stakeholders required, so too is an understanding of how researchers behave in their environment, how they interact with one another and how they value each other’s contributions, such that a successful social layer can be constructed that will facilitate open discourse.

## Open Access

The Web combined with scholars’ immutable desire to push the boundaries of their fields should allow for the production of an “unprecedented public good” – the intent of the Budapest Open Access Initiative, which envisaged a two-fold approach to opening academic discourse through author self-archiving after conventional publishing or via author-pays OA journals. As the disruptive effects of the Web become increasingly apparent in other markets, and researchers frequently find themselves unable to access the articles they wish to read, the debate around OA has grown, ultimately leading to a number of developments. The UK government has now accepted the recommendations of the Finch report to support ‘Gold’ OA [44], whereby the cost of publication is covered by the author, but the content is free to access – a significant step forward in OA, although it is not without concerns from those who would rather see ‘Green’ OA supported [46] (since the time of writing the Business, Innovation and Skills Committee has released a report summarised as “government mistaken in focusing on Gold as route to full open access”, which will no doubt have significant implications on the future of OA in the UK [12]).

The outcome of the Finch report raises some concerns about OA and reveals perhaps the true nature of some of the stakeholders within academic publishing. If OA makes research accessible to all and researchers are driven by their “willingness […] to publish the fruits of their research […] without payment, for the sake of inquiry and knowledge” [11], then why is OA not more prolific? Similarly, why is the Finch report required 11 years after the Budapest Open Access Initiative? The answer to these questions likely lies in the mutually dependent relationship that exists between researchers and journal publishers. Assuming researchers are self-interested economic agents, their actions will reflect what is best for them; in the instance of the publishing researcher this would be securing a publication in a high impact journal, enhancing their reputation and status in the community, which in turn may lead to increased pay, but comes at the cost of potentially keeping knowledge to one’s self until the moment when publishing the research can extract the greatest benefit.

A counter to this argument would that be an OA article may receive more citations than a non-OA article [52], thus broadening and enhancing recognition, but ultimately it is the power of a journal’s brand and its history that influence publishing academics. Gold OA journals are often expensive to publish in and being relatively recent additions to the publishing ecosystem, are often yet to attain high impact factors and thus are not pursued by researchers.

The alternative is green OA in which a journal permits the author to self-archive their paper in an OA repository. Green OA serves to illustrate two pertinent issues. Firstly, green OA is often viewed as an unsustainable and illogical approach to academic publishing – with researchers making their own deposits of their papers after journal publication, publishers are concerned about falling subscriptions and a flawed business model. Studies have indicated that this may not be a valid concern, especially if the publishers continue to add some value that is not present in the researcher’s pre- or post-print deposit [4], for example by facilitating the peer review process. Whilst this trend has not emerged in disciplines which have achieved full or near 100% green OA, it is nevertheless an understandable (if perhaps uninformed) concern from publishers who may envisage a less stable and more open to competition environment in which to operate as the number of niches to fall back on diminishes.

This perceived flaw of green OA along with the problem of embargo periods delaying access to research has resulted in many in the academic research community identifying green OA as a potential hindrance rather than a boon to academic publishing, with researchers uneasy at the prospect of the existing publishing system being tampered with. Thus green OA also indicates the extent to which OA has mischaracterised the problem that afflicts scientific publishing: it is heavily dependent on journal publishers to carry out the tasks they conduct, cementing their traditional role in scientific publishing, and thus simultaneously dampening the disruptive effects of the Web. Journal publishers, in protecting their own business models whilst serving researchers to the best of their ability before the Web, have created a stable and reliable approach to scholarly discourse; which in turn makes it difficult to alter and improve upon.

OA has done much to improve academic publishing, but its ultimate goal is to seemingly take the power and disruptive force that is the Web and apply it to improving the accessibility of research whilst at the same time endeavouring to maintain the status quo. Making research accessibility more open is a worthwhile pursuit, yet fails to grasp the full potential the Web has to offer. Recent developments have seen the emergence of a third type of OA, typically called “Diamond” OA [19], which starts to focus on more than just issues of accessibility. Diamond OA professes to offer a free to read and free to publish approach, via the form of an “overlay journal”, as proposed by the Episciences project [10]. An overlay journal is derived from a list of links to preprints on a repository such as arXiv, a service which Episciences intends to provide along with a technical platform for organising the peer review process [10]. It is also hoped that each article would be accompanied by a page for comments and reviews of the article, encouraging further scholarly activity. Therefore this new approach to OA goes beyond dealing with issues of accessibility, as noted by Stevan Harnad, who disagrees with prepending new “colours or precious metals” to OA, stating that OA is “not about cost-recovery models (nor about peer-review models); it is about research access” [23].

Ultimately Harnad comes to the conclusion that overlay journals are simply gold OA journals, funded by either subsidies or eventually author fees [23], yet his comments on some of the wider aspects of the publishing process and how they do not pertain to OA, reveal why OA alone may not be enough when considering the potential of the Web to disrupt the system in place. Diamond OA begins to take advantage of some of the Web’s properties with its approach to organising peer review and allowing readers to comment and review published articles, facilities made possible by beginning to disintermediate the journal publisher. Nevertheless the Web could be used to shake up the academic publishing system even further to create a more agile and open research landscape and combat other problems that afflict scholarly discourse.

It should also be noted, that whilst OA does primarily concern “research access” rather than “cost-recovery models” [23], the impact of widespread adoption would have repercussions for other stakeholders, primarily the learned societies that accompany many disciplines and use publishing to fund their other activities, often identified as public goods and recognised as contributing towards the “international standing of UK research” [17]. The role of learned societies is not one academic communities would wish to see diminished, yet shifts towards OA threaten their business models. However this does not mean conversely that a push towards open methods of dissemination should not be pursued; instead new approaches to discourse present new opportunities for learned societies and the chance to adapt their business models, in what is an inevitably “uncertain publication environment” as the effects of the Web become ever more evident [17].

## Reward and Recognition

An analysis of methods which have been used in other markets to spark disruptions and a look at a significant and somewhat successful approach to opening up scholarly discourse that is on-going and making continuous developments; has led to a number of conclusions. It is clear that many of the main stakeholders in the scholarly discourse market have little impetus to change and in some instances may wish to avoid changes, with present business models serving sufficiently well for researchers and publishers alike. However the model of discourse that has developed over the years prior to the Web is outmoded and costly from a researcher’s perspective, yet despite these problems it has remained remarkably stable to outside disruptive influences. This stability may partly be attributed to the fact that whilst it is imperfect it is ultimately functional, allowing researchers to communicate their ideas with one another and in its provision of other useful services such as providing mechanisms for ensuring published work remains of a suitable standard, and for ranking both fellow researchers and their institutions. Another possible contribution to the robustness of the system may stem from the network of stakeholders, in which arguably the most important stakeholders, from whom the entire system hangs upon, the researchers themselves, find many of the problems with the system obfuscated by other actors, such as institutional libraries for example (who have to deal with the more explicit financial costs of dealing with academic publishers).

Thus if the concept of disintermediation is to be not just applied to scholarly publishing as an idea, but also implemented to instigate a change in the approaches taken to academic discourse, it is necessary to understand the researchers, their complex roles as both consumers and producers of scholarly discourse and their motivations and concerns that affect their behaviour. Disintermediation requires researchers to adapt their approaches to scholarly discourse as they become responsible for making more of their outcomes available to the wider research community, yet it is understandable that should they feel any changes may impact upon their success in the field, they may be resistant to disruptions. Therefore it is vital to frame suggested changes in such a way so as to fit with researchers’ current perceptions, whilst nevertheless attempting to subtly shift the culture that has developed around scholarly discourse. With current perceptions of reward and achievement in scholarly fields being heavily weighted towards the prestige associated with publishing in a journal article, disintermediation will require alterations to the methods by which success is measured in academia; a point raised by the *Science as an open enterprise* report which states, “major barriers it widespread adoption of the principles of open data lie in the systems of reward, esteem and promotion in universities and institutes” [51].

It is often assumed that researchers are motivated solely by their curiosity and desire to push the boundaries of knowledge, and the literature concerning the sociology of science does much to back-up this assertion. Most scientists are sufficiently driven by an interest in the subject they study, yet as with any institutionalised system, there sits a reward structure, which ultimately impacts the behaviours of researchers [3]. With the aim of science being to create new knowledge, “originality is at a premium” [33] and thus it is for new discoveries that scientists are rewarded, which in turn leads to an emphasis being placed on the priority of discovery and in turn the publishing process as a way of laying claim to one’s discovery. As Merton observed, it is with priority and the divisive arguments that can arise when there are conflicting claims over who takes priority of a discovery, that the institutional norms of science can be identified, principally the need to be recognised by one’s peers; the only “property right” a scientist may reserve, having given up the fruits of their labours to the rest of the community [33].

With scientists striving to contribute original knowledge to their corpus, it is only natural that recognition from peers is the currency of reward, with the highest honour being eponymy; yet eponymy by its very nature can ultimately only be applied to a very small minority of the total community [33]. This scarcity of recognition extends further still with many researchers unlikely to ever be acknowledged for their contributions in any significant manner [21] and yet still many researchers are driven to participate and contribute their findings. The other popular incentivising reason for conducting research is that of personal enjoyment, which Gustin develops into the notion of the allure of scientific charisma, “because through it, a man seems to or is thought to come into contact with what is essential in the universe” [21]. Fundamental to both the concept of recognition and charisma is the act of publication: for many researchers, “getting things into print becomes a symbolic equivalent to making a significant discovery” [33] and without a publication, “a new discovery is incomplete – it is not a work” [21]. The process of peer review lends further significance to publications, legitimising a researcher’s standing in the community and elevating certain members of the community to gate-keeper status, in turn a form of reward [3].

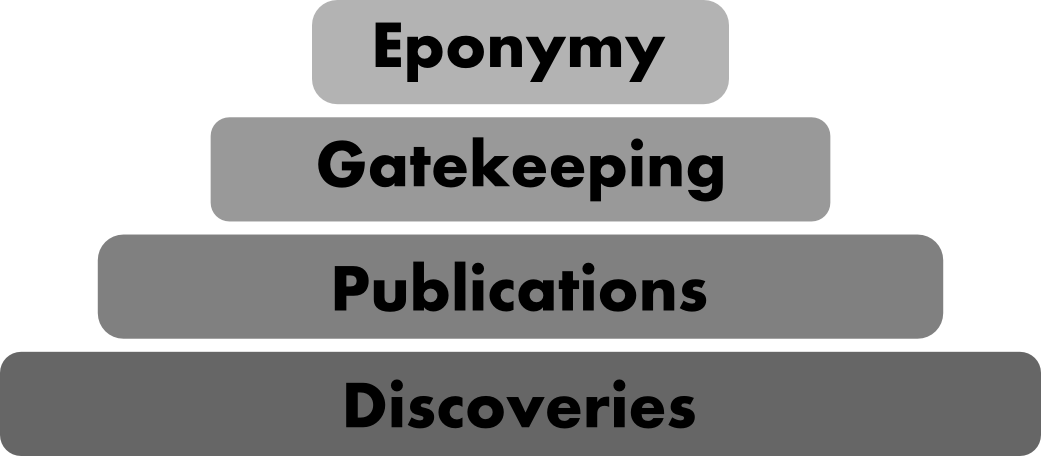


Figure 5 – The levels of recognition that a researcher may work for. Recognition is the only premium in academia by which researchers can measure their success. Many will make discoveries and secure publications, but fewer will go on to become gatekeepers of knowledge and fewer still will achieve fame and eponymy in their field.

Whilst it should be noted that many of the ideas presented above predate the Web and scientific practice has moved on in recent years with the introduction of new methods and technology, the institutional norms of science can still be observed today, with researchers still concerned over issues of priority and maximising the value of their research findings. Arguably the pressure to demonstrate novel and valuable results is greater than ever before with significant pressure from funding agencies and institutions to demonstrate the impact of research. The “publish or perish” atmosphere that has been cultivated in academia has resulted in academics publishing to not merely become part of a wider community and be associated with the charisma attached, but to “gain grants, promotions and tenured positions” [1]. This has seemingly had an overall effect of muddying the waters of reward and recognition mechanisms in research, with concerns such as an author’s h-index influencing a researcher’s publishing (and citing) patterns [27]. The nature of modern research problems has also resulted in a changing “concept of authorship” [43] as a result of large collaborative projects and the introduction of citizen science. This in turn can reduce the “visibility of individual role performance” [34] and thus likely lessen the feeling of reward achieved by securing a publication. Developments in the way in which papers are authored may represent the start of a change in the way in which individual researchers perceive their position in their community and how they value what they do; presenting an opportunity for modern and open approaches to scholarly discourse, reward and recognition.

Thus with the institutional norms that place great significance on the need to publish, alongside the manner in which recognition acts as reward, there is little incentive for researchers to experiment with the way in which they publish, for fear of jeopardising their potential future careers. It is this aspect of scholarly publishing that the OA movement failed to take into account: “as long as scientists rely on publishing in high-impact journals to secure funding, there is no incentive to switch to new, open-access models” [39]. This statement perhaps does not ring true as much as it once did when it was first made in 2003 with some OA journals becoming increasingly prestigious, but it nevertheless indicates that the reward and recognition structure in place in scholarly publishing hugely influences the flexibility of the system and the scope with which it can be experimented with; the same paper raises the issue of early career researchers: “PhD students have their career to make and should not be sacrificed on some altar of idealism” [39]. The strong correlation between the act of publishing and career progression helps to explain that whilst scholarly dissemination may not be considered to be very efficient and can at times hinder every day science, it is not perceived as a significant problem by a majority of the scientific community who depend upon it to advance their careers.

# Methodology and Hypothesis

From analysing the literature and observing the open knowledge and open science communities, it is clear that a malaise is impacting upon scholarly publishing. Some problems are well documented and defined, such as the serials crisis [8] and thus lots research has been undertaken to investigate solutions to such problems. Other concerns, such as the effectiveness with which researchers can present their achievements in a field, or the extent to which data and knowledge goes unpublished because it does not suit the style desired of conventional journal articles, are ultimately much trickier to address.

The approach taken thus far has been one of considering the various aspects of academic publishing and manner in which they are related to one another and interact with each other to form a larger system. An examination of academic publishing in this way reflects structural functionalism and has allowed the idea of disintermediation to arise through an understanding of the society that academic publishing has formed; with the project thus far being particularly influenced by the norms and institutions that have arisen in this sector. However the structural functionalism perspective does not allow for a finer analysis of the problem and potentially restricts future investigations. As a result different epistemologies and methodologies will be employed in future.

Having arrived at the idea of disintermediation, it is believed that to solve many of the problems that afflict academic publishing, a shift in the roles of publishing is required that places more control and responsibility in the hands of the researcher, whilst simultaneously making these new demands work for the researcher such that they make it easier for a researcher to achieve their personal goals rather than adding an extra burden. Thus a working hypothesis for this project may be that: “Disintermediation allows for an open and effective approach to disseminating scholarly resources and accomplishing the roles that this process encompasses”. Such a hypothesis states that by allowing researchers to disseminate their work amongst themselves, without the intermediary of the publisher, should lead to not only more open access to science, but also make the subsidiary roles and side effects of publishing, namely the reviewing of work, demonstrating one’s abilities and assessing impact more easy and effective to carry out.

However, a hypothesis that looks at disintermediation from such a broad viewpoint is not a pragmatic one and would be very difficult to test in a sufficiently rigorous fashion, requiring collaboration from a significant number of researchers and a shift in the way these researchers go about publishing their work. As already discussed, with the phenomena under investigation being responsible for assessing a researcher’s career, it would be immoral to experiment with them and jeopardise a researcher’s future in their field [39]. Similarly, whilst the most effective way to investigate new approaches to publishing would be to mandate new approaches from the more authoritative perspective of the funding councils, such an approach is neither pragmatic nor is it desirable when one of the primary aims is to establish how researchers act towards open approaches to publishing whilst minimising external pressures.

Thus future research will adopt two strands that will be used to investigate disintermediation and its application to academic publishing at different levels. The first strand proposes a model for how disintermediation might operate in reality, looking at the new roles of different stakeholders and the technologies that may be required to implement such a scheme. This strand of the research will be validated by peer review from experts within the community to assess whether the models proposed are a viable alternative to publishing practice, even if difficult to implement or realise in the near future. It will also investigate what services are already becoming available that could be used to facilitate such a publishing model and analyse the extent to which the academic community is embracing these services and whether this is impacting upon the way researchers think about their approaches to academic publishing. Drawing parallels between the disintermediation approach to publishing and existing services should not only influence the direction of disintermediation in scholarly discourse, but also help validate this new approach from the perspective of the research community.

The second strand of research aims to investigate if some of the requirements of disintermediation can be satisfied in reality, via the examination of some disintermediation issues on a small scale that would be interoperable with current research and dissemination practices, adopting an approach that examines how it might be possible to encourage the end goal of disintermediation in smaller, iterative steps. Disintermediation has two main requirements: the need for researchers to disseminate the work they do amongst their peers and beyond; and the fulfilment of the various roles of publishing. As such, an investigation into researchers’ motivations and what incentivises them is required, as this will both help to establish how researchers may be encouraged to contribute and provide new mechanisms for rewarding researchers and assessing their output. To investigate these areas of academic publishing a number of tools and “widgets” will be developed that aim to understand how researchers may partake in the publishing process and what facilities will be required for them to operate in the publishing environment, devoid of the influence of the traditional publisher. Prototypes of the tools or widgets will be built and released to undergo an evaluation by the research community or experts in the field, with feedback being considered both with respect to the environment in which they have been deployed and their potential as part of disintermediation.

## Disintermediation in Practice

An examination of the stakeholders and their interactions has led to two main conclusions that will inform much of the future work of this project: first that researchers are the major stakeholder and a large majority of the interactions that occur between any two stakeholders largely depends on their work at some point; and secondly that researchers act as both producers and consumers in the market for scholarly discourse. This leads on to the notion of disintermediation, the process of reducing the role of one stakeholder who merely facilitates a process between two others; with the most significant efficiencies coming from the removal of the journal publisher. However the extent to which disintermediation can take place is reliant on how effectively new technologies, in this instance the Web, can facilitate the roles of old intermediaries.

In the combined context of scholarly publishing and of this project, the success of disintermediation lies with the Web’s ability to allow researchers to perform the roles traditionally associated with the journal publisher, namely: the distribution of research; the provision of some sort of editorial function (both with respect to content and formatting) and the administration of a reviewing system for quality control. Not only do these functions need to continue in some form, but the needs and wants of the researchers themselves need to be adequately fulfilled for new systems to develop and be sustainable alongside the well-established conventional methods.

1. **Reward and Recognition Page**

To substitute the first function, the Web can allow researchers to host their work on a server from which it is then made available to a much wider audience; a page which demonstrates a researcher’s contribution to their field and is able of conveying their achievements and capabilities to whomever may be interested. Such a page should essentially work for the researcher, helping them to progress with their career, be it furthering their status in academia or applying for jobs in industry, ideally presenting the researcher’s work in the context that best suits the current viewing audience.

However, such a solution is far from being simple: the work needs to be “distributed” in some way such that others are made aware of its presence; it would be helpful if there was an indication as to the level of quality of the work on offer and the significance of its contribution to the field; and infrastructure and support is required to make such a service operational (including support for both technical requirements and researcher digital literacy). Not only does such a service need to fulfil these criteria, but would also need to offer a convenient experience to users, with any attempt to collate a researcher’s work being a streamlined solution that integrates with pre-existing tasks [9], preferably harvesting content as it is generated (via integration with tools such as electronic lab notebooks) or as it is being formulated into more traditional forms of discourse. It is fair to assume that researchers do not want to be tasked with more administrative roles and would instead rather be able to spend a larger proportion of their time conducting research; a system which respects this preference will more likely be favoured by the academic community. Ultimately it is hoped that such a page would act as the authoritative representation of the researcher on the Web, in a sense that might mimic proposals suggested by the Semantic Web community when real world resources cannot be adequately dereferenced from their URI [47] and users are instead redirected to a page about the resource they are looking for.

However, whilst such a service demonstrates how disintermediation may operate to serve certain stakeholders in academic publishing, it is unclear how it would be implemented. Possible routes of service provision are numerous, yet require new burdens to be placed on existing stakeholders or involve services which are currently available but are often perceived to be ineffective compared to preferred, traditional processes (for example the use of a research-centric social networking site). A service would need to fulfil the dual requirements of being both a convenient site to use which does not impose further administrative duties on researchers, and also carry an air of authority and credibility. Numerous sites already exist that allow a researcher to develop an online profile, but these can be time consuming to cultivate and curate, and it could be argued none have yet reached a critical mass of users that lends them the legitimacy required to be properly acknowledged by academic communities. Thus two potential stakeholders would likely be appropriate facilitators of this role: academic institutions such as universities or learned societies.

In the case of the former, any researcher is almost certainly already going to be affiliated with an institution and as a result likely possess some sort of profile on the institutions website, which itself will likely run a repository to store its research outputs. However at present a researcher’s profile is likely to be no more than a brief biography and a collection of links to further work, rather than a complete collection of their work to date and a detailed indication as to their influence and significance to their field. Whilst it may attempt to act as a homepage for a researcher, it will likely not be the sole resource a researcher would direct those who are interested in their work to. Institutional repositories are also often unwieldy to use [28], a finding that originates from earlier in the development life of many institutional repositories, but may nevertheless still hold true for many researchers today.

Nonetheless, the potential for a much more flexible and informative page describing a researcher and their work is present and coupled with the institutional environment has an opportunity to provide a number of benefits for the individual and the institution. Not only can an individual’s research activity be collated, but so too can information pertaining to projects that are run within the institution that the researcher is involved with. The process of harvesting research activities could also be tied in to administrative procedures required by the institution saving researchers’ time and introducing administration efficiencies for the institution. In the short term however this may impose significant costs on the institution that would need to develop it, and in the long term maintain it. Other problems arise when considering that an institution may not wish to be responsible for hosting a page that presents a researcher’s achievements accomplished outside of their position at the institution; and the complexities and inefficiencies that may arise from different institutions implementing different solutions (although the competition aspect this may introduce may ultimately result in the provision of better services).

The other stakeholder in a position to implement such a service would be the learned societies that are typically associated to a particular discipline. By engaging with this new form of scholarly discourse, the learned societies are provided with an opportunity to build upon their status within the community and develop a new business model for themselves in a market which is currently under threat from the Web and the proliferation of Open Access (an even greater concern in the context of the findings of the Finch report [26]). They may also be able to use their disciplinary expertise to assist with the development of platforms for communication, leveraging their previously held positions in the community to add value in a new market [14]. However, problems are once again present with this approach also, with the transfer to a disintermediated environment potentially being very costly for learned societies during the development phase, and with the paradigm shift in publishing possibly discouraging those researchers who prefer traditional approaches to publishing, there remains a possibility that researchers may not continue to renew their membership to support such an enterprise (and those who are not already members may also suffer when it comes to dissemination). An argument can also be made that this fails to properly represent disintermediation, with a single stakeholder in each academic community being responsible for dissemination of scholarly articles, although the concern is perhaps lessened by the non-profit nature of these stakeholders.

