David J. Brown
Richard Boulderstone

The Impact of
Electronic Publishing

The Future for Publishers and Librarians

K · G · Saur München 2008
Contents

Abstract xv

Acknowledgements xix

Chapter 1 Background 1
  1.1 Introduction ................................ 1
  1.2 The Book .................................. 1
  1.3 Purpose of the Book ............................ 2
  1.4 ‘Change management’ .......................... 4
  1.5 Target Audiences ............................ 6
  1.6 Definitions ................................ 6
  1.7 Objectives .................................. 7
  1.8 Approach adopted ............................. 7

Phase One

Chapter 2 Industry Evolution 11
  2.1 Tragedy of the Commons......................... 11
  2.2 Frustration Gap ................................ 12
  2.3 Publisher and Library dissentions .............. 13
  2.4 Big Deals .................................. 15
  2.5 The Tipping Point ............................ 16
  2.6 Open Access ................................ 17
  2.7 The Long Tail ................................ 19
  2.8 Disenfranchised Researchers ................. 20
  2.9 Knowledge Workers and the Thinkforce ....... 23
  2.10 Emergence of Search Engines ................. 24
  2.11 Something is Good Enough .................. 25
  2.12 The new market for research material ....... 26
  2.13 Overall Trends ............................. 26

Chapter 3 End User Behaviour 29
  3.1 Change in User Behaviour ..................... 29
  3.2 Who are the users? ......................... 30
  3.3 Typology of Users .......................... 31
  3.4 Information Overload ....................... 32
  3.5 Research Studies ........................... 33
### Contents

3.5.1 Industry wide Studies ........................................... 33  
Tenopir/King research ........................................... 33  
Collection development ........................................ 34  
3.5.2 Library sourced initiatives ................................. 35  
The eJUS report on E-Journal Users ............................ 35  
Faculty Attitudes at Univ California ............................ 37  
3.5.3 Publisher commissioned studies ............................ 37  
CIBER Studies .................................................. 37  
Elsevier/Mabe research ......................................... 39  
3.6 Author versus Reader .......................................... 43  
3.7 Digital Natives and the Millennium generation .......... 43  
3.8 Forecasts ..................................................... 44  
3.8.1 The Outsell View ......................................... 44  
3.9 User perceptions of Value  .................................. 45

### Chapter 4  Measuring the Value of Information 47
4.1 Background ..................................................... 47  
4.1.1 Peer Evaluation ........................................... 47  
4.1.2 Citation Analysis .......................................... 47  
4.1.3 Document Downloads ............................. 49  
COUNTER ..................................................... 50  
SUSHI .......................................................... 51  
Effect of Robots ............................................... 52  
Case Study – The Mesur Project .............................. 52  
4.1.4 Focus Groups and investigating individual usage patterns 53  
4.1.5 Document Delivery statistics ............................. 54  
4.1.6 Questionnaires ........................................... 55  
4.1.7 Triangulation .............................................. 55  
4.1.8 Scientometrics ............................................ 55  
4.2 Research Assessment Exercises .............................. 56  
4.2.1 The United Kingdom’s RAE ............................... 56  
4.2.2 Criticisms of UK’s RAE ................................. 57  
4.2.3 UUK report looks at the use of bibliometrics .......... 58  
4.2.4 Australia’s research assessment exercise (RQF) .......... 60  
4.3 The Future of Bibliometrics .................................. 60

### Phase Two

### Chapter 5  Electronic Information Industry Structure 65
5.1 How much Information? ................................. 65  
5.2 The Information Industry .................................. 66  
5.3 Corporate Size ............................................... 66  
5.4 The Scientific, Technical and Medical Information sector 67  
5.5 Challenges facing the information industry ............... 70