Whilst a reward and recognition page that represents a researcher’s findings would be a useful device in achieving disintermediation, it is unclear how such a service would ultimately be provided. However the philosophy of openness that accompanies the ideas behind disintermediation, mean that in practice, the implementation of these pages should not matter. By publishing researchers’ findings as open data, the specifics of the infrastructure that make such a page possible are no longer important and instead services could be built on top of the data available, helping the researcher make best use of their data. For example an institution may wish to provide a service which can take a researcher’s data and help the researcher to automatically generate a professional development report in a quick and streamlined manner. Alternatively, a researcher could deploy a feed that keeps tracks of developments made by certain researchers or in certain topics to keep abreast of the latest research in a field. The benefits that are made available by adopting this approach are numerous and should ultimately help researchers by allowing their findings to be demonstrated in the most effective way possible; help institutions by being able to more easily demonstrate the overall output of its research groups; and help learned societies by creating a new business model.

1. **Crowd Sourced Editorial**

The editorial role carried out by a journal publication is two-fold. First journal editors are responsible for choosing which articles will make for appropriate contributions to a specific issue of the journal and allocate reviewers. Secondly copy editors are used to assist with the presentation of an article to ensure that it is fit for publication with respect to its layout, style and formatting. Both are important roles in the production of a journal valued by the community it aims to serve, yet arguably the value added here does not equate to the financial costs imposed upon institutional libraries.

Thus for scholarly dissemination to follow a broadly similar approach to conventional publishing once disintermediation has taken place both of these roles will be substituted in some way. However, the disintermediation approach proposes that researchers have individual pages that represent their achievements, rather than these works being spread across a number of journals or conference proceedings. Therefore the role of a journal editor is no longer required, but this presents the problem as to how interesting pieces of research are brought to the attention of the wider community. One route to identifying interesting pieces of research is to use the many eyes of the Web to highlight items of interest, by allowing viewers to simply rate an article as interesting or not, with those that garner a number of votes slowly rising to the top of their field and being seen by more viewers (in a system not dissimilar to the popular social news site Reddit (http://www.reddit.com/)). A potential disadvantage of this approach would be the requirement of some central site which would be responsible for allowing this service to take place, but this does not need to threaten the decentralised nature of disintermediation; it can merely be a service built on top of the data that has been made available, acting as an “overlay” journal [10] or implemented by existing publishers in a value adding capacity whereby the value added exceeds the cost from the perspective of the consuming researcher.

The second role of ensuring appropriate formatting and style is applied to research outcomes in order to make it easier for others to make use of them may be a trickier problem to resolve without imposing new costs upon researchers. Whilst strict formatting guidelines may no longer be necessary if a researcher is making their work available via their own personal means rather than through a published journal, it is likely that for work to continue to be taken seriously by the wider community it needs to maintain a sufficient standard. This may provide another opportunity for learned societies to assert their relevance in the modern research environment, by suggesting standards and styles that researchers should follow for a myriad of research outputs relevant to their discipline; or publishers could go further still by continuing to offer a copy-editing service for a fee, allowing them to continue to profit from those areas in which they do add value to research.

The role of editing, in whichever guise it takes does however have the potential to be outsourced to the “crowds” of the Web and the wider research community. It is possible to assume that the need to compile collections of research outputs on a particular theme, as a journal does, is no longer necessary in a world where researchers can (and increasingly do) search for articles using a search engine rather than perusing journals. The need for an editor is increasingly unnecessary and perhaps instead is being replaced by a demand for search and efficient and effective methods for tagging research such that it becomes much easier to find and categorise; highlighting once again how the conventional approaches to dissemination are often outmoded.

1. **Publish then Filter Peer Review**

The final role that would need to be fulfilled is that of ensuring published work is of a sufficient standard so as to be beneficial to the academic community and has been conducted in a rigorous, sound and correct manner. At present this is achieved via peer review, a process organised by journal publishers but carried out by researchers. Thus with disintermediation it is merely the process of organising peer review that a substitute needs to be found for, although should other forms of quality control emerge as a result of disintermediation these would not be unwelcome.

The very nature of disintermediation and the philosophy that underpins it suggests that a “publish then filter” approach to peer review is desirable. Disintermediation operates under the assumption that all research outputs should be made more widely available as there is a chance they will be useful to someone in the academic community, even if that audience is very small, or by having the outputs available it will help with the interpretation of some other, higher form of the research that has been published. Thus only a system which reviews work after it has been published in some manner would be appropriate, for which there are two possibly routes for an implementation.

The first option would be to once again employ the many eyes of the Web and simply use social media sites to act as a channel of debate [32]; convenient locations as many scholars are already using these sites to discuss their work. However this raises concerns that such a scheme would result in a deluge of publications from which it would be difficult to identify the high quality pieces of work [50]. There would also remain the possibility that many in the research community would not give discussions held on social media sites the same level of significance as to the traditional peer review process, with social media’s potential lack of structure or rigour (not to mention that this process would exclude those voices who are not present on social media sites).

The alternative option has already been briefly touched upon and would require the development of an online platform that would help researchers organise the peer review process. Once again, whilst this strays away from the disintermediation idea by resorting to an intermediary to manage a certain process, it nevertheless hones in on an area where previously the roles performed by intermediary stakeholders do add value, allowing these value-adding services to be retained but (hopefully) at a more reasonable price to the research community. Few implementations of journal-free peer review platforms exist at present, but interest in them is growing [40], with one of the first platforms of this kind, LIBRE [38], “a free, multi-disciplinary platform that allows academic authors to invite expert peers to formally review their work” [40]. The identities of authors and reviewers are open for all in the community to see, to ensure transparency and to assist collaboration between the author and the reviewer, with the intent of making the review process much more effective at improving submitted work. However at present it is unknown how the research community will embrace a platform like LIBRE and there will be little obligation or incentive to contribute a review (especially if a researcher is also compelled to do reviews for a well-known journal or renowned conference). It should also be noted that LIBRE is just one, well developed platform for organising peer review, and numerous other platforms may begin to emerge that operate using different models.

### A Model for Disintermediation

Having identified the major roles which would need to be substituted in the absence of a journal publisher, it is clear that a number of alternative solutions exist, each of which places new pressures on certain stakeholders. Perhaps the most pragmatic approach at implementing disintermediation therefore would arise by decentralising many of the major roles, but ensuring that the processes of publishing occur in a well-defined manner such that they are interoperable and compatible with those performed by others. To achieve such a goal would require collaboration from a number of different stakeholders and the use of a wide range of technologies as illustrated in Figure 6. The model depicted in Figure 6 is just one interpretation of how the services provided by the stakeholders of academic publishing could be reconfigured to offer a more open approach to discourse and involves learned societies, who are trusted to act as benevolent representatives for their discipline, taking on a larger role by providing services which would otherwise be left wanting in the absence of traditional journal publishers. These roles include both the provision of copy editing that would assist researchers in producing accessible research outputs and infrastructure that would help standardise research outputs such that new services can be built more efficiently on top of them.

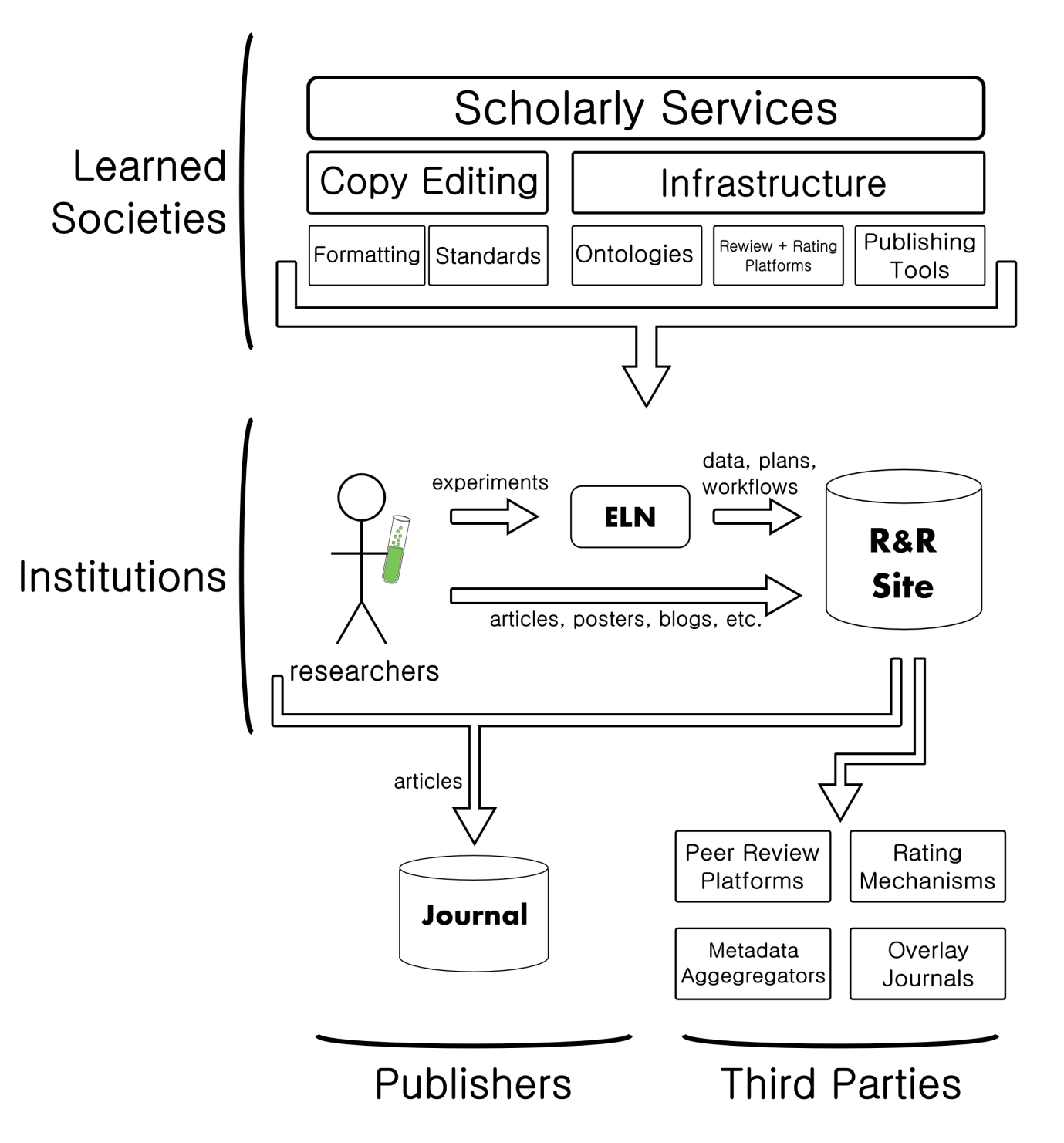


Figure 6 - The roles and services provided by stakeholders in the new scholarly environment. The role of the trusted, not-for-profit, learned society has been expanded upon to provide a number of services to the research community that fall under the dual services of copy editing and infrastructure. Institutions are responsible for providing researchers with the resources they need to continue their work as they presently do, but also for hosting the outputs of their research on a ‘Reward and Recognition’ site. The lower left portion of the diagram represents journals and researchers’ continuing practice of submitting articles, and the lower right portion represents new services that can be built upon the open infrastructure that is the collection of researcher ‘Reward and Recognition’ sites.

Having established the role of learned societies, the demands placed upon research institutions becomes clearer. Institutions are expected to employ researchers and provide them with the facilities they required to do their work, much as they do in the present approach to scholarly publishing. However in this new model, the institution also provides each researcher with a ‘Reward and Recognition’ site which their ELN (assuming a chemistry context) can push content to. Naturally this would require more infrastructure required at the institutional level, although with the widespread use of repositories, much of this is already in place and this approach has the advantage of spreading the burden of hosting research outputs across a distributed network of such sites rather than in one central location.

Finally it is worth noting that journal publishers continue to operate and draw the attention of researchers in the new model. As illustrated in Figure 6, in the new model for publishing, researchers are able to pull together research outputs from both their ELN and their ‘Reward and Recognition’ site during the composition of journal article submissions. Complementary to this process is the provision of a range of services by any “third party” stakeholders who wish to build tools and services that operate on the open infrastructure that has arisen as a result of the ‘Reward and Recognition’ sites as provided by institutions, creating a whole new sector which is able to add value in the scholarly publishing market.

For the model of publishing described above to manifest itself in reality would be require the adoption of open data standards and the application of semantic web technologies in such a manner that ensures the research outputs made available can be used effectively [20]. Many of the typical concerns that are raised about effective data use are not a problem in this instance: the provision of Internet access and services for accessing data are not a problem for researchers who have institutions to manage these concerns for them. If researchers are expected to develop their own “reward and recognition” pages which represent their contributions then a degree of digital literacy would no doubt be required, especially when considering the various requirements that academic publishing demands: “provenance, quality, credit, attributions and methods to provide the reproducibility that enables the validation of results” [2]. In most instances it is hoped that these problems could be offset by making many processes automatic via sophisticated uses of tools researchers are already familiar with, such as ELNs (with the added benefit of increasing uptake from researchers via the employment of technology that streamlines their day-to-day tasks).

The other requirement to make the model above work in an effective manner is the need for a consistent and powerful way in which to describe the outputs that are being produced and published. Underpinning the published works using the semantic web allows for the research outputs to be described in such a manner, with the CiTO “ the Citation Typing Ontology” [48] looking like a viable candidate for describing many of the interactions between resources which would be present in the publishing model described. The CiTO ontology contains a wide range of properties for describing how different scholarly resources may be connected to one another, but is ultimately quite generic (although this trait does also make it quite flexible and powerful) and so in many cases it is likely a subject specific discipline would also be required to better describe the picture which emerges as researchers upload their work. These ontologies may be generated by the community as it begins to build a collection of readily available research outputs or may be built on existing ontologies which have already been devised, such as that described by oreChem [30]. It is likely however that the Semantic Web technologies that underpin this publishing model would need to remain hidden from many users who may be unfamiliar with the technology and ultimately dissuaded from engaging with it if it is so required, further compounding the problem and subsequent need for support that institutions and learned societies will need to overcome.

### Online Academic Publishing Framework

This area of the research may also be further supplemented by conducting an analysis of existing tools and websites that have been introduced to the academic publishing landscape in recent years. These sites and services, such as reference management package Mendely, or research-centric social networking site ResearchGate, to name just two examples, have enjoyed various degrees of success and an investigation into the nature of such tools and their comparative success may reveal insights into how researchers are adapting to the Web and its increasing impact upon academic publishing. This will require both the development of a classification schema that can define what each service has to offer to its users, and metrics for measuring their success in disrupting academic publishing.

### Conclusion

It is clear that the technology is in place for disintermediation to work in practice and if implemented could result in a number of benefits. These benefits fall in to one of three themes that emerge from disintermediation: the facilitation of research; the creation of new opportunities; and the introduction of efficiencies. If disintermediation were to be implemented using some of the ideas proposed then over time one would observe a radical shift to the research environment with free and open dissemination of a wide range of research activities, which should allow researchers to not only lay claim to their work, but also to demonstrate their skills and achievements in a more effective manner that helps to make clear their overall impact on a range of areas, rather than just those impacts that are typically measured by institutions and academia. Not only do existing research practices benefit, but the open discourse proposed by disintermediation should make it much easier to identify new research areas and possibilities for collaboration, allowing for flexible and agile methods to be applied to research, and for institutions to more easily engage with a broader audience. Finally, despite some new costs in the short term, over time disintermediation should reduce costs imposed upon researchers and research institutions and provide other stakeholders with an opportunity to focus on those value-adding areas that are most appreciated by the wider community. Institutional libraries should be able to save on subscription costs to journals and ultimately cancel their subscriptions to third part metadata aggregators, whose role can now be accomplished by open alternatives. Researchers should also be able to easily report their progress back to their funders and auditing tasks could be facilitated by tools that take advantage of research outputs being openly published, allowing researchers to spend more time with their research rather than administrative duties.

However the ideas outlined here, whilst illustrating a number of benefits, nevertheless represent what would be a huge paradigm shift for the academic community, not only in the degree to which many processes are changed, but also in the number of stakeholders such changes would have an impact upon. As a result this will make implementing the changes very difficult in practice and it is likely that many of the proposals outlined here would not materialise for many years. Developments would likely take place iteratively and with the emergence of several third party sites that purport to offer new and innovative services within academic publishing, the beginnings of such changes may be starting to take place. How these will evolve over time however is currently unclear.

## Widgets

As has already been discussed, researchers are not driven by many of the more conventional factors that motivate people at work. It can be argued for example that they do not respond directly to financial incentives, but nevertheless are keen on securing more funding for research projects. They are also interested in contributing new ideas, originality and in achieving a place in the annals of science, which in turn has engendered a deep concern over the violation of the institutional norms that have been slowly established over hundreds of years of scientific practice. As a result any attempts to promote a shift in the culture of academic dissemination that do not stem from mandatory edicts announced by research funders, may be best served by taking a more subtle approach. The combination of researchers’ desire for more funding whilst adhering to institutional norms creates a challenging barrier to surmount and so if new approaches are to be encouraged they must neither threaten routes to funding nor stray away from publishing terms and practices that researchers are familiar with. The fundamental desire for researchers to be recognised by their peers for their contributions suggests that a restructuring of the reward and recognition system that operates in academia would lie at the heart of devising any new approach to scholarly dissemination. And, as noted by Harnad in his somewhat dated, but nevertheless often pertinent paper: “old ways of thinking will not be corrected by futuristic proposals; only convincing demonstrations of the potential power, productivity and scope […] will capture the scientific community’s allegiance and participation” [22]. Thus all of these various complexities and requirements of publishing demand a tangible and pragmatic approach of investigation.

Therefore to investigate how the mechanisms of reward and recognition can be altered to influence researchers’ approaches to publishing, a series of tools are proposed that allow researchers to interact with one another and demonstrate their contributions in new ways. However, the allure of traditional routes to dissemination is very influential and will in most cases likely affect the extent to which a new method is adopted. Thus the tools are also designed to assist researchers with their day-to-day research workflows, as well as helping researchers use the large amounts of data that are available to them via the Web in a more efficient manner. It is hoped that through use of the tools, researchers should be able to build a collection of their own research and the related research of others that serves them both from the perspective of a researcher who wants to work efficiently and make use of the large amounts of data now available to them via the Web, and the perspective of the researcher who wishes to show the outputs of their work to numerous different stakeholders and minimise the time they spend on the necessary administrative tasks.

The need to provide a range of avenues through which value can be added is further emphasised by the numerous sites and metrics that are already on offer for researchers to upload their research outputs and to measure their success by; it is hoped that by offering tools that assist researchers it is possible to encourage new degrees of openness and better scientific practices through the provision of functional incentives rather than potential benefits. The advantage of this approach is that it provides incentives to adopt new methods, whilst simultaneously allowing a degree of flexibility to be built into the system which, if carefully managed, should not hinder researchers’ attempts to secure traditional publications and enhance their reputation via that route: it provides a complementary way in which to enhance one’s reputation rather than a substitute method. The path taken here is not dissimilar to that of Mendeley, the reference management platform that was able to collect large amounts of metadata concerning papers and their citations, by providing its users with an efficient and simple way in which to organise their PDF collections.

With this approach in mind, this project aims to consider the following working hypothesis: “Tools that operate in an open context and facilitate day-to-day research workflows, encourage researchers to adopt more open approaches to disseminating their work”. To investigate the truth of this hypothesis an approach will be used that introduces new mechanisms into publishing, whilst not directly disrupting existing ones. A number of “widgets” are proposed that can be embedded in to existing websites to enhance the dissemination process whilst also assisting researchers with their own concerns such as planning their work or enhancing their status in their community. Once developed, these tools will be deployed in appropriate fashion and then evaluated to gain a greater understanding of their impact upon the academic community and how they may evolve in future to better serve researchers.

The proposed tools are as follows: a “Rate My Data” widget that allows users to rate and provide feedback on a scholarly artefact that is represented by the Web page the widget is embedded in; and a provenance and graph drawing widget, “ChemConnections” that assist users in the process of mind-mapping projects as they browse resources and demonstrating their connections. The widgets are being made in collaboration with the Royal Society of Chemistry (RSC) and their website ChemSpider Synthetic Pages (CSSP); a site that provides “a free database of practical synthetic procedures, provided by the community for the community” [45]. Being a community driven website that depends on scholars from a particular field to make contributions, many of the ideas proposed as part of this project are a natural fit: the site is inherently open by being free to access and by publishing the findings of other scholars; it is looking at new ways in which researchers can contribute knowledge to the community by making research outputs that may previously go unpublished discoverable; and it aims to offer new ways in which researchers can present themselves to their peers. The proposed widgets are designed to encourage members of the chemistry community to engage with the site in the hope of achieving the sites goals and thus in turn provide the aforementioned services to its users.

A third and final tool operates outside of the RSC and provides researchers (or indeed any user) with a convenient method for making works more openly available. The “Image Redactor” allows users to simply redact documents or presentations made with the popular Microsoft Office tools, replacing images contained within with those sourced from the Web and licensed under a Creative Commons (CC) licence [13]. The purpose of this tool is to investigate user attitudes towards copyright and issues it may raise and the extent to which this might concern researchers. The tool is also interesting in that it presents an opportunity for researchers to make a previous work more open, rather than helping to create something as being open from the moment it is conceived. The ability to make previous resources open may be important when considering the paradigm shift required to affect a change in publishing; with more resources becoming available openly, the more likely it may be that a new publishing model gains momentum and users.

### Rate My Data

The purpose of the “Rate My Data” widget is to provide users with a quick and simple mechanism for reviewing each other’s works, adopting an approach that is similar to the manner in which browsers of Reddit can vote a story up or down, to ultimately determine its prominence on the site. Whilst in the academic context we may assume that users would be searching for items relevant to their work rather than looking for something others consider interesting as with Reddit, the simple rating mechanism purports to be a new way in which an author can demonstrate their value to the community via small contributions which independently may not be citable or feature in a citable scholarly artefact, but nevertheless may be of suitable significance so as to influence another researchers work. Thus the “Rate My Data” widget provides a tool through which other researchers can casually acknowledge the contributions of others and provide a metric through which authors can demonstrate their activity and value in a scientific community; a feature which may be particularly useful for those researchers who are yet to establish a publication record but have been making contributions.

As demonstrated by the wireframe represented in Figure 7, the “Rate My Data” widget display a score alongside the voting buttons that represents the aggregation of all votes cast. When the score is clicked, the voting history for the artefact is displayed, indicating which users have voted and in what direction. It is hoped that this feature will both prevent abuse of the system, with users’ voting patterns being made visible, but will also allow authors to demonstrate if influential members of the academic community have voted up their work, helping an author validate the value of their work to the field (see Figure 8 for voting history wireframe).

|  |  |
| --- | --- |
| **“Rate My Data” Widget – Start** | |
| **RateMyData1.png** | |
| **View Description**   * Small embeddable widget for rating the scholarly artefact on the page either up or down, applicable to CSSP, ChemSpider, Articles, etc. * User ratings are public allowing other readers to see who rated the artefact up or down. | |
| **View Purpose**  Allows user to rate scholarly artefact either up or down, whilst displaying the current aggregate score. Score can be clicked to display previous ratings from different users. Rated items are stored in user’s personal collection for future recall. | |
| **Accessible from**  Landing view of any scholarly artefact on sites such as CSSP, ChemSpider or Articles. | |
| **Functionality** | |
| Up Arrow | Up votes the artefact, adding +1 to its score and adding it to the user’s collection of rated artefacts, used to establish a personal record of useful artefacts and be reminded to provide a “review” of the artefact at a later date if appropriate. |
| Down Arrow | Down votes the artefact, -1 to its score. |
| Score | Represents the aggregate score of all users’ votes. When clicked displays voting history view. |

Figure 7 – “Rate My Data” Widget - Start

|  |  |  |
| --- | --- | --- |
| **“Rate My Data” Widget – Voting History View** | | |
| **View Description**   * Pop up display that lists those who have voted on an artefact and the nature of their vote (up or down). * Appears when the score on the ‘Start’ view is clicked. | |  |
| **View Purpose**  To display the names of those who have voted and the direction of their vote, which in turn may have an influence on the scholarly “weight” of the artefact (e.g. the artefact may be endorsed by a high ranking member of the community). | |
| **Accessible from**  Embeddable rating widget ‘Start’ view | |
| **Functionality** | | |
| Name + Icon | Clickable to show a user’s comment when available | |
| Scrollbar | Scrolls through list of ratings | |

Figure 8 – “Rate My Data” Widget – Voting History View

A common problem with many attempts to change scholarly discourse is the issue of new tools and methods using researchers’ precious time which could otherwise be spent either in the lab conducting or research or on pursuing traditional routes to publication that are much more likely to be looked more favourably upon. This widget has been intentionally designed to be as simple as possible so as to alleviate these concerns with users’ interactions being consolidated to a single click to leave their opinion. However the widget has opportunities to offer further advantages that add value for the user and assist them with their day-to-day research. For example the widget could incorporate a book marking feature that allows researchers to store a record of the artefacts they have voted up for future reference, helping a researcher to form a collection of useful resources.

With the “Rate My Data” widget designed to provide a very simple mechanism for conveying one’s thoughts on a scholarly artefact, it consequently lacks the ability to channel more sophisticated feedback that may be required to satisfy the rigours of academia. As a result there is a risk that little value may be placed in the scores that artefacts accrue, as elements that are perceived to be crucial to the process of academic dissemination such as peer review are missing. Thus, this introduces a number of interesting research questions, namely: how does a “Like” system, similar to that found on Facebook, operating in a scholarly context compare to the social media context; and how does an aggregated score compare with a citation, with regards to building influence and recognition? With these questions in mind, it is perhaps more valuable to consider the “Rate My Data” widget as a starting point for engaging academic communities with unconventional methods of academic dissemination.

The “Rate My Data” widget also provides an opportunity to explore other avenues of research. It could be attached to a feedback widget that allows users to provide a comment along with their vote; allowing for an interesting comparison of the effects of a longer, but more considered voting process with that of a quicker voting mechanism. This would provide an opportunity to investigate the effect of a more open and “light weight” approach to peer review. Further research could also explore the effectiveness of allowing authors to request feedback on particular aspects of their work, tailoring feedback received on these smaller scholarly artefacts and thus creating a learning process that provides the author with a greater incentive to share their work.

It should be noted that whilst the “Rate My Data” widget is planned to operate within the context of CSSP, the concept could be applied across a range of disciplines and areas of study. Smaller research outputs can be elicited across the research workflow and so “Rate My Data” could be used in a number of situations, creating spinoffs such as “Rate My Research Proposal”, “Rate My Methodology”, “Rate My Workflow”, to name just a few. However for the widget to applied in different contexts, new platforms would need to be developed that would allow the hosting of different, micro research outputs.

#### Rate My Data Progress

Initial development work has commenced on the “Rate My Data” widget alongside the RSC’s own efforts in engaging the chemistry community to collaborate online via the manipulation of reward and recognition mechanisms. Therefore, with the aim of development both to introduce a new widget yet also work alongside existing tools in development, a “Rate My Data” widget is under development which allows users to vote as they submit feedback in the form of a comment. The RSC has already developed a widget for allowing researchers to leave comments on a page that represents a research object, and “Rate My Data” currently exists as an extension to this widget, allowing the user to submit their feedback via a vote that gives a single click interpretation of their comment, be it either an up vote, down or neutral (see Figure 9).

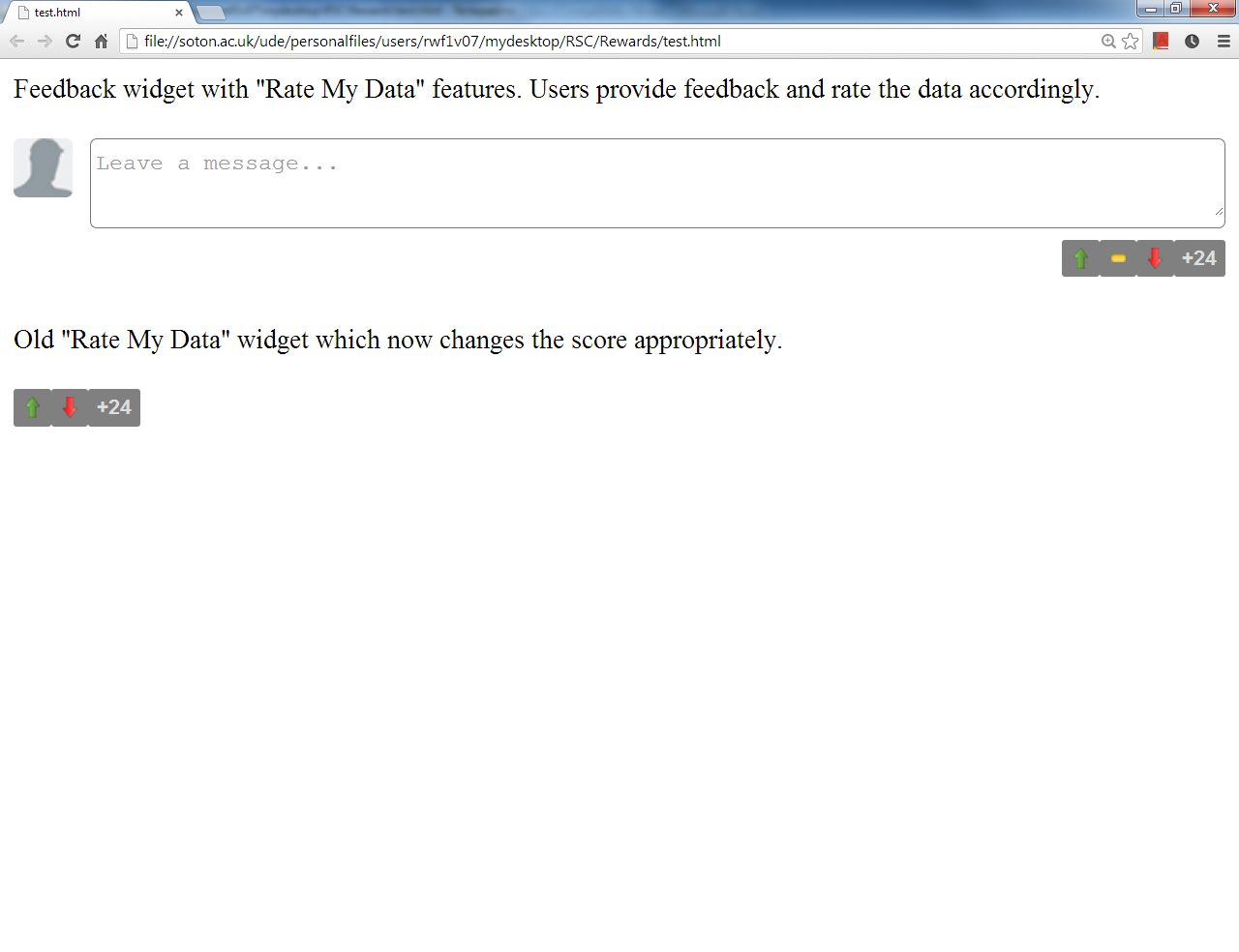


Figure 9 - Screenshot of the RSC feedback widget with submit buttons that allow a user to rate the data as they provide feedback. A choice of voting up, down or providing a neutral response is available with the overall score reflecting voting history.

The development of the widget in this manner helps ensure that not only do the new widgets comply with plans already in motion at the learned society, but also that the in this instance some qualitative feedback is given alongside the vote, a step which it is hoped will encourage a voter to reflect on the research object that they are rating and provide an opportunity to provide constructive feedback if necessary.

### ChemConnections

The second widget proposal aims to leverage the Web’s potential for disseminating research and information about research in a manner that is not simply akin to pushing PDFs and spread sheets containing data around the community, but to link all of these various resources to build a bigger picture of the research landscape. By linking different scholarly artefacts and providing meaning to these links it is possible to provide detailed information about resources and how they were derived, as well as opportunities to determine which researchers and projects are responsible for different findings, and a chance to discover related work that may further inform the reader’s research. Consequently, it should be easier to identify potential areas of collaboration among research groups; permit new ways of demonstrating the extent to which one is engaged with the academic community; and, perhaps most valuably, a platform through which to implement provenance metadata. Whilst the benefits of linking data are beginning to be seen in a range of different sectors [6], establishing links between existing datasets and then continuing this procedure on in to the future requires significant time and effort from numerous stakeholders.

Thus the “ChemConnections” widget purports to offer an incentive to provide linking data to researchers as they browse resources when planning and conducting their work. The widget is embedded on pages that represent scholarly resources (much like the “Rate My Data” widget) and assists users by providing a mind-mapping tool. Users can link the current page to any other and specify a type of connection that links the two users. This will allow users to form a network of any scholarly resource which is represented by a URI, including other researchers, reactions, papers. With these underlying connections then made available to other users it opens up the opportunities to discover resources and contacts known by other researchers.

|  |  |
| --- | --- |
| **“ChemConnections” Widget – Start** | |
|  | |
| **View Description**   * Small embeddable widget for allowing users to explore or make additions to the provenance graph for the scholarly artefact the page represents. * Allows users to view the provenance graph of the artefact. * Allows users to add the scholarly artefact to their own provenance graph. | |
| **View Purpose**  Allows users to interact with provenance graphs, either for the artefact represented by the page or to edit their own graphs in the context of the artefact currently displayed. | |
| **Accessible from**  Landing view of any scholarly artefact on sites such as CSSP, ChemSpider or Articles. | |
| **Functionality** | |
| Large image button (left) | Opens the ‘Provenance Graph Viewer’ window. Image displays an icon to represent a provenance graph. |
| Small image button (right) | Opens the ‘Provenance Graph Addition’ window. Image displays a ‘plus’ icon to represent addition. |

Figure 10 - "ChemConnections" Widget - Start

As shown by the wireframe sketch in Figure 10 the embeddable widget is once again simply designed, with just two buttons: one to view the graph for the page’s corresponding resource and another to add the resource to the user’s person graph as used for mind-mapping or demonstrating one’s influence and position in the community. Figure 11 illustrates another wireframe that would appear when the user chooses to add the resource to their person graph. The user must specify the URI of another resource which they wish to establish a connection to and define the type of relationship between the two resources.

|  |  |
| --- | --- |
| **“ChemConnections” Widget – Provenance Graph Addition** | |
|  | |
| **View Description**   * Input form for adding new items to a provenance graph. * Combo-box allows users to select a type of relationship to indicate how a new item attaches to the current artefact. * User can copy and paste a URL or DOI to attach a new object. | |
| **View Purpose**  To add new entries to a scholarly artefact’s provenance graph in a consistent and persistent manner, using URLs, DOIs and a vocabulary of relationship types, to create triples that can be added. | |
| **Accessible from**  Widget’s ‘Start View’ | |
| **Functionality** | |
| ‘Name’ label | The name of artefact on the current page; the subject of the triple being added to the graph. |
| Type of relationship combo-box | Allows the user to specify a type of relationship in which the new artefact is linked to the one on the current page; the predicate of the triple being added to the graph. |
| Text box | Allows user to input a URL or DOI which links to another scholarly artefact; the object of the triple being added to the graph. |
| ‘Add’ button | Adds the specified details to the provenance graph. |

Figure 11 - "ChemConnections" Widget - Provenance Graph Addition

Once provenance graphs have been created these may be viewable either from the perspective of the artefact or from a user’s perspective. The graphs offer a potential to apply different views to the data that has been mapped, allowing users to see the connections between scholarly resources in different contexts. For example, a user may want to view their personal network of connections to other researchers and use this to share an item of work with a number of their peers, who then in turn could be granted permission to share the resource further still among people in their network. Figure 12 demonstrates a wireframe diagram for a very simple set of connections that illustrates how resources of completely different types could be related to one another, providing the opportunity to present information on how a piece of work was conducted, by whom it was carried out (and thus may have authored more work in this area) and who the work is funded by (a potentially useful resource for funders who wish to examine the overall impact of a particular programme).

|  |  |
| --- | --- |
| **“ChemConnections” Widget – Provenance Graph Viewer** | |
|  | |
| **View Description**   * Displays the provenance graph for the scholarly artefact represented by the page. * Shows more details on the artefacts and their connections. * Can export a provenance graph as an image | |
| **View Purpose**  To allow users to view the provenance graph they have mapped (possibly with additions from other users to, if made openly available) and understand the relationships between different artefacts. Users can view more information on different relations, navigate between artefacts or export the graph. | |
| **Accessible from**  Widget’s ‘Start’ view. | |
| **Functionality** | |
| Provenance graph (left) | Allows user to pan around the graph and select different items on the graph |
| Information pane (right) | Displays more information on the currently selected artefact and it’s relationships |
| ‘Export’ button | Exports the provenance graph as an image |

Figure 12 - "ChemConnections" Widget - Provenance Graph Viewer

Once again this widget invites a number of research questions which if answered would inform on many of the issues that surround disintermediation. The widget offers an opportunity to examine what factors may motivate researchers in creating the connections that provide provenance information when they may not see any direct benefit; to establish to what ends researchers in the lab may use provenance graphs; and to see if provenance graphs are an effective medium for demonstrating one’s work. The widget, like “Rate My Data” could also have the potential to be extended to investigate further avenues of research. For example, users could link abstract concepts to scholarly artefacts, such as being able to specify the recommended audience for a resource, which in turn may lead to the development of more effective browsing methods that would be beneficial to both readers and authors.

### Image Redactor

The third tool proposed operates quite differently from the previous two widgets and does not feature as part of the RSC website, nor does it have a chemistry focus. The Image Redactor is a tool that allows users to take existing pieces of work and redact them such that they can be distributed under a more permissive licence (typically one of the Creative Commons (CC) variants). It is hoped by offering a redacting tool, authors will feel much more inclined to disseminate their work more widely online without fear of any copyright infringing material appearing in their presentations. For example, a university lecturer may be able to use copyrighted material when presenting a lecture to a small class of students, but should they wish to make the lecture more widely available after the lecture so that their students or any other interested parties could learn from it in the future, they may be prohibited from doing so because of the presentation contents. By using the tool, they could redact their presentation such that infringing images are removed and find replacement images which are guaranteed to be freely available. Whilst this task could be done manually by the presentation author, it is hoped that the convenience the tool offers makes it more likely that an author would pursue a more open approach to dissemination; and thus not only should the tool help to alleviate fears of copyright infringement, it should help to establish a greater corpus of materials that are free for others to use and build upon (CC licence variant permitting).

This tool reflects a slightly different approach to looking at some of the problems which may face disintermediation if implemented. The earlier analysis of the academic publishing market and the roles played by various stakeholders indicates that a number of burdens and tasks are supported by certain intermediaries, and many of these issues would still need to be overcome once disintermediation takes place. A publishing researcher would be faced by many barriers that may discourage them from taking the self-disseminating approach to communicating their research impact; and issues of copyright represent one such barrier, which to many researchers may initially seem insurmountable. However the purpose of this tool is to examine if an app-like structure (that is to say presenting a tool or service which carries out a singular task in a streamlined manner) is a useful approach for eroding potential barriers to publication by easily eliminating certain problems (in this case copyright concerns). By constructing a tool that works with Microsoft Word or PowerPoint documents, software which many researchers will be familiar with, the tool minimises its impact on the researcher’s normal workflow and allows them to write work in much the same manner as before, only encouraging an author to redact their work if and when it becomes necessary.

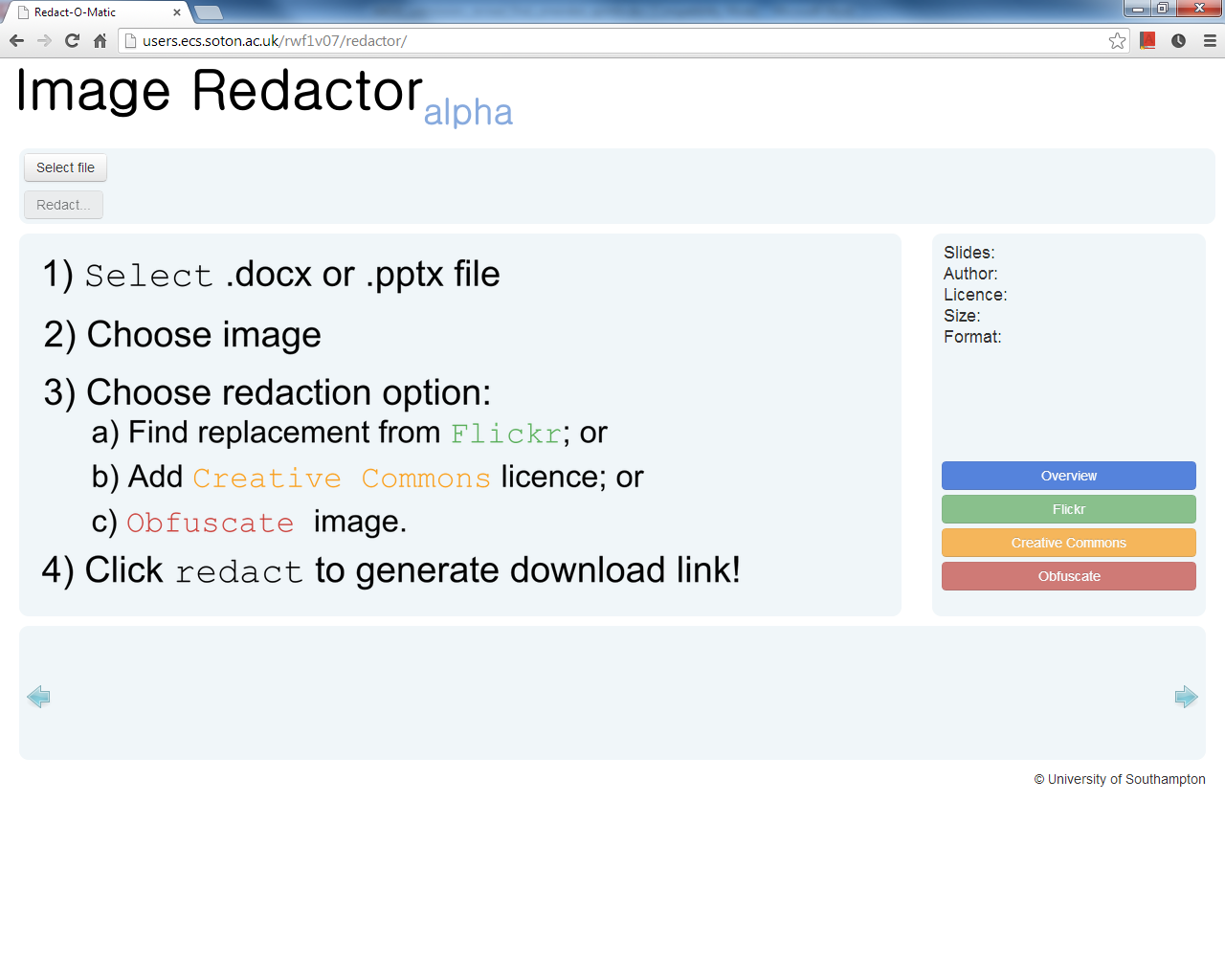


Figure 13 - The image redactor introductory screen, informing users how to use the service and the options available them to redact their uploaded document. Images can either be replaced with alternative, licenced images from Flickr, have licence metadata added to them or be obfuscated in some manner.

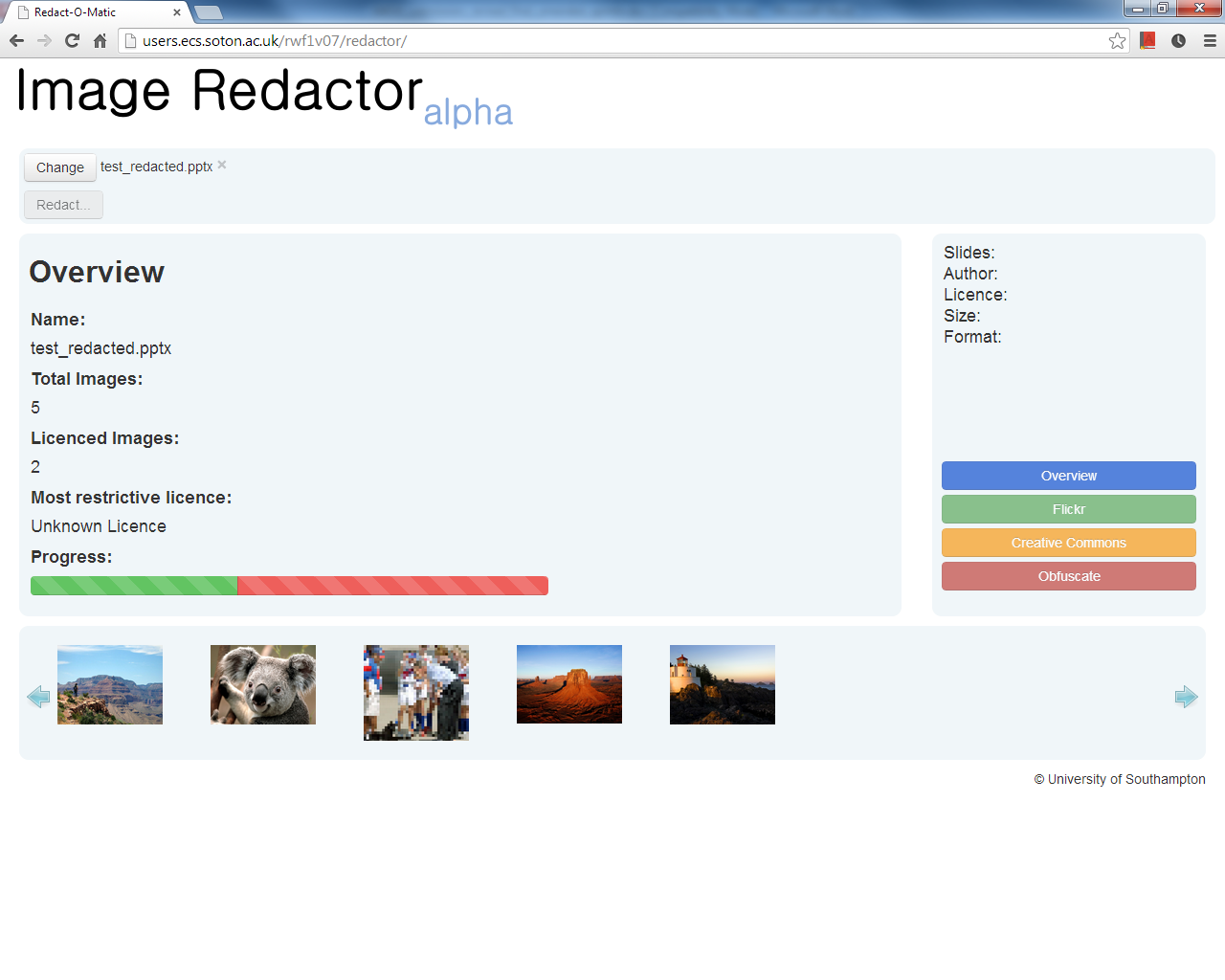


Figure 14 - Screenshot of the Image Redactor. Images from an uploaded document are displayed along the bottom of the screen, for users to select and then redact. An overview screen displays how many images have been redacted as a proportion of the total number of images in the document and the most restrictive licence found in the document.