### Chapter 6  The Key Players 73
6.1 Industry Overview ......................................... 73  
6.1.1 Overall Scholarly Trends ............................... 73  
6.2 Structure of the Journal Publishing System .............. 73
<table>
<thead>
<tr>
<th>Chapter 6</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Market Estimates</td>
<td>75</td>
</tr>
<tr>
<td>6.3.1 STM Publishers</td>
<td>76</td>
</tr>
<tr>
<td>Key Stakeholders</td>
<td>78</td>
</tr>
<tr>
<td>6.4.1 Publishers and Information Providers</td>
<td>78</td>
</tr>
<tr>
<td>Leading Publishers</td>
<td>78</td>
</tr>
<tr>
<td>Learned Society Publishers</td>
<td>78</td>
</tr>
<tr>
<td>University Presses</td>
<td>79</td>
</tr>
<tr>
<td>Life Cycle of scholarly communication</td>
<td>80</td>
</tr>
<tr>
<td>The Future of the Big Deal</td>
<td>82</td>
</tr>
<tr>
<td>Trend towards Open Access</td>
<td>82</td>
</tr>
<tr>
<td>Versioning</td>
<td>83</td>
</tr>
<tr>
<td>Other Challenges facing Publishing</td>
<td>86</td>
</tr>
<tr>
<td>Hybrid Journals and document delivery</td>
<td>86</td>
</tr>
<tr>
<td>Refereeing</td>
<td>87</td>
</tr>
<tr>
<td>Peer Review in scholarly journals</td>
<td>88</td>
</tr>
<tr>
<td>Alternative review procedures</td>
<td>89</td>
</tr>
<tr>
<td>Publishers and the ‘Valley of Death’</td>
<td>89</td>
</tr>
<tr>
<td>The Prisoner's Dilemma</td>
<td>90</td>
</tr>
<tr>
<td>Future of Publishing</td>
<td>91</td>
</tr>
<tr>
<td>Research Libraries</td>
<td>92</td>
</tr>
<tr>
<td>ARL statistics</td>
<td>93</td>
</tr>
<tr>
<td>UK university library expenditure</td>
<td>93</td>
</tr>
<tr>
<td>Librarian relationship to their customers (users)</td>
<td>94</td>
</tr>
<tr>
<td>The European Digital Libraries</td>
<td>95</td>
</tr>
<tr>
<td>The Future of the Librarian</td>
<td>96</td>
</tr>
<tr>
<td>Understanding the new user</td>
<td>98</td>
</tr>
<tr>
<td>Other Stakeholders</td>
<td>98</td>
</tr>
<tr>
<td>Collaboratories</td>
<td>98</td>
</tr>
<tr>
<td>Funding Agencies</td>
<td>99</td>
</tr>
<tr>
<td>Government involvement</td>
<td>100</td>
</tr>
<tr>
<td>Case Study</td>
<td>100</td>
</tr>
<tr>
<td>Joint Information Services Committee</td>
<td>100</td>
</tr>
<tr>
<td>National Priorities on toll free or toll paid</td>
<td>101</td>
</tr>
<tr>
<td>Emerging Competition</td>
<td>102</td>
</tr>
</tbody>
</table>