The tool has been designed to extract images form a .docx or .pptx file (the Office Open XML file format used by Microsoft Office 2007 and subsequent editions) and assist the user in finding a replacement image that can be licensed using a CC licence. Users are presented with three options for redacting an image: they can specify a licence for the existing image, should they own the image and wish to make it freely available; find a replacement image from the image hosting site Flickr (http://www.flickr.com/) licensed under any one of the CC licences; or insert a placeholder image in the previous image’s place should the image be copyright infringing but a suitable alternative cannot be found. All replacement images have licence information embedded in the image’s EXIF metadata where possible.

Deploying the tool allows a number of research questions to be answered in relation to the model of disintermediation. Whilst disintermediation proposes a new approach to scholarly discourse, it is important to consider that more conventional approaches will continue to operate and will, for almost all researchers, be the preferred choice for publishing. This tool provides an opportunity to assess researchers’ natural propensity to engage with open discourse. By allowing an author take a resource made for one purpose and redact it so that it can be shared to a wider community, in a manner which is fairly quick and efficient, but provides no other benefits to the user other than a CC licensed version of their work, it is should be possible to ascertain the role of copyright and it’s potentially inhibiting effects from the perspective of the researcher. With a fundamental part of disintermediation focusing on a researcher’s ability to develop an online profile from which they can demonstrate their achievements and collate their work; tools that make research legally available to a wider audience will be necessary, once the publishing safety-net role of a corporate publisher and their editors has been removed.

As part of the redacting process, images from documents are temporarily uploaded to the Image Redactor server so that licencing information can be written to the metadata existing images if required. This presents yet another research opportunity, to investigate whether by providing users with a tool that assists them to one end; it is also possible to build a collection of CC licensed images that have been used in documents. Whilst repositories of CC licence images do already exist, the tool provides a simple method for uploading images and declaring a licence while simultaneously providing some other benefit. It may also be the case that if the collection of images grows, the tool may become better at helping the user to find new images and offer a more varied and rich pool of images to make use of. If it is possible to build a repository of resources that are open and clearly licensed by offering a tool that is originally designed for some other purpose, it may have significant implications for disintermediation in scholarly discourse, providing an insight in to how individuals can act for their own end, but contribute to a greater, public good (possibly encouraging the uptake of concepts such as “overlay journals” [10]).

### Conclusion

For disintermediation to be successful many of the roles that publishers perform will instead need to be adopted by researchers to ensure they are still carried out. Whilst at present researchers do still perform many of these functions, new tools will be required to facilitate the organisation that publishers provide and to exploit the new value-adding opportunities that disintermediation offers. The tools described here offer just a small example of what might be possible in the disintermediated publishing landscape, but aim to serve as starting points for researchers, learned societies and institutions to begin engaging with the principles and ideas that disintermediation embraces. The tools do not fully act as replacements for any one task that publishers might provide, but instead explore how researchers may be encouraged to contribute more information and knowledge to the community and offer a chance to reflect on the way in which they value the various mechanisms of reward and recognition.

# Discussion

The ideas and concepts proposed by applying the notion of disintermediation to academic publishing, raises a number of interesting points for discussion. Disintermediation can be interpreted in lots of different ways and will have varying meanings among stakeholders depending on how they perceive their role and the roles of others within the academic publishing market. Not only is it possible to interpret disintermediation in different ways but it can also be implemented in varying manners too, with different stakeholders undertaking varying degrees of effort to support the changes required. The widgets proposed will go some way in establishing to what extent researchers are willing to stray away from conventional publishing practices and what may motivate them, but these form only one small, iterative step to a new approach to scholarly publishing.

At the crux of the issue of disintermediation is the role of the researcher and their attitudes towards publishing, both conventional and unconventional. For a new approach to be successful the processes and their outcomes need to be valued by the researcher community and to understand this it is necessary to perceive the costs of using such a system from the perspective of the researcher and its risks as well as the value it may have or be seen to have at some point in the future. To understand the complex relationship between risk, cost, ephemeral value and future value that a typical researcher may possess requires a means by which these different concerns can be measured and an understanding of what metrics researchers use when making their decisions. It would seem in academic publishing at present, the various metrics that have arisen to measure what a researcher has produced and the fulfilment of its intended benefit or impact has become conflated over time, with metrics such as impact factors holding undue prominence over the real impact on society or academia an item of research may have. Only by acquiring a better understanding as to how one might measure the various processes that are taking place in scholarly discourse and how these compare to the metrics used by individual stakeholders (and the contradictions that may arise between different stakeholders’ metrics) will it be possible to truly manipulate academic publishing in a manner that may allow it to be embraced by those who have the ability to push through changes.

# Future Work

Having arrived at the idea of disintermediation after a consideration of academic publishing and its many facets, future work will continue pursuing the two strands of research outlined above. The model for a disintermediated publishing environment will be elaborated on further to provide a more comprehensive description as to how it might be implemented, the technologies it will need to take advantage of and the role of the various stakeholders. Understanding the role of the stakeholders is a key part to disintermediation and how these roles are conveyed to stakeholders may be a crucial factor in how a new approach to publishing may begin to be find acceptance. Future work on this strand will aim to validate the model proposed through the medium of peer reviewed articles and expert feedback, as to actually implement the model itself would be impossible. Feedback may be sourced from stakeholders from numerous positions within the model including: those who work on an institutional level, such as repository managers, librarians, and those who are responsible for research policy; representatives of learned societies who may be able to offer feedback on the role of the learned society in the new publishing model; and possibly from those who are already pushing innovative new technologies into the academic publishing market.

An element of the future work in this area will be a continuation in gaining a deeper understanding of academic publishing and how it might be possible to move away from the concerns of accessibility that OA has largely dealt with (and made significant improvements to) and start looking at other innovations that are taking place. To improve the chance of new approaches to publishing being adopted by the community, an understanding of what different stakeholders value and the metrics they use to measure this value (or perhaps the metrics they should be using but do not) would be hugely beneficial. Therefore future work will look at some of the aforementioned new and novel services that have been launched and look at their varying degrees of success and acceptance from academic communities to ascertain how these fit within researchers’ perceptions of the academic publishing sector. If successful, this work should also go on to inform the evaluation of the tools as they are developed and validate the disintermediation model by demonstrating where similar techniques are being employed in the existing publishing market.

With regards to the tools, development will continue with close collaboration with the RSC and ChemSpider to deploy the widgets in an environment where they may be used by researchers and an analysis can be conducted concerning their usage. It is hoped that the “Rate My Data” widget may be used on the ChemSpider website where it will have the opportunity to be evaluated by both the ChemSpider development team and the ChemSpider user base. Use of Web analytics will provide an indication as to the impact it has had upon the community to properly assess whether a feedback mechanism of this kind would be successful in the academic environment. Further testing would ideally focus on testing the effectiveness of the widget in varying configurations, for example to examine if more constructive feedback is provided by users if the original author can specify the area of their work they wish to receive feedback on, or if by displaying other user’s comments feedback may become more insightful. To test the long term impact ratings may have for publishing authors will be difficult to determine, thus most likely demanding the gathering of qualitative feedback from users (which may be possible to source via the ChemSpider website itself) to assess initial reactions.

The “ChemConnections” widget will likely require a significantly greater development time but early evaluation may be possible through the use of focus groups, in which researchers may be shown the early development stages of the tool and be provided with an opportunity to use it. This should provide valuable feedback and help influence the development of the widget such that data stored as triples and provenance information may be presented to researchers in a useful and engaging way. Whilst the benefits of creating linked data are well understood by the linked data community and the use of provenance in scientific research has been well documented, the development of this widget will provide a good opportunity to understand the researcher perspective on these issues, both with respect to how researchers may envisage its usage and also what motivational elements may be required to create this data in the first instance.

The third widget, the Image Redactor, is the furthest through its development stage at present and may be deployed in the near future, with the University of Southampton’s institutional repository ePrints providing a suitable environment in which to evaluate its usage. Further development will look at adapting the Image Redactor such that it has the potential to be deployed on a range of websites as a pop-up, allowing its usage to be explored on any site which may be hosting a Word or PowerPoint document. Feedback on the Image Redactor will be sought through both user and expert evaluations, to improve upon the user experience and provide features deemed useful by both researchers and the Open Knowledge community. Once again, much of this feedback may be able to be collected via the websites in which the tool is embedded.

Once both strands of the research have been evaluated and built upon, a reassessment of the disintermediation approach to publishing will be considered to understand to what extent changes to academic publishing may be feasible, how such changes should be implemented and how disintermediation may need to adapt to reflect the findings of this future research.

# References

1. Altman, L. and Melcher, L. Fraud in science. *Environmental science & technology 286*, 6383 (1983), 2003–2006.

2. Bechhofer, S., Buchan, I., De Roure, D., et al. Why linked data is not enough for scientists. *Future Generation Computer Systems 29*, 2 (2011), 599–611.

3. Ben-David, J. and Sullivan, T.A. Sociology of Science. *Annual Review of Sociology 1*, (1975), 203–222.

4. Berners Lee, T., De Roure, D., Harnad, S., et al. Journal publishing and author self-archiving: Peaceful Co-Existence and Fruitful Collaboration. (2005).

5. Berners-Lee, T. and Shadbolt, N. There’s gold to be mined from all our data. *The Times*, 2011.

6. Bizer, C., Heath, T., and Berners-Lee, T. Linked Data - The Story So Far. *International Journal on Semantic Web and Information Systems2 5*, 3 (2009), 1–22.

7. Buckley, P.J. and Casson, M. *The future of the multinational enterprise*. Macmillan London, London, 1976.

8. Butler, D. The writing is on the web for science journals in print. *Nature 397*, 6716 (1999), 195–200.

9. Cabrera, A. and Cabrera, E.F. Knowledge-Sharing Dilemmas. *Organization Studies 23*, 5 (2002), 687–710.

10. Centre pour la Communication Scientifique Directe. Episciences.org. 2013. http://episciences.org/.

11. Chan, L., Cuplinskas, D., Eisen, F., et al. Budapest Open Access Initiative. 2002. http://www.opensocietyfoundations.org/openaccess/read.

12. Commons Select Committee. Government mistaken in focusing on Gold as route to full open access, says Committee. 2013. http://www.parliament.uk/business/committees/committees-a-z/commons-select/business-innovation-and-skills/news/on-publ-open-access/.

13. Creative Commons. Creative Commons. 2013. http://creativecommons.org/.

14. Eysenbach, G. From intermediation to disintermediation and apomediation: new models for conusmers to access and assess the credibility of health information in the age of Web2.0. *Studies in Health Technology and Informatics 129*, 1 (2007), 162–166.

15. Fyson, R., Coles, S., and Carr, L. Dissemination through disintermediation. *The Third Annual Digital Economy All Hands Conference*, (2012).

16. Fyson, R.W., Coles, S., and Carr, L. AltOA: a framework for dissemination through disintermediation. *Proceedings of the 5th Annual ACM Web Science Conference*, ACM (2013), 79–88.

17. Gardner, R. Open access and learned societies. In N. Vincent and C. Wickham, eds., *Debating Open Access*. British Academy for the humanities and social sciences, 2013, 13–29.

18. Gowers, T. The Cost of Knowledge. 2012. http://thecostofknowledge.com/.

19. Gowers, T. Why I’ve also joined the good guys. 2013. http://gowers.wordpress.com/2013/01/16/why-ive-also-joined-the-good-guys/.

20. Gurstein, M.B. Open data: Empowering the empowered or effective data use for everyone? *2First Monday 16*, 2 (2011).

21. Gustin, B.H. Charisma , Recognition , and the Motivation of Scientists  ’. *American Journal of Sociology 78*, 5 (1973), 1119–1134.

22. Harnad, S. Scholalry skywriting and the prepublication continuum of scientific inquiry. *Psychological Science 1*, 6 (1990), 342–344.

23. Harnad, S. On “Overlay Journals”, “Epijournals” and “Diamond OA.”2013. http://openaccess.eprints.org/index.php?/archives/974-On-Overlay-Journals,-Epijournals-and-Diamond-OA.html.

24. Harvard University. Faculty Advisory Council Memorandum on Journal Pricing. 2012. http://isites.harvard.edu/icb/icb.do?keyword=k77982&tabgroupid=icb.tabgroup143448.

25. Hitchcock, S., Brody, T., Hey, J.M.N., and Carr, L. Digital Preservation Service Provider Models for Institutional Repositories. *D-Lib Magazine 13*, 5/6 (2007).

26. Jump, P. Open access will cause problems for learned societies’ journals, accepts Finch. *Times Higher Education*, 2013. http://www.timeshighereducation.co.uk/422395.article.

27. Kelly, C.D. and Jennions, M.D. The h index and career assessment by numbers. *Trends in ecology & evolution 21*, 4 (2006), 167–70.

28. Kim, J. Finding Documents in a Digital Institutional Repository : DSpace and Eprints. *Proceedings of the American Society for Information Science and Technology 42*, 1 (2005).

29. Kronick, D.A. *Scientific and technical periodicals of the seventeenth and eighteenth centuries*. Scarecrow Press, 1991.

30. Li, N., Zhu, L., Mitra, P., Mueller, K., Poweleit, E., and Giles, C.L. oreChem ChemXSeer : A Semantic Digital Library for Chemistry. *Proceedings of the 10th annual joint conference on Digital libraries*, ACM (2010), 245–254.

31. Line, M.B. The Publication and Availability of Scientific and Technical Papers: an Analysis of Requirements and the Suitability of Different Means of Meeting Them. *Journal of Documentation 48*, 2 (1992), 201–219.

32. Mandavilli, A. Trial By Twitter. *Nature 469*, (2011), 286–287.

33. Merton, R.K. Priorities in Scientific Discovery: A Chapter in the Sociology of Science. *American Sociological Review 22*, 6 (1957), 635–659.

34. Merton, R.K. Behavior Patterns of Scientists. *The American Scholar 38*, 2 (1969), 197–225.

35. Michaels, D. Politicizing Peer Review: The Scientific Perspective. In W. Elizabeth and R. Steinzor, eds., *Rescuing Science from Politics: Regulation and the Distortion of Scientific Research*. Cambridge University Press, Cambridge, 2006, 220–221.

36. Murugesan, S. Understanding Web 2.0. *IT professional 9*, 4 (2007), 34–41.

37. O’Connor, D. Apomediation and the Significance of Online Social Networking. *The American Journal of Bioethics 9*, 6-7 (2009), 25–7.

38. Open Scholar. LIBRE | Liberating Research. 2013. http://www.libreapp.org/.

39. Owens, S.R. Revolution or evolution? *EMBO reports 4*, 8 (2003), 741–743.

40. Pandelis, P. New forms of open peer review will allow academics to separate scholarly evaluation from academic journals. *Impact of Social Sciences*, 2013. http://blogs.lse.ac.uk/impactofsocialsciences/2013/08/20/libre-project-open-peer-review-perakakis/.

41. Peters, D.P. and Ceci, S.J. Peer-review practices of psychological journals: The fate of published articles, submitted again. *Behavioral and Brain Sciences 5*, 2 (1982), 187–195.

42. Priem, J. and Hemminger, B.M. Scientometrics 2.0: Toward new metrics of scholarly impact on the social Web. *First Monday 15*, 7 (2010).

43. Ravetz, J. Sociology of science: Keep standards high. *Nature 481*, (2012), 25.

44. Research Information Network. Finch Report | Research Information Network. 2012. http://www.researchinfonet.org/publish/finch/.

45. Royal Society of Chemistry. ChemSpider Synthetic Pages. 2013. http://cssp.chemspider.com/.

46. Russell Group. Government response to the Finch Report. 2012. http://www.russellgroup.ac.uk/russell-group-latest-news/151-2012/5324-government-response-to-the-finch-report/.

47. Sauermann, L., Cyganiak, R., and Völkel, M. *Cool URIs for the semantic web*. Kaiserslautern, 2007.

48. Shotton, D. and Ciccarese, P. CiTO, the Citation Typing Ontology. 2013. http://vocab.ox.ac.uk/cito.

49. Smith, J.W.T. The Deconstructed Journal. *Proceedings of the ICCC/IFIP Conference on Electronic Publishing*, (1997), 73–84.

50. Smith, R. Classical peer review: an empty gun. *Breast cancer research : BCR 12*, 4 (2010), 1–4.

51. Society, T.R. *Science as an open enterprise*. London.

52. Swan, A. *The Open Access citation advantage*. 2010.

53. Tay, K.B. and Chelliah, J. Disintermediation of traditional chemical intermediary roles in the Electronic Business-to-Business (e-B2B) exchange world. *The Journal of Strategic Information Systems 20*, 3 (2011), 217–231.

54. Trevorrow, P. and Volmer, D. a. Dispelling the myths surrounding the Research Excellence Framework. *Rapid communications in mass spectrometry : RCM 26*, 4 (2012), 399–402.

55. Tucker, R. Disintermediation: The disruption to come for Education 2.0. 2010. http://radar.oreilly.com/2010/05/disintermediation-risks-trends.html.

56. Ware, M. Institutional repositories and scholarly publishing. *Learned Publishing 17*, 2 (2004), 115–124.

57. Ware, M. *Peer review : benefits , perceptions and alternatives*. London, 2008.

58. Wigand, R.T. Electronic Commerce : Definition , Theory , and Context. *The Information Society: An International Journal 13*, 1 (1997), 1–16.

# Appendix A – 9 Month Report

The following is a 9 month report conducted for the Web Science DTC. It forms a comprehensive overview of the aims of the project as it stood in the first 9 months and a review of various subjects in the literature, including academic publishing, market failures, disintermediation, knowledge creation, game theory, collective behaviour theory and knowledge management. It goes on to detail some preliminary studies that guided initial research and ultimately leads to the position that PhD has arrived to at present.

Uncensoring Open Science: conflicts of open data – 9 Month Report

# Introduction

Academic publishing has the potential to undergo a revolution, one that may change scientific discourse for the better and in turn enhance the productivity of the research community. The potential exists for research to become more readily available to its public funders and more accessible and informative to the researchers that use it – and all at a lower cost than current publishing models. The tool that enables these benefits to be realised is the Web and the scale at which it allows us to communicate with one another. Yet despite the advantages that affordances of the Web allow, much of academia fails to exploit it, showing a preference instead for using the Web to push around traditional journal publications. This research aims to examine why the flexibility and scale that the Web can provide, is not fully utilised in academic publishing and to identify possible methods of instigating change, ultimately focusing on one method in particular and putting it into practice to study the results and implications.

The potential for developments in the field of academic publishing goes beyond simply adopting an Open Access (OA) publishing model and reproducing what is already available but at low cost (although this is still greatly beneficial) - what outputs are published and the form research outputs take can also be built upon. The act of publishing can be applied across all scientific endeavours, not just the final research output that is the journal article. Encouraging a culture of open data and transparency has led to numerous benefits in other sectors which may be mirrored by the academic sector – fostering creativity, efficiency and thus greater productivity. In turn, if the output of the entire academic community is to increase, and there is already considered to be a “deluge of data” created by current e-science techniques (Hey & Trefethen, 2003), new methods of presenting research and data will be vital to both manage the quantity produced, but to also make it more useful.

Despite the possibility of these advancements and the advantages they may bring, they have yet to be fully realised, for a number of reasons. There are a number of stakeholders in the field of academic publishing: there are researchers who both produce and consume content, funding bodies that pay for the research and the publishers who disseminate it. Ultimately it is the researchers who drive the entire process as they both put value into, and take value from the system and so it is the researcher that needs to be the target of any study that looks into how publishing may be changed, for it is they who will determine the success or failure of any developments.

Previous research indicates that the degree to which an academic community has moved away from the traditional journal article publishing model varies among disciplines (Björk, et al., 2010) (Swan & Sheridan, 2005), although there is little research identifying why this might be. It is apparent that some disciplines need greater open access advocacy than others and the reasons for this may be complex, deriving from sociological and epistemological differences between disciplines. With academics being influenced by a diverse range of factors, determined on a disciplinary basis, it may be wise to focus on a single discipline when investigating and testing different advocacy techniques and potential publishing tools, rather than develop a solution where one size fits all. To this end, the subject which will be the focus of this research is chemistry, a subject that previous studies have indicated have not been the most keen to use the OA publishing model (Björk, et al., 2010) and may reap huge rewards from publishing knowledge in more flexible formats. Previous research has also indicated, that whilst an overwhelming majority of scientists within chemistry believe OA will bring benefits to their field, this proportion is smaller than any other discipline (Dallmeier-Tiessen, et al., 2011).Therefore chemistry provides an interesting case study – a practical led discipline which may benefit a great deal through advances in publication techniques (with the potential for data to be published in formats more appropriate than papers of prose). Not only can a change to the publishing process become more convenient to researchers allowing them to spend more time in the lab than at their computers; chemistry also produces vast quantities of data much of which goes unpublished. The affordances of the Web allow all this data, including vast data sets produced by experiments and potentially undesired results, to be exposed, with greater transparency leading to greater efficiency. With more information available to researchers, it should become easier to reproduce results, a tenet of the scientific method that underpins all chemistry research and thus improve the productivity and quality of researchers’ work.

Tools for publishing OA research are readily available and there are plenty of OA journals to publish in, yet despite this OA does not hold an influence in the chemistry community – it is clear that a “build it and they will come” approach is not suitable. It is therefore necessary to investigate what differences exist between the disciplines so that an OA solution can be provided that best suits the needs of those who will use it the most, both in a production and consumption capacity: the researchers. By comparing disciplines alongside examining socioeconomic drivers for researchers’ behaviour it should be possible to create a tailor made solution that chemists will want to use and benefit from.

The Web is the technological development which has made all of these other potential advances possible, allowing researchers to communicate with each other freely and easily in a near limitless manner. However, as described above, this does result in social implications; the technology does not fit simply into the everyday working practice of researchers. Thus an interdisciplinary approach is required to engineer a solution from both technological and sociological perspectives, creating a tool that provides the benefits of the OA publishing model whilst simultaneously providing the functionality and scale that meets the needs and desires of chemists.

# Literature Review

The aforementioned research problem touches upon a number of different areas. Firstly there is the issue of academic publishing and its current state. What are the problems with the current methods of publishing and what other possible solutions are available? How will these be implemented and will they be recognised by researchers? The perceptions and attitudes of researchers are also important factors to consider, both with regards to old methods and techniques, but to new approaches and what they might have to offer to.

Other topics outside the realms of publishing should also be investigated. The decision making process researchers make when they have conflicting demands placed upon them – researchers need to collaborate with their colleagues and yet they also need to further their own careers. With regards to collaboration, how do researchers work in a group lab environment and manage their various professional relationships with their colleagues, supervisors and funders. An understanding of how these interactions play out may inform potential methods of how best to elicit knowledge sharing among researchers. An examination of other disciplines that have been more inclined to embrace open access publishing techniques, such as physics (Björk, et al., 2010) and computer science (Spohr & Cimiano, 2011) may also off insight by being able to identify what properties are present in these disciplines that results in researchers being favourably disposed towards actually enacting more open methods of publishing, than researchers in chemistry.

## Academic Publishing

Before further analysis of the problem and routes to a possible solution can be established, it is necessary to define precisely what is meant by academic publishing and why the current model may seem flawed to some. There are numerous methods for disseminating the results of one’s research, with the most predominant across all disciplines being to publish the outcomes in the form of a journal article.

Publishing results in journal articles has been established practice for hundreds of years, with one of the first academic journals, *Philosophical Transactions of the Royal Society*, first published in 1665 (Kronick, 1991). The form of journal articles has changed very little since – namely an introductory section that describes the problem at hand; description of an approach or methodology that provides new information on the problem; and finished with some concluding remarks. Whilst this format is undoubtedly a useful one for conveying a summary of the research carried out and has been used successfully over many years to advance scientific progress, it is becoming increasingly recognised that it is not without its flaws. The model that was originally designed to convey accounts of experiences and letters describing observations has not progressed at the same pace as the scientific developments it has enabled, resulting in vast quantities of data either going unpublished or being represented in unsuitable ways; a prime example of this being instances where complex chemical data is presented alongside prose descriptions of results.

Whilst the presentation format itself may not be ideal, the main concern with the current publishing model is one of expense. Subscribing to the myriad of journals available has become a very costly burden that university libraries have to endure; and yet the Web undermines any barriers to communication, with individuals now able to broadcast to the world at next to no cost. For hundreds of years, the only way to distribute scientific findings was either to talk face-to-face by giving a lecture or to publish the results in a scientific journal, which universities could subscribe to, to provide to their researchers. Subscription fees are necessary to cover a variety of costs, including fixed costs such as editorial costs and manuscript management; variable costs of paper, licensing and distribution; and overhead costs such as premises and management cost (The Wellcome Trust, 2004). Naturally, many of these costs are still present when publishing journals electronically, and whilst some variable costs are removed, others are added, namely the cost of hosting articles and maintaining servers resulting in only very slight savings (The Wellcome Trust, 2004). Conversely, a report published by the International Network for the Availability of Scientific Publications (Morris, 2006) states that electronic publishing can reduce costs although this will be nullified if print editions continue alongside – a product that most readers still demand (Morris, 2006). The INASP report highlights a number of other benefits to electronic journal publishing which may make it seem like a more attractive offer to publishers: international reach, opening up the demand side of the market at little extra expense to supply; greater speed of publication, allowing articles to be available as soon as they are ready; additional functionality, even basic extras such as linking to other articles improves the digital reading experience; and finally publishing online creates new opportunities and the potential for new business models, for example article-by-article publication (Morris, 2006).

To capitalise on these advantages the journal publishers must make them available to consumers too and studies have found that electronic journals do provide a number of benefits for libraries, students and researchers. Provision of electronic journals leads to a broadening of the range of journals read by students, less time spent identify and locating articles and cheaper on a per use basis (Hansen Montgomery & King, 2002). However universities and other consumers of academic publications are unable to fully benefit from the advantages brought by electronic journal publishing – despite proving to be much better value for money, the business model that requires universities to subscribe to journals is too costly, with even the most affluent institutions, such as Harvard University, being unable to support the costs of the myriad of journal subscriptions required by their researchers (Harvard University, 2012). The Web, whilst having a positive impact upon academic publishing, has also had a disruptive one, which the community has not been able to respond to in the most appropriate manner, an experience similar to that felt in any industry in which the Web has exerted its presence.

### Market Failures and Disintermediation

As previously shown, there are many advantages to going digital in academic publishing. As with all digital goods, supplying a product online results in negligible marginal costs (the cost of producing one extra unit of the good) and thus the cost of purchasing a unit should decrease as only the fixed costs of producing the initial good need to be covered. However this creates a market failure – there is no incentive to create the first instance of the good as one cannot sell at a price to recoup the start-up costs (Quah, 2003). The power of the Web to massively reduce the marginal cost of supply for a product is a phenomenon that has appeared in numerous sectors and had a dramatic effect on each, forcing numerous industries to adapt their business models.

The music industry for example can now distribute its products online and even remove the idea of purchasing a product at all with services such as Spotify, with which users can stream music but do not own it. Music publishers’ business models have adapted to reflect this, with more revenue being sourced from live performances and by taking advantage of the lower cost to distribute music (Weller, 2011). In some instances the business models have changed significantly, through disintermediation in which the artists deal directly with their customers, bypassing the service provided by the publisher (Weller, 2011). Opportunities for disintermediation have appeared throughout a number of sectors, for better or worse. For example, in health, disintermediation may be perceived as providing misinformation and can create conflicts between physician and patient (Eysenback, 2007), whereas in other sectors such as governance it is seen as key to “strip out layers of redundant or non-value-adding process and bureaucracies from service delivery” (Dunleavy, 2010). The determining factor on whether disintermediation is seen as a step forward or not, and thus affects its chances of happening, is the value-added by the intermediary. In academic publishing the intermediary, the publisher, is perceived to add value in the form of the reputation that publishing in a prestigious journal can attract, as well as providing other benefits such as the organisation of the peer review process and the editing and formatting of articles.

Where music publishing differs from academic publishing is the role the intermediaries perform. In music, they perform little function other than the distribution; they do not provide a guarantee of quality or trust in the music, consumers choose the music informed by their personal preferences. The same luxury cannot be afforded in academic publishing, when reading a research paper, one wants to be assured that the content within meets certain expectations with regards to the quality of the research and its validity, preferably before one reads the paper and reaches their own conclusions. It is this role that journal publishers point to in defence of their charging high prices for electronic journal subscriptions despite the low marginal cost of production.

Another industry that bears some similarity to academic publishing is the newspaper industry, which has also been significantly affected by the Web and its unique method of providing content. Whilst consumers often choose news sources that fall in line with their own political perspectives, the news is nevertheless a service which consumers need to be able to trust for it to be successful and useful, thus possessing a faint similarity to academic publishing. Newspaper publishers however have little choice to accept the free provision of content that the Web affords – as some news outlets provide news online for free, due to the extremely low cost of searching for other providers online (Smith, et al., 1999), consumers can find free content elsewhere and newspapers which traditionally charged for their content by selling papers, are now in a situation where they must find revenue from new sources to complement what they have lost in a reduction of paper revenues (Weller, 2011).

The Web is transforming news provision, but not just because of the economics of the Web, but once again partly due to disintermediation, with citizen journalists breaking and disseminating news on sites such as Twitter (Kwak, et al., 2010), showing that such a process is possible even when intermediaries perform an important role. However it is important to make a distinction here, as Clay Shirky does (Shirky, 2009) between form and function. The form, newspapers, is changing to a (typically) free to distribute, digital format, websites. However the function that the newspapers and now websites act as an outlet for, that is to say journalism, and it is the effect of the Web on the function that matters, not the form. Disintermediation may result in a deficit of investigative journalism, an expensive task that requires experience, tasks which citizen journalists may not be able to fully enact upon (although they may still be able to assist, by both providing information and used as a crowd sourcing mechanism to trawl through large data sources) (Houston, 2010). The same distinction can be made in academic publication: journal articles are simply the form scholarly discourse has taken, but the function of research is much more important and does not hinge upon journal articles – it does however depend upon the flow of knowledge and information for future research to build upon. Other forms of communication may still provide this function, eroding the dominance that journal articles wield within scholarly communication and in turn within the academic community itself concerning matters of reputation and funding. Where academic publishing may differ from journalism is in the source of value-adding intermediaries: investigative journalism requires news agencies that are prepared to fund expensive investigations and acquire sources and contacts; the consumers of academic publications already have these resources established being the producers themselves, creating potential for disintermediation.

### Adding Value

The analysis of the music and newspaper industries shows that disintermediation can happen in instances where the intermediaries do not add sufficient value and it is clear that all of the tasks required during the academic publishing process can be achieved without the need for a publisher. Yet clearly publishers do add value, for not only are they still perceived to be relevant in a market which has the potential to function without them, but they are thriving, with one of the largest publishers, Elsevier, earning profits totallin

g £768m, 37% of revenue, a figure which has risen year on year (Reed Elsevier, 2012). Whilst publishers are responsible for organising the peer review process and for editing and presenting academic work, their most vital contribution comes from their name – the number of publications in high impact journals is a key factor in determining an academic’s, or even a department’s success, with journals able to bestow a reputation upon those who have published therein.

At present, the degree of impact that either a journal or an author can possess is predominantly based upon citations – how many times has the article been cited by other articles, giving an indication as to how useful it has been as a seed for future research. The impact factor for a journal is calculated annually by the Institute for Scientific Information (ISI), by dividing the number of times articles published in the preceding two years have been cited, by the number of articles published in total in those two years (Thomson Reuters, 1994). However, by grouping a number of articles together in this way, the validity of the impact factor is lessened; it only represents an average article within the journal, one which in reality may not exist. As commented upon by Woodside (2009) summarising the findings of Mafulli (1995) and Vinkler (1986), “the most cited half of the articles are cited, on average 10 times more often as the least cited half…Even the uncited articles are then given full credit for the impact of the few highly cited articles that predominantly determine the value of the journal impact factor”.

Therefore, whilst journals are perceived to add value by bestowing reputation, this as an argument for retaining the current publishing model is still flawed. Other measurements of quality have are available, such as the Hirsch index, an index proposed in Hirsch’s 2005 paper, that aims to quantify the “cumulative impact and relevance of an individual’s scientific research output” (Hirsch, 2005). In his paper, Hirsch sets out the index simply: “A scientist has index *h* if *h* of his or her *Np* papers have at least *h* citations and the other (*Np* – *h*) papers have ≤ *h* citations each” (Hirsch, 2005). This has the effect of only rewarding a scientist for their cited works and not punishing them for their infrequently cited papers; however, disadvantages also exist, with the rewarding element restricted, “once a paper belongs to the h-core, it does not matter how many more citations it will receive” alongside numerous other problems as illustrated by Egghe (2010). The upshot is that articles are judged by their authors rather than the company they keep in journals. But even with this change, a researcher’s impact on their scientific community, and thus their reputation, is only measured by the articles they have authored and fails to consider the broader range of academic outputs, a pertinent issue when discussing disintermediation and knowledge dissemination via the Web.

An academic in a community may contribute far more knowledge than the sum of the journal articles authored, with knowledge being spread through a variety of different mediums; for example blogging has become a very popular tool for disseminating ideas, be they either short comments, reviews of other academics’ works or more traditional, journal style articles (Mortensen & Walker, 2002). The more open and dynamic nature of some of these forms of self-publication, attract a greater number of comments and generate more discussion than is possible with journal articles (or if possible is less likely to attract a vibrant community), which has resulted in academic discourse thriving in the blogosphere and social networking sites such as Twitter thanks to the high speed communication and feedback that it can provide (Mandavilli, 2011). The validity of discourse in these environments may be challenged however; the Web being such an open forum allows anyone to make comments irrespective of their knowledge on the topic they are remarking upon and as such may fail to gain sufficient respect as a medium for scientific discourse. Whilst efforts are being made to encourage researchers to use channels such as Twitter, including the publication of a “guide to using Twitter in university research, teaching, and impact activities” (Mollett, et al., 2011), it is very difficult to measure the value a researcher adds to the scientific community by contributing various sized and scoped chunks of knowledge.

If less conventional contributions are to count towards a researcher’s reputation then new metrics will need to be devised, perhaps based on algorithms, which are already proven to work well on the Web, such as Google’s PageRank algorithm which is used to return the results for its market dominating search engine (Page, et al., 1998). With more journal articles alone appearing on the Web, it may be sensible to adopt the PageRank approach to assessing quality of search results (Bollen, et al., 2006). Bollen, et al. (2006), after trialling PageRank as a journal impact metric on the dataset of the 2003 ISI Journal Citation Reports and comparing with the ISI Impact Factor, identified two types of journals: there are “popular journals” which are “cited frequently by journals with little prestige” and thus have a high impact factor; and there are “prestigious journals” which are “not frequently cited, but their citations come from highly prestigious journals”, possessing high PageRank values. Ultimately, a pattern between the two types of journals was not found and the paper concludes recommending a variety of journal status metrics to measure impact rather than just one (Bollen, et al., 2006). It is clear that if one metric cannot adequately indicate the impact different journals have on their respective fields, one metric certainly cannot be used to measure an individual’s contribution. As such an array of metrics should perhaps be used to create an index for researchers and their contributions across a range of different media, although the implementation of such a system would be a complex task and potentially difficult to reach a consensus upon.

The topic of discussions arising on Twitter leads on to another area where journals add value during the publishing process, they organise the peer review process which is considered to be essential in ensuring that only valid and innovative research is published. Yet the peer review process has come under much criticism in recent years, as numerous examples have been found where mistakes that lead to erroneous conclusions have been published (often going on to have a large impact when the results reach mainstream media publications such as tabloid newspapers) (Colquhoun, 2011) and professionals in some fields having been reported as claiming “cliques” are forming which are suppressing some works and supporting inferior ones (Henderson, 2010). One of the bigger problems with peer review that has been discussed for some time is that of confirmation bias: reviewers are much more likely to agree with arguments that support commonly held beliefs and theories, making it much harder for new, disruptive ideas to be published and gain a foothold (Mahoney, 1977). Mahoney (1977) recommends training academics to be better peer reviewers may help alleviate the problem, but admits it would not be a perfect solution, and it becomes weakened further still when considering the far greater number of papers being published today than compared to 35 years ago; training all peer reviewers would be a costly exercise and with such a large body of reviewers across all journals and domains, mistakes will still be made. And, as Mahoney (1977) observes, a consensus may still be prejudiced if the pool of researchers share the same “ideological or methodological biases”. One study has also demonstrated the lack of attention given to the peer review process by resubmitting 12 already published psychology papers by authors from prestigious departments published in highly regarded psychology journals. When resubmitting, the papers were given fictitious authors and departments, but sent to the same group of journals, whereby only 3 were detected as resubmissions and 8 of the remaining 9 papers were rejected (Peters & Ceci, 1982). The reasons given for rejection were typically “serious methodological flaws”, but as the papers were the same as ones previously accepted, this would point towards a bias in favour of recognised and respected names and departments (Peters & Ceci, 1982).

It can therefore be argued that journal publishers are failing to add value in this part of the publishing process either. The actual act of peer review is conducted by academics and has notable flaws alongside a lack of accountability, which together can result in poor articles being published and biases creeping in. As already touched upon, Web services such as Twitter can be used as a channel for debate about research and papers, and have been shown to “weed out sloppy work” that could previously had a major impact on the field in question (Mandavilli, 2011). Whilst the many eyes of the Web can be an efficient and thorough approach to the peer review process, it can nevertheless only happen after a paper has been published for all to see and thus where it has the potential to already make an impact. A shift from pre to post publication peer review may not be a welcome change in academic communities, who are content with their present mechanism, which may work albeit with imperfections, and at the very least prevents a deluge of publications from which it might be difficult to identify the high quality pieces of work (Smith, 2010). One international survey of academics found that post-publication peer review is viewed as a supplementary method for validation and verification rather than a substitute for the current review mechanisms, due to its tendency to “encourage instant reactions and discourage thoughtful review” (Ware, 2008).

Nevertheless, post publication peer review is an emergent behaviour that can be accredited to the Web, and if it can help to provide a complementary approach to the current peer review system, which has been known to have its failings, then it is an approach which should be encouraged and recognised. The ability of social media sites such as Twitter to provide a platform where members of any academic community can contribute their feedback is a powerful tool for both authors of papers and those that read them. As such, use of communication channels such as these should be encouraged, but also monitored to ensure that feedback is appropriate and acted upon where necessary. If disintermediation is to spread in the academic publishing market, then this process must happen at all levels; opportunities for post publication peer review should be available and able to be acted upon. The openness of these post-publication review platforms is of critical importance in fostering a discussion: as commented upon by Jonathan Eisen, academic editor-in-chief of PLoS Biology, from (Mandavilli, 2011), “Who in their right mind is going to log on to the PLoS One site solely to comment on a paper? I guarantee that there are more comments on Twitter about a PLoS paper”. However, this does not mean that journal publishers may not be able to have a role. Comments made would need to be monitored, organised and structured to increase their utility, as well as ultimately be acted upon where appropriate.

Whilst there are few identifiable areas where journal publishers add value during the process of scholarly dissemination and seemingly do not add sufficient value to demand the subscription fees that are charged, there are opportunities where value can be added and yet is not at present, as has been observed with the potential to help facilitate and act upon post publication peer review. Publishers remain significant players in academia, yet this is largely attributed to a time when publishers were essential in collecting works together and making them available to those who are interested. However, as we have seen in other markets, pre-existing players need to adapt to survive the disruptive forces of the Web and thus journal publishers need to find new ways in which they can add value, justifying their presence and the fees they charge, now that academics have the potential to disseminate research themselves.

A representative of the RSC discussed the future of journals and how they might adapt at an RSC Roadshow event held at the University of Southampton in February 2012. They indicated that in the future, the time to publication would be significantly reduced, with an article appearing in print within 40-50 days of being accepted and would appear on the Web the day after acceptance. Such practice shows that the journal publisher does recognise the fast pace of communication that the Web enables and is allowing research to be made available whilst it still has time to make its biggest impact. The RSC also identifies several areas where journals can be found to add value to the publication process, primarily a journal offers: registration and quality; certification; dissemination; archiving; and discoverability. It is with the assuredness that journals offer these services through which they are likely to maintain market dominance in providing a channel for scholarly discourse; substitutes do exist but may struggle to gain trust of the consumers.

A further opportunity for journal publishers to add value is by helping to shape and produce the journal articles of the future. At present, articles follow a standard format that has been used for hundreds of years, yet with digital technologies this is no longer necessary; findings can be shown in as great or little detail as desired by the consumer. The RSC has its own perceptions on the article of the future, which includes links to authors for finding related works and figures which can be downloaded and examined in greater detail alongside the text. Whilst these are steps forward from simulating journal article digitally on a screen, further advances could be made. The granularity of the information being communicated could vary, as shown to great success by examples such as the coexistence of micro blogging services such as Twitter alongside traditional blog services provided by Blogger and WordPress as various channels for self-publishing; and online music outlets such as iTunes and Amazon MP3 that allow for the purchase of individual music tracks or whole albums. The ability to publish chunks of scholarly material at a size appropriate to the work in question may result in a lot more information being made available and utilise the long tail property of the Web.

## Knowledge Model Framework

The process of knowledge generation functions differently across a range of academic disciplines. In some instances this may be expected and reflect the varying natures of the subjects being studied; for example, the natural sciences investigate different types of phenomena to the social sciences and thus different methodologies are required. Differences also exist between some disciplines which are broadly similar, for example, physics and chemistry: both try to observe and understand the nature of the universe, yet nevertheless have very different approaches to how they reach this understanding. Years of academic study have resulted in the “siloing” of disciplines (Miller, et al., 2008) each of which embrace socio-technological developments, such as the Web, in their own way, exploiting the affordances of the Web to differing extents.

The aim of research in any discipline is to further the understanding of the discipline, creating new knowledge which is then tested in some away to assert its validity before being disseminated so that future research may be conducted, thus building up understanding iteratively. Elements of this process may be enhanced by the Web, but only by taking a holistic approach to knowledge creation in disciplines is it possible to identify where the Web can make a difference and in what manner. By modelling this flow of knowledge for a number of different disciplines it is possible to not only show where the Web applies (and also where its absence may be felt) but also to compare disciplines and identify what elements in one discipline’s knowledge cycle may be appropriated for use in another.

### Framework Foundations

Due to the iterative nature of research, it is fair to assume that any model that describes how knowledge may be created would be cyclical: knowledge is created, represented and disseminated so that it can go on to inform future knowledge generation. Therefore the ideas which are represented in one phase of the cycle go on to inform the other: as knowledge is created, its origins determine how it is represented, and it cannot be disseminated to others until it has been formalised in some way so as to give it meaning in a wider context. This cyclical nature lies at the core of the knowledge model framework.

Epistemology, the branch of philosophy that examines how knowledge is acquired (amongst other things), forms the basis of the knowledge creation aspect of the framework, proving ideal as each discipline possesses its own epistemology that sets out how new knowledge is generated. Features of epistemologies include such ideas as to whether knowledge is *a priori* (deriving from “rational insight into abstract objects”) or *a posteriori* (based upon empirical observations) (Landesman, 1997, p. 160) with both being employed for different needs within a discipline. Different disciplines may also utilise different branches of epistemology, such as empiricism, rationalism and constructivism.

Having established the grounding of a discipline using philosophy, the creation aspect of the model also describes how these manifest themselves in reality. For example, in the natural sciences knowledge is typically created through the processes associated with the scientific method, with experiments being planned and repeated to check their validity (Repko, 2008, p. 104). The results from these manifestations can then be represented in an agreed upon formal matter, the next phase in the knowledge framework allowing them to be disseminated to other members in the wider academic community - the third element of the model.

Within academic communities, standards for disseminating research have evolved over time, with the form of journal articles assuming a dominant position for many disciplines (except possibly for books, which are typically reserved for foundational knowledge rather than cutting edge research). However other methods are available, with academic conferences being prominent in some disciplines as well as less conventional methods of publishing data which are now possible, such as the publication of research objects. It is worth noting that this area of the model brings in a new phase for the framework, as the dissemination of knowledge brings focus on to a discipline’s relationship with society. Concerns of epistemologies, data gathering and representing knowledge are of relevance to those within the field; however the outcomes and impact of research and its application has a wider significance. This in turn brings a new element of the knowledge cycle which concerns how the discipline operates in relation to society, such as how the process of creating knowledge is funded and the organisations which conduct research, including their internal structures and motivations.

This element of the cycle, labelled “production” (with its focus on knowledge creation as part of a much larger, socio-economic system) concerns itself with the manner in which the various communities that make up the discipline are structured and how these various communities may be funded. As found by the *Information handling in collaborative research* report (Research Information Network, 2009), factors such as the “intended outcomes” of different communities, the size of organisations conducting research and simply just geographic location can all have an effect upon the course of the research. Such factors, whilst they may not have a direct impact on the discipline’s epistemologies, they do have an impact on how research is communicated and may present some barriers which hinder the free flow of information.

### The Framework

Figure 15 depicts the relationship between the aforementioned phases of the knowledge cycle: creation, representation, dissemination and production.

Figure 15 – The flow of knowledge. Knowledge is created, before it can be represented in some form, allowing it to be disseminated, ulitmately leading it back into its creation process. The way in which a discipline is structured within society is reprsented by “Production” which is both affected by the creation and dissemination of knowledge, and itself has an effect upon dissemination, with the makeup of the community affecting the way in which research is distributed.

This model may be applied to any discipline; however the contents of each phase of the model may vary among disciplines as they detail the inner workings of each discipline as well as describing the discipline’s relationship with society at large.

### Knowledge Creation

Knowledge creation is split in to two subcategories: epistemology and practice. The epistemology subsection details how a discipline may (or may not make) use of *a priori* and *a posteriori* knowledge and which branches of epistemology are utilised by the discipline and how. The practice element concerns what research methodologies are employed by the discipline, for example how qualitative or quantitative studies may be employed.

1. **Creation**
   1. **Epistemology**
      1. **A-priori**
         * **Examples (e.g. theoretical physics)**
      2. **A-posteriori**
         * **Examples (e.g. previous experiment observations)**
      3. **Branches (e.g. empiricism, rationalism, constructivism)**
         * **Examples (e.g. rationalism – theoretical physics; constructivism – experiment plans; empiricism – experiment observations)**
   2. **Practice**
      1. **Research methodologies**
         * **Examples (e.g. interviews, surveys, experiments, ethnography)**
      2. **Aims**
         * **Examples (prove theory, create patents, draw conclusions)**

### Representation

The representation section is used to illustrate how knowledge within a discipline is stored and used and is strongly influenced by the knowledge creation section.

1. **Representation**
   1. **Form**
      1. **Example forms method can take and associated information (e.g. experiment plan, workflows, metadata, prose)**
   2. **Results**
      1. **Examples results (e.g. graphs, datasets, research objects, observational notes, audio recording)**
   3. **Models**
      1. **Example model features (e.g. qualitative elements, quantitative elements, mathematical foundations)**
   4. **Laws**
      1. **Grounding for law (e.g. mathematics, prediction)**

### Knowledge Dissemination

Dissemination is concerned with how the representation of knowledge is communicated to other researchers in the discipline and so is influenced by both how knowledge is typically represented by the discipline, but also by how the discipline operates on a broader scale (as well as having an influence on this wider picture itself).

1. **Dissemination**
   1. **Papers**
      1. **Paper Medium (e.g. journal articles, repositories, preprint archives)**
         * **Access Methods (e.g. open access, subscription)**
   2. **Data**
      1. **Data Medium (e.g. prose, online, Semantic Web)**
         * **Access Methods (e.g. open access, subscription)**
   3. **Conferences**
      1. **Conference Purpose (e.g. networking, publishing)**
   4. **Patents**

### Knowledge Production

The final section of the model relates to how knowledge is produced from the perspective of society and economics rather than the discipline itself. As such it has an effect on the workings of the discipline and how knowledge is disseminated, as well as being influenced by these factors too.

1. **Production**
   1. **Community**
      1. **Organisations (e.g. universities, institutes, industry)**
         * **Funding Sources (e.g. charities, research councils, profits)**
      2. **Community Structure**
         * **Research drivers (e.g. prove theory, earn profit)**
         * **Community cohesion (e.g. collaborative, competitive, geographically distributed)**
   2. **Stakeholders**
      1. **Reason for interest/Stake (e.g. profit, reputation, curiosity, impact)**

## Framework and the Web

The primary purpose of the framework is to identify not only how knowledge is created and disseminated through a discipline but also to see how and where the Web and its affordances can be utilised. Academics within disciplines can already promote themselves without using the Web so it may be considered unreasonable to expect that the Web has value as a medium for distributing knowledge by default. There is no doubt however that it can ensure knowledge reaches as many people, as cheaply as possible, than any other form of communication previously and has great potential for use by any discipline.

The following list illustrates a number of ways in which the Web can have an impact upon the research generation process, highlighting which sections of the model may see changes as a result (thus highlighting which utilities may be provided together to provide a more unified and useful experience).

* Provision of previous research. Typically achieved through journal articles and conference proceedings. Whilst these are available on the Web, often they are behind pay walls, limiting access, a problem which is more common for industry than it is academia (Research Information Network, 2011). Open access publishing removes this problem although has other problems and perceptions associated with, such as a lack of support from the publishing industry (who would lose revenue) and a lack of trust from academics who are wary of a lack of rigorous peer review process (which is not always fit for purpose in the journal publishing system (Smith, 2010)). **Sections effected: Creation, Dissemination**
* Previous research may be presented in formats which lend themselves better towards creating new research. This may be achieved through Semantic Web technologies allowing different aspects of previous research to be organised effectively, such as the system proposed by oreChem (Borkum, et al., 2010). Metadata, such as provenance information may be provided to make published research clearer to understand and easier to reuse. **Sections effected: Creation, Representation**
* The Web may allow researchers to get in contact with authors of previous research or other researchers wishing to collaborate or who are working in a related field. This process could be aided using Semantic Web technologies, linking research to researchers and other works which may have cited it. **Sections effected: Production, Representation, Dissemination**
* A Web service in which the research registers what they are currently researching and are notified of changes in the topic as they happen – including new papers which have been published, new projects being started, papers which are being cited by others in the field and relevant conferences, seminars and lectures. Could synchronise with other services such as LinkedIn and Twitter. **Sections effected: Creation, Representation, Dissemination, Production**
* Various documented benefits of eScience, such as better handling of the “data deluge” and collaboration among research groups (De Roure & Frey, 2007). **Sections effected: Creation, Representation, Production**
* Social networks for sharing and comparing plans for various investigations, such as myExperiment (Goble, et al., 2010) which can also be published as products of research. **Sections effected: Creation, Representation, Dissemination**
* Online communities which have the effect of joining disparate communities such that they can collaborate and work together to achieve a common goal, taking advantage of each others’ natural, competitive advantages and resources. **Sections effected: Dissemination, Production**
* Crowd sourcing can be used to both provide data for research and be used to handle tasks which computers are not so adept at tackling (Kittur, et al., 2008). **Sections effected: Creation, Production**
* Context aware publishing, providing appropriate viewers with appropriate content, creating a publishing system with a greater degree of flexibility. **Sections effected: Dissemination, Production**

### Framework Comparisons

In the *Patterns of information use and exchange report* (Research Information Network, 2009), whilst studying a number of case studies over the course of 5 days, the report investigators identified a number of activities taking place in accordance with Charles Humphrey’s Knowledge Transfer Cycle.

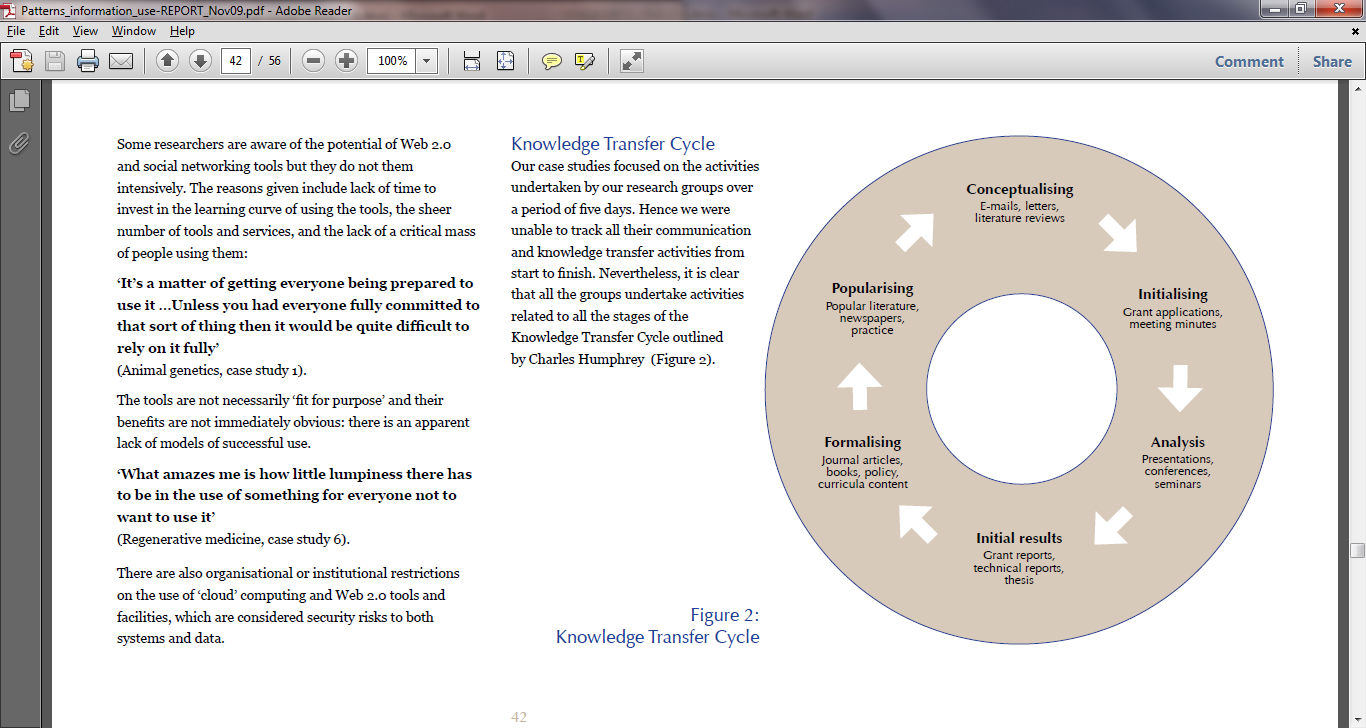


Figure 16 - Charles Humphrey's Knowledge Transfer Cycle

This model views the flow of knowledge from a researcher’s perspective and how knowledge is used by a researcher throughout the process, as opposed to the broader, interdisciplinary approach taken by the knowledge model framework proposed above. However the two models do coincide in certain aspects, with many of the activities described in the Humphrey’s cycle being applicable to phases in the framework. The notion of “Conceptualising” fits nicely into “Creation” phase of the framework which states what is required for the foundations of new research. These newly identified activities stem from the researcher perspective taken by the Knowledge Transfer Cycle, thus revealing more informal influences on the creation process such as emails and Skype calls, rather than the stricter demands as set by the discipline’s epistemology that the knowledge model framework is influenced by. The formalising of initial ideas into a more solid grounding for research is however covered by the next phase in the cycle “Initialising”.

The *Patterns of information use and exchange* report (Research Information Network, 2009) makes a case for adding a new stage in to the cycle called “Pre-analysis” alongside the existing “Initialising” stage. This phase takes into account issues of planning for an experiment, in which much information is shared. Examples of the flow of knowledge as given by the report include what consumables may need to be ordered, health and safety information, equipment and reagent information from industry and demonstrations by colleagues in training sessions (Research Information Network, 2009). It could be argued that the knowledge model framework encompasses these elements in the “Representation” section, which allows for knowledge such as experiment plans to be accounted for, but there may be value in broadening the spectrum of types of knowledge that may be represented by the framework.

The next phase of the Knowledge Transfer Cycle is “Analysis” and concerns presenting research, along with the “Initial Results” phase. These ideas are covered in “Dissemination” in the framework, but once again a discrepancy emerges between the degrees of formality for different types of dissemination. The Transfer Cycle talks of presenting research to colleagues, whereas the framework details publications such as conference proceedings and journal articles. The Transfer Cycle covers these methods in “Formalising”.

Finally the “Popularising” phase focuses on the process of work changing over time from being cutting-edge research, to mainstream, accepted knowledge. This is conveyed through the research being taught in school, or reported in the popular press, and ultimately being put into practice. Whilst elements of this are included in the framework’s “Creation” section, with knowledge being used as the grounding for future research, this can be further expanded upon in “Production”. At present, the focus is on how the discipline’s community is organised, how different groups interact and how research is funded. However the discipline may have a wider relationship with society its relationship with the media and the extent to which it is engaged with the public for example.

*Studies on Subject-Specific Requirements for Open Access Infrastructure* (Meier zu Verl & Horstmann, 2011) identifies several common steps for research lifecycles: “data collection, processing enriching, archiving and re-using”. By observing work groups from a range of disciplines, the report’s authors not only determine how researchers make use of knowledge in their day to day work, but also how it is used collaboratively, thus providing a picture as to the structure of the wider research may be disseminated to other researchers, but also considers that data may be reused before it is formally published in some way and re-used by others from different disciplines (for example data collected as part of ICT research may be re-used by social science or humanities). The temporal element of dissemination is one which is missing from the framework, as well as an opportunity to enumerate examples of other disciplines’ use for research outcomes.

However the report only examines academic communities from the perspective of their primary goal, conducting research and scholarly work. It fails to look at other elements, such as those predicated by sociology, psychology or economics, which are also required to design open access systems. It is not sufficient to simply establish user requirements from infrastructure – it is also necessary to establish how this infrastructure will fit within the community and to try and predict the consequences of its introduction upon those that use it and those that are affected by it.

### Future Work

Having compared the knowledge framework with a couple of other approaches to modelling the flow of knowledge, it is clear that there are still many factors in the creation, dissemination and uses of knowledge to be accounted for. Therefore, future work will examine the different perspectives that can be applied to the four phases of the framework, adding a number of dimensions to each section. As we have already seen, the model paints a different picture depending from whose perspective the flow of knowledge is observed, different stakeholders have different interests: the researcher will be focussed on the practicalities of knowledge generation, whereas the research council funding the research may have different priorities, such as identifying impact and value for money. However, it is likely that different stakeholders will not all be interested in different sections of the model, thus the dimensions may not be applied universally across the model, but only where appropriate.

## Game Theory

Analysis of disciplinary differences has led to the conclusion that the products of much chemistry research possess a not insignificant commercial value. Chemistry has a lot to contribute to profitable industries such as medicine and energy and as such the research environment is quite competitive, with charities and industry also conducting research and patenting their outputs. With a researcher’s career dependent on the research they produce, a competitive environment may result in some researchers feeling reluctant to disclose their knowledge until they can do so under the protection provided by a journal. However, this protection comes at a cost, the journal will typically hold the copyright for the piece of work, preventing the original author from distributing it as they see fit.

These attitudes towards sharing one's work emerged from the responses given by a preliminary study conducted at the University of Southampton Chemistry department, a questionnaire distributed to early career researchers. When asked about the impacts of publishing research outputs online and not in a journal, participants responded with concern about IP theft coming from not just the commercial sector, but also their fellow researchers, citing the “novel and commercial” aspect of some research being the key issue. Such attitudes also emerged from a much wider distributed survey in the Advocacy to Benefit from Changes report, which conducted a nationwide survey of chemists under the auspices of the RSC and EPSRC (Burns, et al., 2009).

However if chemistry researchers are to share more of their work, either be it between lab partners or more widely, it is important to understand the social dynamics at play that underpin their decision making. Game theory aims to simulate what decisions players will make under certain conditions and can be used to determine what factors are affecting certain situations, and ultimately be applied to lessen or enhance certain aspects in a scenario to encourage the desired choices. In this context, chemistry researchers are faced with a choice as to whether to publish their work and share it with colleagues or to keep it private until a later stage where it can be published in its complete form in a journal; in short they may either collaborate or compete and a successful chemist will strike the right balance between the two. This dichotomy is mirrored by the common game theory scenario of the prisoner’s dilemma.

The prisoner’s dilemma involves two players who have allegedly committed a crime. They are taken to separate rooms (from which they cannot communicate with one another) and asked to either confess or deny to committing the crime. The resulting prison sentences for each player vary depending on the combination of responses given, as shown by Table 2.

|  |  |  |  |
| --- | --- | --- | --- |
|  | | **Player 2** | |
| **Confess** | **Don’t Confess** |
| **Player 1** | **Confess** | **5,5** | **0, 10** |
| **Don’t Confess** | **10, 0** | **1, 1** |

Table 2 - Prisoner's dilemma payoff matrix, showing the years in prison for each player based upon their response and the other player's.

The above payoff matrix indicates that a player’s dominant strategy would be to confess to the crime, resulting in them either getting 0 or 5 years in prison (as opposed to the risk of getting 1 or 10 years). However if both players choose to confess they get 5 years each – should both players not confess they will just get 1 year each. The option where both do not confess is the socially optimum outcome, and yet will not emerge if both players take their own personal, dominant strategy. Without knowing what the other player is choosing there is an inherent risk in taking the “Don’t Confess” option, and even if both players can communicate with one another, there is an issue of trust: will one player keep their word or exploit the trust of the other.

Morphing the prisoner’s dilemma to fit the chemistry perspective, players have a choice to share their data or to keep it to themselves. This situation is represented in Table 3.

|  |  |  |  |
| --- | --- | --- | --- |
|  | | **Player 2** | |
| **Share** | **Don’t Share** |
| **Player 1** | **Share** | **8, 8** | **3, 10** |
| **Don’t Share** | **10, 3** | **2, 2** |

Table 3 - A payoff matrix for chemists in possession of some data. The numerical values are arbitrary but are intended to represent some sort of quantifiable benefit of the potential actions.

The above payoff matrix has been filled with arbitrary values which have been designed to reflect the value of various benefits which may be derived from taking a certain course of action. Should both players choose to share their data, then the socially optimum result is achieved; both get to build upon their own work, whilst potentially benefitting from the work provided by the other player (potentially also opening up opportunities for feedback and collaborations). However, the payoff is greater for the player who does not share whilst their counterpart does – they receive the advantage of seeing another player’s work, but are the sole beneficiaries of their own achievements. When both choose to share, neither player will receive any kind of advantage with regards to rewards such as promotion as both are in the same position (other than natural biases which may exist, that is to say one player may be better at their job than the other). For the player who chooses to share, whilst the other is not does still derive some benefit through any opportunities or feedback that may arise from having work in a more public domain. In the final situation where neither player chooses to share, the benefits received are merely a product of the players’ own work being available to them and no advantages of possessing information about other works are created.

A similar problem arises in the world of business, in which corporations find themselves in a state of needing to both compete and collaborate with fellow corporations in a market. These conflicting tensions were given the name “co-opetition” by Loebecke et al. (1999) in a paper which makes proficient use of the prisoner’s dilemma to illustrate how decisions are made when contradicting ways forward are presented. Sharing knowledge among organisations is a complex issue without any simple solutions, principally because normal market mechanisms fail when dealing with knowledge transfer – an issue which may arise in the academic context also. A potential buyer of knowledge is uncertain of its true value until the knowledge has been revealed, upon which moment they no reason to pay for it any more (Buckley & Casson, 1976). It is therefore important, in any knowledge exchange environment, for a preview of the knowledge in question to be made available to lower the risk associated with purchasing something which could be of little use.

At present, academia solves this problem by providing abstracts alongside journal articles or conference proceedings. However, knowledge for which there is little demand is still pushed on academic institutions in the form of journal bundle subscriptions, in which a group of journals is purchased in one transaction, with some of the bundled journals having much more worth than others. The issue of providing some information, albeit not all of it, may also need to be addressed when considering units of knowledge that do not fall within current conventions; for example vast datasets that may have a variety of different uses, or much smaller items shared among a small group of colleagues.

Loebecke et al. (1999) use the prisoner’s dilemma to try and determine when knowledge will be shared in this environment in which traditional economic mechanisms break down. Table 4 shows a payoff matrix adapted from (Loebecke, et al., 1999) to fit the academic publishing context, breaking down the factors involved that may influence a decision.

|  |  |  |  |
| --- | --- | --- | --- |
|  | | **Player 2** | |
| **Publish** | **Don’t Publish** |
| **Player 1** | **Publish** | **2r, 2r** | **r, 2r + va** |
| **Don’t Publish** | **2r + va, r** | **r + va, r + va** |

Table 4- Payoff matrix adapted from (Loebecke, et al., 1999) to illustrate the publishing dilemma. (r = basic value of knowledge; va = value added by monopolistic knowledge)

Loebecke et al. (1999) argue that knowledge sharing is only beneficial when there is a high basic value (the value of the work’s content) and low monopolistic value (the value of withholding the work to the individual). It could be considered that holding the work to one’s self is not deemed worth the potential reciprocation that may follow in exchange for making an initial giving gesture. More concisely: collaborative behaviour arises when *r* > *va* and competitive behaviour, when *r* < *va*.

However was originally designed with a business context in mind, the publishing context may provide further ways in which the various benefits can be broken down. For example, the publishing researcher may receive valuable feedback on their work and opportunities to collaborate further, essentially improving its basic value. It is also possible that certain benefits may only be gained when both works are published, allowing them to build upon each other or complement each other in such a way that new knowledge is discovered. This phenomenon is described more formally by Loebecke et al. (1999) as synergy: “the extent to which cooperation yields additional value from interdependent knowledge sharing beyond the sum of the parties’ individual knowledge”. Benefits such as these are represented in Table 5, included as *f*, representing feedback on published work; and *s*, the synergy produced by both works being made available together.

|  |  |  |  |
| --- | --- | --- | --- |
|  | | **Player 2** | |
| **Publish** | **Don’t Publish** |
| **Player 1** | **Publish** | **2r + f + s, 2r + f + s** | **r + f, 2r + va** |
| **Don’t Publish** | **2r + va, r + f** | **r + va, r + va** |

Table 5- A further adaptation of the payoff matrix from (Loebecke, et al., 1999). (r = basic value of knowledge, va = value added by monopolistic knowledge, f = value of feedback on published work, s = synergy produced by published works)

One aspect that the models fails to address is that in reality there are likely to be more than two players in the “game”. The more people there are willing to publish, the greater the multiplier of the value *r* thus the greater the benefit reaped when others publish. Similarly, we can also assume the more people there are in the academic community the value of *f* may be multiplied by some value with respect to the total population size, thus incentivising publication. However, more players may also increase the value of *va*. The more researchers there are in a community, the greater the chance that someone else will be able to put this information to use faster or in a more innovative manner than the initial researcher, placing them at a disadvantage. By the same token, withholding knowledge from a greater number of peers, a researcher may be able to gain greater influence and significance in their field by formally publishing the research in reputable journals.

Therefore the benefits that are derived from either publishing work informally online, or instead through a reputable journal, may swing depending on the overall balance of decisions taken by the various researchers. The more knowledge researchers publish however, the greater the increase in the multiplier of *r* and once a certain threshold has been reached, the risk of publishing is diminished as one player knows that other players are contributing to the shared pool of knowledge and so will also stand to benefit. This situation mirrors that of the public good dilemma, which whilst identifying the problem at hand, does not provide an easy solution. Rational players follow their dominant strategy of not publishing, thus the “deficient equilibrium” position is reached, at which no player has an incentive to publish as they will instantly be put at a disadvantage, a barrier Cabrera & Cabrera (2002) call a “social fence”.

So with the various possible outcomes mapped out, what decisions are players, or researchers, likely to choose? For a single turn we can assume that the dominant strategy will be chosen, and if the case of the deficient equilibrium has been reached, then the option of choosing not to publish will be chosen. However, the model does not incorporate an element of time and thus fails to show how decisions may change as time goes by and players have the opportunity to look back upon past changes. The most successful strategy in an on-going prisoners’ dilemma scenario is tit-for-tat: cooperate on the first move and for all following moves do whatever the other player did (Hutchens, 2007). Tit-for-tat has a number of advantages as a strategy for encouraging cooperation as it is able to foster a trustful environment for if everyone employs the strategy no one player will defect. However, should a player not act cooperatively, the strategy allows for retaliation, meaning it cannot be exploited by those wishing to seek maximum gain from a system.

Therefore whilst players may be motivated by self-interest, cooperation may still emerge providing players are able to recall previous interactions with other players and are able to see that their interactions will continue for the foreseeable future (Hutchens, 2007). The tit-for-tat model may not sustain cooperation on a long term basis, as the ability to retaliate may atrophy over time: “cooperation can flourish if the public-spirited majority can punish freeloaders” (Darwen & Yao, 2002), without this some individuals will exploit the situation; once again drawing parallels with the public good dilemma. This problem is typically solved by the presence of an authority figure, be it an individual or organisation, and thus the elegance and simplicity of the tit-for-tat system is eroded.

The application of the tit-for-tat strategy to the community of publishing chemists may break down slightly however. Typically one would not know who is accessing one’s work until after the fact and if so, there is little guarantee that they would provide anything of relevance in return. It is also impossible to restrict the scope of your interactions with others in the community, if you choose not to publish some data in retaliation of others not publishing, you are penalising everyone else by doing so, meaning the cooperation which is possible through tit-for-tat immediately breaks down. To take advantage of the collaborative nature of tit-for-tat it would be necessary to build a much smaller system that works on a local level and allows for closer, more informed and more controlled interactions between its members. This leads on to a much more organised, knowledge management approach to sharing data.

## Knowledge Management

Influencing individuals to make decisions that are for the greater good of the organisation in which they are members is not an uncommon problem in the realms of knowledge management and business, in which a lot of research has been conducted to help managers elicit an optimum level of productivity from their employees as members of a larger population. The benefits of knowledge management are well understood and desired by those who run organisations, better knowledge management can result in gaining a competitive advantage, increasing market effectiveness and improving product innovation (Cabrera & Cabrera, 2002); with this last advantage being of particular benefit to chemistry research in which researchers try and build upon each other’s work.

However, as has been already touched upon, knowledge management requests that individuals act for the benefit of a larger community rather than for themselves: knowledge is a public good. Public goods are non-excludable; everyone can benefit from them regardless as to whether or not they have made a contribution towards them and thus free-riders emerge, exploiting the good, which begins to deteriorate the more free-riders there are – the dominant strategy for “rational, self-interested individuals” (Olson, 1965). Naturally a solution needs to be devised to stop the free-rider problem or else no one can benefit from the good as there is no incentive to contribute towards it. In some cases, government provides the good, with mandatory contribution towards its provision through the means of taxes. However, in the instance of the an organisation wanting its employees to provide their knowledge so that others may benefit, enforced regulation may not be an option and thus other methods to encourage contribution may need to be implemented.

A number of other reasons why individuals do not want to contribute their experiences to a knowledge management system, beyond the reason that consumers may benefit more than the providers, have also been identified by research in this area. Sources of demotivation include the cost of the time and effort involved in contributing an item of knowledge, which is greater than that required in sharing knowledge informally with a colleague (Ardichvili, et al., 2003), a lack of understanding of the benefits of knowledge sharing both on a personal and organisational level and a feeling of intimidation that comes with making personal thoughts public, a fear of being belittled for making trivial contributions (Cabrera & Cabrera, 2002). Whilst these possible reasons for a lack of participation have been derived from a business environment, they may not be completely transferable to an academic setting. It is fair to assume however that the underlying psychology that influences the decisions made by potential knowledge management users may still help in the design of an approachable, useful knowledge management system for chemistry.

Research by (Cabrera & Cabrera, 2002) indicates that technical solutions tend to be the most effective way of encouraging people to create, disseminate and exploit organisational knowledge – providing a channel through which users can both contribute and received data – typically working better than solutions that involve altering organizational structure or implementing new human-resource policies. This implies that many employees who are asked to contribute to a knowledge management system identify the benefits to sharing knowledge, but the cost-benefit ratio is weighted in favour of costs. A well implemented technological solution should reduce the cost of carrying out the activity it is supposed to facilitate, thus in this context, making it more convenient to contribute and also pull out appropriate knowledge from the system. Research indicates (Cabrera & Cabrera, 2002) that employees cannot integrate knowledge sharing takes into their everyday activities – the provision of a technical solution that has this goal in mind may make a drastic change to the extent at which knowledge is provided, highlighting the importance of building this process into pre-existing ones such that the step is essentially seamless.

Whilst providing technological channels through which employees can conveniently share their knowledge, does help a firm leverage its knowledge potential, a medium for communication alone is not enough to reap the full benefits of knowledge management. The other problem to overcome is that of knowledge hoarding, a result of the “‘knowledge is power’ syndrome” (du Plessis, 2005), with hoarding typically taking place where there is competition. This is an issue which is particularly prominent in the academic environment in which there is a pressure to release better quality and a greater quantity of publications than one’s peers, and thus is a strong disincentive to the sharing of knowledge outside of a channel through which one would be recognised and rewarded (i.e. a journal article or similar publication). Thus it is necessary for organisations to attempt to encourage a transparent, open and friendly culture among those it wishes to share knowledge. One way in which this might be achieved might be to create communities of practice in which like-minded colleagues are brought together to share their work, a scheme which is essentially already in place in the academic context as researchers are members of research groups.

The open culture approach might struggle to counter a further problem faced by those wishing to implement a knowledge management strategy, the hoarding that occurs as a result of “functional silos” within an organisation (du Plessis, 2005). A more effective technique may be to offer more direct incentives for contributing to the knowledge management system, restructuring the manner in which employees are rewarded. Bartol and Srivastava (2002) found that providing extrinsic rewards, either monetary or not, did help to encourage the sharing of knowledge, although the provision of an extrinsic reward did erode intrinsic incentives (the pleasure derived from simply performing the task). Of particular note are the benefits of rewards provided at a team level, which were found to encourage “cooperation and coordination” and develop a “focus on group goals and performance” (Bartol & Srivastava, 2002). Suggestions of a change in the way in which the achievements of academics are recognised have been made before, with calls for universities and research institutes to “support an open data culture by: recognising data communication by their researchers as an important criterion for career progression and reward” (The Royal Society, 2012).

Incentives may be required to both “push” knowledge into the system, but also to “pull” knowledge, a process which may suffer from the “not-invented-here” syndrome (Katz & Allen, 1982), in which users do not wish to use knowledge sourced from elsewhere for fear of feeling inadequate or for not trusting its validity. Financial motivation tends not to work as successfully when encouraging the pulling of knowledge, if not managed carefully it can result in the hoarding of knowledge (Hauschild, et al., 2001). In an academic context, it may be sensible to encourage a supervisor to advise their supervisee to seek knowledge from further afield (especially when conducting interdisciplinary work), possibly giving a student a target number of papers to source from other disciplines.

As has already been noted, much of knowledge management research falls with the business domain, trying to ensure that a corporation makes the most effective use of its employers and that vital knowledge is not lost when an employee leaves the organisation. Whilst there are parallels between an academic community and a business one, other factors also need to be taken into account when considering a knowledge management system that could be used successfully in a research environment. The type of knowledge that needs to be handled may be differ greatly from that found in business – scientific knowledge is difficult to aggregate as it frequently changes and may require specialised expertise to be understood and needs to be codified to be stored, a complex task for cutting-edge research (Bos, et al., 2007). Typically researchers also have greater freedom to pursue high risk/high reward ideas than individuals in other professions and many scientific endeavours are cross-institutional, which in turn creates new knowledge management problems (Bos, et al., 2007).

## Collective Behaviour

Whilst game theory can be used to help predict what decisions individuals may make given a number of factors, and how we might engineer a situation to elicit the desired outcome, it is nevertheless restricted in its ability to model n-players where *n* is greater than 2. For example, it is difficult to measure how penalties and rewards may be distributed among players, when all the other players’ decisions also need to be accounted for, resulting in a very large number of possibilities needing to be modelled as more players are added (Yao & Darwen, 1995).

An alternative approach to looking at the behaviour of many inter-connected individuals is to consider the theories put forward on the subject of collective behaviour. A collective behaviour model can be used when actors have two alternatives (in this case to disseminate or not to disseminate) and the costs and benefits of the action depend on the number of actors who choose the alternative (Granovetter, 1978). Individuals in the system each have a threshold value: the number or proportion of others who must make one decision before a given actor makes a similar decision the point where the net benefits exceed the net costs for that particular actor. This model, as proposed by Granovetter (1978) highlights the importance of understanding not only individual norms and preferences, but also the aggregate actions of the population in question. Comparing the beliefs of the individual and the actual outcomes of the population often reveals paradoxical effects, as they fail to align with the intentions of the individuals who together are responsible for them. The model goes on to describe the various roles that are fulfilled by certain individuals in the population: “instigators”, those who will commit even if no else is doing so; and “conservatives”, who possess very high thresholds and may only commit when 80-90% of others do also. Instigators are already beginning to emerge in the topic of academic publishing and, more precisely, that of the problems posed by the existing publishing models, as shown by the actions of Tim Gowers, the instigator of the recent Elsevier boycott in protest of their rising journal subscription prices (The Cost of Knowledge - http://thecostofknowledge.com/) (Jha, 2012).

Granovetter (1978) identifies possible reasons for why the actions that take place may be as a result of the threshold values people possess rather than their “norms”. Norms are the actions which are carried out because of an individual’s beliefs and attitudes that affect their decisions should they be in a closed environment without any third party influences. Thus norms are only one “causal influence” on behaviour, with thresholds being the other, with the two influences typically being interrelated. For example, in the publishing context, we could assume that a chemist would be happy to publish their work in an open, accessible format and this belief would have the effect of lowering their threshold for publishing – however, despite wanting to openly publish, they may never do so if the threshold is not surpassed, because they are concerned about what impact this will have upon them in the context of the community at large. Factors that may influence one’s norms and thresholds include determinants such as social class, education and social position (Granovetter, 1978); in the case of chemistry these factors may translate to your career (early career researchers may be more protective of their data), peers and supervisors (supervisor may influence a student’s decision on whether to disseminate data or not) and occupation (a lab assistant may be less inclined to publish data if it adds a significant burden to their routine tasks).

Therefore, when advocating the benefits of publishing data more openly, it is important to identify those with the lowest thresholds and take advantage of the aggregate effect. This strategy is reflected in the Advocacy to Benefit from Changes report (Burns, et al., 2009), that suggests targeting local champions, who have little to lose by publishing data openly. Care must also be taken to manage the threshold values of individuals in a population such that an insurmountable divide cannot form between the groups, where the conservative threshold is so high it is impossible for a proportion of the threshold to reach it. Any system that would take advantage of collective behaviour theory would want to encourage users to not only lower their threshold values, but also be aware of when this value has been reached so that they may act as they see fit. This manipulation of thresholds and norms can only be achieved by understanding the determinants as described above and building systems, which as indicated by the paradoxical outcomes of collective action, may differ to what the user might initially specify as expectations and requirements.

# Preliminary Studies

Work has already been conducted to gather some initial thoughts and opinions that belong to the various stakeholders of the publishing process. The data generated thus far has been mostly qualitative, originating from questionnaire responses, interview observations and panel discussions. As such it has been approached with an interpretivist approach, attempting to elicit general comments and themes and thus build to some conclusions and directions for moving forwards from there. The results from these initial data sources are summarised below, along with interpretations as to what impact they may have in future.

## Questionnaire

The first study to be conducted was a questionnaire distributed to early career researchers at the University of Southampton’s chemistry department. The questionnaire was designed with three core topics: publishing, motivations for publishing and implementation of publishing systems. The response rate for the questionnaire was low, with approximately 7% response rate, a result which itself informs about the population sampled, indicating that the topic of dissemination of chemistry research is of little interest to chemists. Those that did respond however gave some useful insights, with some very much in support of more open, informal dissemination and others against new, disruptive means of communication.

It is clear from the survey that researchers place a lot of importance on the role of journals and the power that a successful article can do for their reputation in the field. Linking this back to ideas of disintermediation presented earlier, it is clear that whilst academics have the potential to disrupt the research market, they have little desire to compare to the consumers in other markets such as music, being content with the system currently in place. However this does not mean that a suitable replacement system would not necessarily be successful, just that the flaws in the current system are not deemed sufficient to be fixed by many. Concerns about new publishing models are typically placed with regards to its function in assuring good quality science is conducted and that poor work does not become too prominent and ensuring that researchers do not fall victim to intellectual property theft. The current medium is perceived to perform this function to a suitable level and any changes to the publishing model would be required to meet these needs as sufficiently, if not more so, to build confidence and become a mainstream approach.

Respondents also made it quite clear that they were keen to retain control over the distribution of their work and that it should not become too readily available too soon, preferring to capitalise on its monopolistic value until that has been sufficiently diminished. Conversely, respondents agreed that being able to access the records of other researchers would be useful and that with more work available to view, collaborations will become more likely. Hence most would like the provision of data and yet the value of keeping it to one’s self is too high for this to happen, thus all researchers stick to their dominant strategy, depriving the society as a whole from benefitting. This is a classic example of a public good dilemma and steps would be necessary to try and attain the socially optimum output. Use of the theories put forward by collective behaviour and knowledge research may be the solution to this, encouraging researchers to disseminate their data, but it is likely that a sustained change in attitudes will only be possible by changing the current reward mechanisms; introducing the possibility of adding new benefits to publishing and weakening the monopolistic value. Thus significant changes are likely to be mandated from higher authorities and not emerge from the community as a whole.

## Interview Observations

Notes were taken during a number of interview sessions to better understand how chemistry students and supervisors work alongside each other and how the community acts in practice on a local scale. The interviews that were observed were conducted by Jennifer Rutner, a representative of the company Ithaka (http://www.ithaka.org/) as part of research into the intersection of academic collaboration and the uses of technology involved in these processes. The range of interview participants lead to a number of thematic areas emerging on the topic of the application of technology to chemistry research and how it has changed chemistry practices, which are summarised below.

### Data Management

Data needs to be stored in a persistent and manageable manner, such that even data generated a decade ago can be called upon and used in a publication where necessary. It is important to consider that the ultimate users of data are frequently not those who created it and so metadata is vital to clarify the true origins and intentions of the data and its creator. Systems used to generate data, such as the crystallography service, contain severe bottlenecks, with use of such systems typically being managed manually by a few members of staff who keep their own records of who did what and when. This could lead to difficulties, should the staff leave, with the loss of much tacit knowledge on the workings of the system. It should be worth considering therefore that any knowledge management system that could be implemented on a local scale in a chemistry department, should perhaps not only be able to facilitate the sharing of knowledge between individuals, but also be compatible with larger, more autonomous systems.

### Student Collaboration

Students each have their own methods for storing material that they find works best for them. Information stored can be sourced from a range of places, including magazine articles, journal articles, software and video presentations. The extent to which students collaborate and share this information may be influenced by their supervisors. Some supervisors for example actively encourage sharing by using a shared network drive to which all students have access and can upload papers. Whilst method may be effective within a single research group, it is nevertheless quite informal and closed off to other sources and influences. There is also no formal method in place for managing data in this way.

Another aspect which affects the degree of collaboration between students is possibility of their work being confidential. If a student has an industry sponsor who wishes the work to be kept confidential, this may limit the extent to which they are allowed to discuss their work with fellow students. However underlying chemistry principles can still be discussed without revealing the true nature of the work. This separation of contexts may be a useful feature for a chemistry knowledge management tool, allowing work to be considered and shared whilst stripping away anything which might be irrelevant for a particular context. This concept also ties in to ideas concerning context-sensitive views when presenting work to different stakeholders. A knowledge management in chemistry may therefore have to deal with the multi-faceted nature of chemistry research; some of it is profitable, patentable and confidential, with a business interest; whereas research should also possess an academic interest, which may be just as important as the business perspective to other stakeholders.

### Student-Supervisor Relationship

Both supervisors and students have their own working styles, resulting in a wide range of possible permutations with regards to supervisory preferences, making it difficult, if not impossible, to design a system whereby one size fits all. The main medium through which supervisors communicate with students is through email, but this is not always regarded as the most preferable. Emails, whilst convenient and flexible, may not be the most organised and efficient way of communicating certain information. Students and supervisors have different purposes for communicating with each other: typically students will contact supervisors when they have questions with their current work or come across barriers that are slowing them down, or when they submit reports of their findings they wish their supervisors to read; supervisors typically contact their students when they come across interesting articles they think their students will find useful.

Using emails for all these purposes can result in what appears to be quite a random flow of communication, with little organisation. Another form of communication which is used, in particular with reference to students with an industry sponsor, is through conference calls or Skype calls. Once again communication in this manner is quite random and whilst calls are recorded, minutes are not always taken and so elements of meetings may be lost if recordings are not replayed.

Responses from the interviews indicate that a centralised approach to communication may be more useful for both students and supervisors. One research group already uses a system like this in the form of a shared network drive for the distribution of interesting articles, whilst another group uses DropBox (https://www.dropbox.com/) for a similar purpose. However this only allows for information sharing within a small group and relies on complementary communication methods, such as email. An alternative approach is to use blogs as a means for student-supervisor contact, which could be used for both the student to submit reports and queries to their supervisors and for supervisors to post comments pertaining to items of interest. Blogs would have several advantages over the aforementioned mediums, allowing for both discussions to flourish and to be archived; however they are not without their own difficulties as described by Wolf (2010).

### Academic Publishing

The final theme that emerged from the interview sessions was that of how the Web has affected everyday research practice. One academic interviewed stated that when looking for articles on a research topic their first place to look was Google Scholar (http://scholar.google.co.uk/), which if unsuccessful was followed by Web of Science (http://wok.mimas.ac.uk/). Should neither of these sites provide useful results the academic would then ask their colleagues for advice. Whilst this does provide a much more convenient method of finding articles, the interviewee highlighted that it removed the random element of stumbling across new ideas when searching the library for appropriate journal articles, a process that would be difficult to emulate in the online environment. This problem may be further compounded by the vast amounts of research articles that are published, with daily content alerts being ignored, and pictures used as a means for identifying interesting articles.

The interviewee also indicated that at times they came across relevant articles, but they are behind paywalls, removing any inclination to find them again. This in turn has led to open access methods of publishing being seen more favourably, albeit typically through emailing copies of papers per request rather than through a formal open access publishing method. Gold open access was seen as a poor use of funds, with the cost of publishing high and the resultant article having very little impact, with perceptions indicating that the money would be more wisely spent in labs – a clear indication that should funding bodies wish their research to be published via open access, some funds will need to be earmarked for the publication process. The interviewee highlighted the importance of understanding which open access journals carried the highest impact before undertaking a similar such open access publishing approach.

## Panel Discussion

“Evolution of Science: Open Science and the Future of Publishing” was a panel discussion held in Oxford, which aimed to tackle questions such as “Do we still need science journals in the Internet Age?”, “Is there something better than peer review?” and “How will scientific papers evolve?” (Benjamin, 2012). In attendance at the panel were: Robert Winston, a professor, medical doctor, television presenter and politician; Alicia Wise, representing the publishing company Elsevier; Robert Kiley, representing the Wellcome Trust, a charity that funds health research; Alison Mitchell from the Nature Publishing Group; Cameron Neylon, notable blogger on Open Research; Victor Henning, cofounder and CEO of Mendeley Ltd; and Tim Gowers, Cambridge mathematician, and instigator of the “Cost of Knowledge” petition to boycott publishing in Elsevier published journals. Each of the panellists were asked to give their opinions on the aforementioned topics, before fielding questions from the audience and thus a number of themes emerged amongst the responses.

### The Role of Publishers

The panellists tended to agree that the publishers do play an important, and perhaps more significantly, costly role and as such the money that is needed to support the roles they carry out needs to be sourced from somewhere, with Gold OA being seen as a popular route. The Welcomme Trust support the research funder pays model and suggests that for OA to become mainstream, funders will need to mandate Gold OA publishing and find the funds to make this possible. However they also stress the importance of lobbying researchers to recognise the issue, making a comparison to the recycling of plastics: the facilities need to be in place before plastic can be recycled, but only with sufficient promotion does widespread action take place. Conversely, the Nature Publishing Group believes readers rather than authors should pay for the costs of the value added by publishers as the cost is spread over a larger population, preferring to use Gold OA for publications of interest, but of a lesser standard than typical Nature articles.

The Nature Publishing Group stressed the expense and value of full time editors who are needed to put journals together, with Elsevier highlighting the importance of other publisher roles including facilitating discovery, dealing with ethical issue and conflict resolution (such as plagiarism cases) and their role in improving the efficiency and quality of the peer review process, a process which needs a lot of effort putting into it to get good results out of it. On the topic of peer review, whilst it was noted that the peer review process is not without flaws, Elsevier also noted that researchers think it works and trust it as the best method available, thus justifying investment put into improving the process. Victor Henning, representing Mendeley, also highlighted the roles played by publishers including: validation, curation, filtration and designation of content, all roles which are important for the free flow and efficient use of information.

Other panellists presented some alternative options to peer review and some of the roles that journal publishers fulfil. Tim Gowers highlighted the extent to which papers are now discussed and debated online in preprint repositories such as arXiv (http://arxiv.org/) and Victor Henning indicated that there was a growing demand for collaborative peer review services to be built into Mendeley.

### Future Publications

All panellists present recognised that traditional forms of publishing have become outmoded and the full affordances of the Web are not being taken advantage of. Nature stressed the need to make content as useful as possible, with regards to both how it is stored and accessed as well as from a linked data perspective, enabling research to become “more than the sum of its parts”. Elsevier too, is keen to change readers’ experiences, developing a number of apps to present journal articles, alongside research to develop new sorts of journal articles.

Cameron Neylon, highlighted the importance of technological developments, making a comparison between the invention of the printing press and the Web: with the printing press it is logical to centralise production, whereas the Web systems that work are those where content is open and are built upon a distributed architecture.

Another important issue raised was the “unit of discourse”. Tim Gowers highlighted how the manner in which work is disseminated may reflect how useful it can be: a conversation with a colleague may be too fast and informal, whereas a journal article may be too long in the making and provide a reader with too much information to reasonably respond too. A blog achieves a happy medium between the two allowing for detailed ideas to be expressed but also quick feedback and discussions to commence; although it was noted that the blog format is not without its flaws.

Discussion on the future of publications and how formats and methods may stray from the current norm was vague, with no concrete ideas discussed. However a number of common elements emerged throughout, including the need to make research more approachable and useful from both the perspectives of other researchers and the members of the public, as well as the need to embrace the power of the community and technology. Publication formats should allow for a greater number of interactions between academics (opening up a research question to the community was identified as being more powerful than searching through existing literature) and for greater use of semantic web tools and techniques to develop new forms of scientific enquiry.

# Future Plans

It is clear that the Web has had a disruptive influence on academic publishing and dissemination, yet ultimately this impact has been restricted to emulating many of the previous practices in an electronic form. Many of the new processes sit alongside pre-existing ones and simply serving the same purpose. PDFs are pushed around between users in much the same manner as paper hardcopies and repositories bundle related articles together like journals. The impact on the processes of research and communication with fellow scholars appear to be limited, with a variety of tools being used to varying degrees of success. Developments such as OA have taken a hold and are set to become increasingly mainstream as approaches such as Gold OA become mandated (Finch, 2012). However, as highlighted by findings in both the literature and the opinions of those who work in this field; for significant changes to be made, we need to move beyond issues of openness and strive for an “intelligent openness: data must be accessible and readily located; they must be intelligible to those who wish to scrutinise them; data must be assessable so that judgments can be made about their reliability and the competence of those who created them; and they must be usable by others” (The Royal Society, 2012).

Naturally, looking for new ways in which the Web can be used to enhance the research process presents endless possibilities. However with the focus on research dissemination, there are two main areas to focus on: how scholars communicate with one another; and what they communicate with one other. Research thus far has also indicated that the development of new tools and technologies alone is not sufficient to get the attention and use of communities at large – OA alone provides an example of this. The knowledge model framework described above, and similar models found within the literature show that it is the researchers who lie at the heart of the processes in question - they produce the content, they verify it and they consume it - and thus it is with their demands and attitudes in mind that will dictate the design of new developments.

Therefore future work will go on to look at the development of some prototype systems that may assist researchers carrying out their work, but whilst also exploiting the power of the Web to make research outputs more openly available to the appropriate stakeholders at the appropriate times. Despite the numerous stakeholders in the overall research process, it is the researchers who lie at its heart, providing a market for research dissemination from both the production and consumption perspectives; and it is the researchers whom this research is aimed at helping the most, allowing them to spend more time in the lab than at their computers.

## Personal Publishing

For any future system to be successful many of the roles currently carried out by journals would have to be emulated so as to provide a degree of familiarity and trust in a system and thus encourage uptake. However as has been identified, some of the functions provided by journals are flawed, with the Web possibly providing mechanisms for bettering these functions. Any changes to metrics such as impact factors, or processes such as peer review, would have to be demonstrably better or replicate existing processed albeit with value adding features. Perhaps most importantly however, it is clear from previous studies and interviews with chemistry researchers, that any system would have to fit seamlessly into their day-to-day processes, working with very little input by users to provide valuable products.

A future system might therefore integrate with electronic lab notebooks to automatically produce, publishable “chunks” of data or knowledge, which in turn may be combined in a modular fashion to create larger publications which fit more conventional ideas as to what a publication should look like. As we have already seen getting the right sized “unit of discourse” is critical to success, with blogs typically being an appropriate size to generate debate. Thus a system which can take electronic lab notebook pages and automatically produce an online, publicly readable account of the work going on may be very powerful in leveraging a small portion of the community. However, two criteria must be met for this to be successful: firstly authors must receive credit for publishing work in this manner and it should when fellow researchers make use of it, this should be recognised; secondly the author must have complete control over who has access to it, with permissions being changeable when appropriate to allay fears with regards to IP theft.

## Interactions and Negotiated Openness

To this end, a system which displays to users a network of their peers and research stakeholders will be investigated to examine effectiveness in controlling access to material. By providing a greater level of control over the openness of research, it is hoped that more research will become more widely available than the more binary approach of making results available or not, and offer more flexibility than publishing in a journal or online repository; and by providing users with a network of those who are interested in a piece of work, it should be easier for them to specify who gets to see what. By having a greater understanding of who you know, and in turn who they know and the connections between this myriad of individuals, it is hoped those who wish to contribute a piece of information into the network, can use this knowledge to enhance the effectiveness of the dissemination, ensuring those who interested receive information with an appropriate degree of detail and possibly uncovering new opportunities for interactions and discussions. Chunks of knowledge may be able to “pulse” through a network, revealing less information as it travels, but ensuring those who care receive some notification and opportunity to show further interest. Collective behaviour shows us that for individuals to commit to certain actions, a certain proportion of the community must already have acted in a similar manner – by mapping out the relevant population in a network, not only can we keep the risks of acting down by controlling the size and extent of the network, but it would also be possible to show how many interactions are already going on, encouraging those with higher threshold values to participate. These ideas are backed up by a JISC report: *TechWatch: Preparing for Data-driven Infrastructure* (Hammond, 2012) that prioritises the benefits of sharing data internally before opening it up to external entities.

Game theory tells us that the best way of encouraging a collaborative behaviour is to use a tit-for-tat strategy in which interactions are transparent and meaningful – information which should be visible in this local network of connections. People in the network should be able to identify who has contributed what and thus may feel encouraged to reciprocate actions in the future, whilst also providing the benefit of highlighting any free riders who may be reaping the benefits of others’ interactions. Whilst an unfortunate state of affairs, this may allow contributors to cut out such members when negotiating the openness of their work, should they see this as a necessary course of action, and thus may enable contributions and collaborations to continue on into the future. A history of interactions may also form, allowing users to generate reputations, assisting researchers in not only knowing who has provided information in the past, but also allowing them to be recognised in some manner for their contributions.

## Canvas

With a focus on the dissemination of research outputs, not only do we need to consider how research is spread throughout the community, but also how it is communicated. We have already established a number of criteria; namely that it should be easier to cite contributions beyond those made in journals or at conferences, and that authors should have a greater degree of control over the accessibility of their work in a bid to make more information available to those who might find it useful. However to enable this functionality we have to consider how this will be presented to users; with one possible option being a virtual canvas of a similar style to presentations made using the website Prezi (http://prezi.com/).

A canvas interface would allow the various items that combined form a complete research output to be presented side by side enabling readers to both see how they are used together, but also how they stand on their own merits. The use of physical space to segregate items of work could be combined with depth to illustrate different levels of detail: users zoom in for more detail and out for a more abstract view. Zooming in sufficiently far into individual items might even present the reader with a page from an electronic lab notebook where appropriate. This use of space may allow for more precise citations (as predetermined regions on the canvas could be cited) but also for a greater control over accessibility as different regions within the canvas are marked for different readers. The interface may also lend itself towards collaborative work, allowing multiple users to work on different areas of the overall document, without interfering with each other’s contributions. Regions of the canvas could then be filtered to display different author’s work and the contributions may be cited either individually or collectively.

A canvas interface would not be without its problems however. Creating such documents that flow and present research in an easy to follow and understand manner may be trickier than normal formats. However, this problem may be countered if much of the document can generate itself as researchers work with their electronic lab notebooks. This would also allow us to take advantage of the desire to produce publishable chunks or research outcomes which can be modularly placed together to create more sophisticated publications.

## Further Avenues of Research

Whilst technological solutions are seen as more successful when encouraging new approaches to knowledge management, it is important to consider that the current methods of research dissemination have a long history and thus are deeply embedded within research institutions and communities. The technological innovations proposed above would not only present researchers with a new system to become familiar with, but also present cultural changes which would need to be modelled and analysed to make the transition more palatable to potential users.

The *Studies on Subject-Specific Requirements for Open Access Infrastructure* report also makes an important point which should be considered when developing any Web based system to help facilitate the flow of knowledge within and beyond a discipline’s boundaries: “infrastructure is an opponent to diversity…a collection of rigid conditions or constraints: it is an inherent property and explicit objective of infrastructure to make research uniform” (Meier zu Verl & Horstmann, 2011). The reports response to this problem is to ensure that any infrastructure operates in an “open mode” promising the “permeability of research resources” and supporting the diversity of research practices”, comments which should be considered when designing the proposed system described above.

Naturally, any proposed system that would be implemented would need to be approved by the chemists themselves – convincing researchers to try a new system, which in turn does not work for them, may very well set back attempts to change the publishing process for the better. To this end demos of prototype pieces of software may be shown to focus groups to attain qualitative feedback alongside interviews or questionnaires to gain a greater understanding of the environment in which any system would hope to seamlessly integrate with.

# Bibliography

Ardichvili, A., Page, V. & Wentling, T., 2003. Motivation and barriers to participation in virtual knowledge-sharing communities of practice. *Journal of Knowledge Management,* 7(1), pp. 64-77.

Bartol, K. M. & Srivastava, A., 2002. Encouraging Knowledge Sharing: The Role of Organizational Reward Systems. *Journal of Leadership and Organization Studies,* 9(1), pp. 64-76.

BBC, 2012. *BBC News - Today - Call to share research papers.* [Online]   
Available at: http://news.bbc.co.uk/today/hi/today/newsid\_9712000/9712028.stm  
[Accessed 14 May 2012].

Benjamin, S., 2012. *The Scientific Evolution: Open Science and the Future of Publishing.* [Online]   
Available at: http://duraspace.org/scientific-evolution-open-science-and-future-publishing  
[Accessed 26 July 2012].

Björk, B.-C.et al., 2010. Open Access to the Scientific Journal Literature: Situation 2009. *PLoS One,* 5(6), pp. 1-9.

Bollen, J., Rodriguez, M. A. & Van de Sompel, H., 2006. Journal Status. *Scientometrics,* 69(3), pp. 669-687.

Borkum, M., Lagoze, C., Frey, J. & Coles, S., 2010. *A Semantic eScience Platform for Chemistry.* Brisbane, Queensland Australia, s.n., pp. 316-323.

Bos, N. et al., 2007. From Shared Databases to Communities of Practice: A Taxonomy of Collaboratories. *Journal of Computer-Mediated Communication,* Volume 12, pp. 652-672.

Buckley, P. J. & Casson, M., 1976. *The Future of the Multinational Enterprise.* New York: Holmes & Meier.

Burns, L., Dennis, N., Kahn, D. & Town, B., 2009. *Advocacy to Benefit from Changes: Scholarly Communications Discipline-Based Advocacy,* s.l.: Publishing Directions.

Cabrera, A. & Cabrera, E. F., 2002. Knowledge-Sharing Dilemmas. *Organization Studies,* 23(5), pp. 687-710.

Colquhoun, D., 2011. *Publish-or-perish: Peer review and the corruption of science.* [Online]   
Available at: http://www.guardian.co.uk/science/2011/sep/05/publish-perish-peer-review-science  
[Accessed 24 July 2012].

Dallmeier-Tiessen, S., Darby, R., Goerner, B. & Hyppoelae, J., 2011. *Highlights from the SOAP project survey. What Scientists Think about Open Access Publishing.* [Online]   
Available at: http://arxiv.org/abs/1101.5260v2  
[Accessed 16 May 2012].

Darwen, P. J. & Yao, X., 2002. Co-Evolution in Iterated Prisoner's Dilemma with Intermediate Levels of Cooperation: Application to Missile Defense. *International Journal of Computational Intelligence and Applications,* 2(1), pp. 83-107.

De Roure, D. & Frey, J., 2007. *Three Perspectives on Collaborative Knowledge Acquisition in e-Science.* Hyderabad, India, s.n.

du Plessis, M., 2005. Drivers of knowledge management in the corporate environment. *International Journal of Information Management,* 25(3), pp. 193-202.

Dunleavy, P., 2010. *The Future of Joined-Up Public Services,* London: 2020 Public Services Trust.

Egghe, L., 2010. The Hirsch index and related impact measures. *Annual Review of Information Science and Technology,* 44(1), pp. 65-114.

Eysenback, G., 2007. From Intermediation to Disintermediation and Apomediation: New Models for Consumers to Access and Assess the Credibility of Health Information in the Age of Web2.0. In: K. A. Kuhn, J. R. Warren & T. Leong, eds. *Medinfo 2007: Proceedings of the 12th World Congress on Health (Medical) Informatics; Building Sustainable Health Systems.* Amsterdam: IOS Press, pp. 162-166.

Finch, J., 2012. *Accessibility, sustainability, excellence: how to expand access to research publications,* s.l.: s.n.

Goble, C. A. et al., 2010. myExperiment: a repository and social networking for the sharing of bioinformatics workflows. *Nucleic Acids Research,* Volume 38, pp. 677-682.

Google, 2012. *Google Scholar Metrics.* [Online]   
Available at: http://scholar.google.com/intl/en/scholar/metrics.html  
[Accessed 23 July 2012].

Granovetter, M., 1978. Threshold Models of Collective Behavior. *American Journal of Sociology,* 83(6), pp. 1420-1443.

Hammond, M., 2012. *TechWatch: Preparing for Data-driven Infrastructure,* s.l.: JISC Observatory.

Hansen Montgomery, C. & King, D. W., 2002. Comparing Library and User Related Costs of Print and Electronic Journal Collections. *D-Lib Magazine,* 8(10).

Harvard University, 2012. *Faculty Advisory Council Memorandum on Journal Pricing.* [Online]   
Available at: http://isites.harvard.edu/icb/icb.do?keyword=k77982&tabgroupid=icb.tabgroup143448  
[Accessed 7 June 2012].

Hauschild, S., Licht, T. & Steain, W., 2001. Creating a knowledge culture. *McKinsey Quarterly,* pp. 74-81.

Henderson, M., 2010. Problems with peer review. *BMJ,* Volume 340, pp. 738-740.

Hey, T. & Trefethen, A., 2003. The Data Deluge: An e-Science Perspective. In: F. Berman, G. Fox & T. Hey, eds. *Grid Computing: Making the Global Infrastructure a Reality.* Chichester, UK: John Wiley & Sons.

Hirsch, J. E., 2005. An index to quantify an individual's scientific research output. *Proceedings of the National Academy of Sciences of the United States of America,* 102(46), pp. 16569-16572.

Houston, B., 2010. The future of investigative journalism. *Daedalus,* 139(2), pp. 45-56.

Hutchens, D., 2007. *Trust and the Prisoners Dilemma.* [Online]   
Available at: http://www.davidhutchens.com/Biz%20Writing/articles/trustandthepriso.html  
[Accessed 20 February 2012].

Jha, A., 2012. *Academic spring: how an angry maths blog sparked a scientific revolution.* [Online]   
Available at: http://www.guardian.co.uk/science/2012/apr/09/frustrated-blogpost-boycott-scientific-journals  
[Accessed 9 April 2012].

Katz, R. & Allen, T. J., 1982. Investigating the not invented here (NIH) syndrome: a look at the performance, tenure, and communication patterns of 50 R&D porject groups. *R&D Management,* 12(1), pp. 7-19.

Kittur, A., Chi, E. H. & Bongwon, S., 2008. *Crowdsourcing user studies with Mechanical Turk.* New York, s.n.

Kronick, D., 1991. *A Scientific and Technical Periodicals.* Metuchen: Scarecrow Press.

Kwak, H., Lee, C., Park, H. & Moon, S., 2010. *What is Twitter, a Social Network or a News Media?.* Raleigh, NC, 19th international conference on World wide web.

Landesman, C., 1997. *An Introduction to Epistemology.* Cambridge: Blackwell Publishers Inc.

Loebecke, C., Van Fenema, P. C. & Powell, P., 1999. Co-Opetition and Knowledge Transfer. *ACM SIGMIS Database,* 30(2), pp. 14-25.

Mafulli, 1995. More on citation analysis. *Nature,* Volume 378, p. 760.

Mahoney, M. J., 1977. Publication prejudices: An experimental study of confirmatory bias in the peer review system. *Cognitive Therapy and Research,* 1(2), pp. 161-175.

Mandavilli, A., 2011. Trial By Twitter. *Nature,* Volume 469, pp. 286-287.

Meier zu Verl, C. & Horstmann, W., 2011. *Studies on Subject-Specific Requirements for Open Access Infrastructure,* Bielefeld: Universitätsbibliothek Bielefeld.

Miller, T. R. et al., 2008. Epistemological Pluralism: Reorganizing Interdisciplinary Research. *Ecology and Society,* 13(2).

Mollett, A., Moran, D. & Dunleavy, P., 2011. *Using Twitter in university research, teaching and impact activities,* London: LSE Public Policy Group.

Morris, S., 2006. *Getting Started in Electronic Journal Publishing,* s.l.: International Network for the Availability of Scientific Publications.

Mortensen, T. & Walker, J., 2002. Blogging thoughts: personal publication as an online research tool. In: A. Morrison, ed. *Research ICTs in Context.* Oslo: University of Oslo, pp. 249-279.

Olson, M., 1965. *The Logic of Collective Action.* Cambridge, MA: Harvard University Press.

Page, L., Brin, S., Motwani, R. & Winograd, T., 1998. *The PageRank Citation Ranking: Bringing Order to the Web,* s.l.: Stanford InfoLab.

Peters, D. P. & Ceci, S. J., 1982. Peer-review practices of psychological journals: The fate of published articles, submitted again. *Behavioral and Brain Sciences,* 5(2), pp. 187-195.

Quah, D., 2003. Digital Goods and the New Economy. *CEPR Discussion Paper.*

Reed Elsevier, 2012. *Reed Elsevier 2011 Results.* [Online]   
Available at: http://www.reedelsevier.com/mediacentre/pressreleases/2012/Pages/reed-elsevier-2011-results-announcement.aspx  
[Accessed 11 June 2012].

Repko, A., 2008. *Interdisciplinary Research.* 1st ed. Los Angeles: Sage Publications Inc.

Research Information Network, 2009. *Patterns of information use and exchange: case studies of researchers in the life sciences,* s.l.: British Library.

Research Information Network, 2009. *Patterns of information use and exchange: case studies of researchers in the life siences,* s.l.: s.n.

Research Information Network, 2011. *Information handling in collabortive reserach: an exploration of five case studies,* s.l.: s.n.

Shirky, C., 2009. *Newspapers and Thinking the Unthinkable.* [Online]   
Available at: http://www.shirky.com/weblog/2009/03/newspapers-and-thinking-the-unthinkable/  
[Accessed 1 June 2012].

Smith, M. D., Bailey, J. & Brynjolfsson, E., 1999. Understanding Digital Markets. In: E. Brynjolfsson & B. Kahin, eds. *Understanding the Digital Economy.* Cambridge, MA: MIT Press, pp. 99-136.

Smith, R., 2010. Classical peer review: an empty gun. *Breast Cancer Research,* 12(4).

Smith, R., 2010. Classical peer review: an empty gun. *Breast Cancer Research,* 12(4).

Spohr, D. & Cimiano, P., 2011. Information and Communication Technology. In: C. Meier zu Verl & W. Horstmann, eds. *Studies on Subject-Specific Requirements for Open Access Infrastructure.* Bielefeld: Universitätsbibliothek, pp. 69-124.

Swan, A. & Sheridan, B., 2005. *Open access self-archiving: An author study,* s.l.: s.n.

The Royal Society, 2012. *Science as an open enterprise,* London: The Royal Society.

The Royal Society, 2012. *Science as an open enterprise,* London: The Royal Society.

The Wellcome Trust, 2004. *Cost and business models in scientific research publishing,* Cambridge, UK: SQW Limited.

Thomson Reuters, 1994. *The Thomson Reuters Impact Factor.* [Online]   
Available at: http://thomsonreuters.com/products\_services/science/free/essays/impact\_factor/  
[Accessed 11 June 2012].

Vinkler, 1986. Evaluation of some methods for the relative assessment of scientific publications. *Scientometrics,* Volume 10, pp. 157-177.

Ware, M., 2008. *Peer review: benefits, perceptions and alternatives,* London: Publishing Research Consortium.

Weller, M., 2011. *The Digital Scholar.* London: Bloomsbury Academic.

Wolf, K., 2010. Bridging the distance: the use of blogs as reflective learning tools for placement students. *Higher Education Research & Development,* 29(5), pp. 589-602.

Woodside, A. G., 2009. Journal and author impact metrics: An editorial. *Journal of Business Research,* Volume 62, pp. 1-4.

Yao, X. & Darwen, P. J., 1995. An experimental study of N-Person Iterated Prisoner's Dilemma games. *Progress in Evolutionary Computation,* Volume 956, pp. 90-108.

# Appendix B – AltAC Executive Summary

The following executive summary was written shortly after the submission of the AltOA: Dissemination Through Disintermediation paper that was presented at the WebSci 2013 conference and aims to present many of the findings found within that paper in a more concise manner. This represents the latest draft of the summary which attempts to outline the ideas behind disintermediation of academic publishing, the benefits it can bring to numerous stakeholders and the potential role of open data in scholarly discourse, under the collective title of “AltAC”, “Alternative Academic Communications”.

Executive Summary – AltAC (Alternative Academic Communications)

1. **Introduction**
   1. The research sector has failed to fully exploit the affordances of the Web, with much research and related metadata remaining inaccessible to relevant stakeholders.
   2. Research involves numerous stakeholders: disciplines, researchers, funding councils, publishers, institutions and metadata aggregators, third party data providers and various administrators.
   3. The researcher plays a fundamental role that underpins all other interactions, responsible for producing, reviewing and building upon research outputs.
   4. To affect a positive, disruptive influence using the Web, the roles and desires of all stakeholders need to be considered.
2. **Disintermediation**
   1. Throughout the research process a number of interactions are conducted between stakeholders via intermediaries. For example journal publishers organise the dissemination of articles and the peer review processes among researchers.
   2. Disintermediation, the removal of intermediaries is facilitated by the Web, allowing many stakeholders to communicate directly and openly with one another, removing many of the inhibiting effects imposed by certain intermediaries.
   3. In a disintermediated environment, each researcher would publish their own research activities online for other stakeholders to see, whilst retaining ownership and intellectual property rights.
      1. Research activities may include (as examples): publications, publishing datasets, grant proposals, peer reviewing and teaching.
   4. The benefits of disintermediation fall into three categories: research facilitation, creating opportunities and introducing efficiencies.
   5. The extent of disintermediation should be appropriate for each context and where intermediaries add value this value should be retained.
   6. For example, the role of journals should not be fully disintermediated, with high impact journals continuing to be used to highlight important research outputs.
3. **Open Data**
   1. Open data provides a mechanism to achieve the benefits of disintermediation.
   2. Research activities can be published in a timely and discoverable fashion, allowing researchers to assert ownership over their research whilst also making it easier to build upon the work of others, improving the overall efficiency of the research process.
   3. Researchers can cultivate an online presence independent of their journal publications, creating a profile which accurately reflects their contributions to the field.
   4. Publishing research data openly allows information to be presented to other stakeholders in a context-sensitive manner.
   5. The University of Southampton is in a unique position to build upon its existing Open Data Service, expanding this to include data on research activities, to ultimately create the foundations of a powerful, researcher centric social machine.
   6. Example use cases of how disintermediation can improve research include:
      1. Automatically generate a professional development report at the click of a button, instantly harvesting a researcher’s output from their open data profile.
      2. Keep up to date on a research area with an automatic feed that alerts a researcher to all developments in a topic within the university.
      3. Establish collaborations with other universities or research groups through quick and easy sharing of data and knowledge, minimising administrative overheads.
      4. Trace the provenance of research results, revealing more information than can be expressed in a journal article to more accurately reflect the work conducted.