Chapter 7 Publication Formats | 103 |
<p>| Journals and e-Journals | 103 |
| 7.1.1 Multi author articles | 106 |
| The evolution of the electronic journal | 107 |
| Article Readership | 109 |
| Concerns about Journals | 110 |
| Electronic journal use | 110 |
| Purpose of Reading | 111 |
| The Value of reading Journals | 112 |
| e-Journals in industry | 113 |
| Future of the Journal | 114 |</p>
<table>
<thead>
<tr>
<th>Chapter</th>
<th>Title</th>
<th>Pages</th>
</tr>
</thead>
<tbody>
<tr>
<td>7.2</td>
<td>Books and e-Books</td>
<td>115</td>
</tr>
<tr>
<td>7.2.1</td>
<td>The e-Book phenomenon</td>
<td>115</td>
</tr>
<tr>
<td></td>
<td>Ebrary results</td>
<td>117</td>
</tr>
<tr>
<td></td>
<td>Oxford Scholarship Online</td>
<td>117</td>
</tr>
<tr>
<td></td>
<td>e-Books on other platforms</td>
<td>118</td>
</tr>
<tr>
<td></td>
<td>A Strategy for Book Digitisation</td>
<td>118</td>
</tr>
<tr>
<td>7.3</td>
<td>Document Delivery</td>
<td>120</td>
</tr>
<tr>
<td>7.3.1</td>
<td>The market for article supply</td>
<td>120</td>
</tr>
<tr>
<td></td>
<td>InterLibrary Loans</td>
<td>122</td>
</tr>
<tr>
<td></td>
<td>Subito analysis of document delivery</td>
<td>124</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Title</th>
<th>Pages</th>
</tr>
</thead>
<tbody>
<tr>
<td>8.1</td>
<td>Legal Initiatives</td>
<td>127</td>
</tr>
<tr>
<td>8.1.1</td>
<td>Creative Commons</td>
<td>127</td>
</tr>
<tr>
<td>8.1.2</td>
<td>Science Commons</td>
<td>127</td>
</tr>
<tr>
<td>8.1.3</td>
<td>JISC and SURF’s Licence to Publish</td>
<td>128</td>
</tr>
<tr>
<td>8.1.4</td>
<td>Future of Copyright</td>
<td>129</td>
</tr>
<tr>
<td></td>
<td>Orphan Works</td>
<td>129</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Title</th>
<th>Pages</th>
</tr>
</thead>
<tbody>
<tr>
<td>9.1</td>
<td>Globalisation of Research</td>
<td>131</td>
</tr>
<tr>
<td>9.2</td>
<td>Movement of global funds for research</td>
<td>132</td>
</tr>
<tr>
<td>9.2.1</td>
<td>Regional variations</td>
<td>132</td>
</tr>
<tr>
<td></td>
<td>Industrial R&amp;D</td>
<td>133</td>
</tr>
<tr>
<td></td>
<td>Academic R&amp;D</td>
<td>133</td>
</tr>
<tr>
<td>9.3</td>
<td>Implications on scholarly publishing</td>
<td>134</td>
</tr>
<tr>
<td>9.3.1</td>
<td>Worldwide Trends in Article Output</td>
<td>135</td>
</tr>
<tr>
<td>9.3.2</td>
<td>Trends in Three Major Publishing Regions</td>
<td>135</td>
</tr>
<tr>
<td></td>
<td>Europe</td>
<td>136</td>
</tr>
<tr>
<td></td>
<td>United Kingdom</td>
<td>136</td>
</tr>
<tr>
<td></td>
<td>Asia</td>
<td>136</td>
</tr>
<tr>
<td></td>
<td>China</td>
<td>137</td>
</tr>
<tr>
<td></td>
<td>Australia</td>
<td>138</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Title</th>
<th>Pages</th>
</tr>
</thead>
<tbody>
<tr>
<td>10.1</td>
<td>Sources of Funds for Research</td>
<td>141</td>
</tr>
<tr>
<td>10.2</td>
<td>Research Trends</td>
<td>142</td>
</tr>
<tr>
<td>10.3</td>
<td>The changing R&amp;D process in large corporations</td>
<td>143</td>
</tr>
<tr>
<td>10.3.1</td>
<td>Social Collaboration</td>
<td>143</td>
</tr>
<tr>
<td>10.4</td>
<td>Behavioural Trends</td>
<td>144</td>
</tr>
<tr>
<td>10.5</td>
<td>Specific Disciplines</td>
<td>147</td>
</tr>
<tr>
<td>10.5.1</td>
<td>Physics and Mathematics</td>
<td>147</td>
</tr>
<tr>
<td>10.5.2</td>
<td>Astronomy</td>
<td>148</td>
</tr>
<tr>
<td>10.5.3</td>
<td>BioSciences and Medicine</td>
<td>149</td>
</tr>
<tr>
<td></td>
<td>United Kingdom</td>
<td>149</td>
</tr>
<tr>
<td>10.5.4</td>
<td>General Considerations for Biomedicine</td>
<td>150</td>
</tr>
<tr>
<td>10.6</td>
<td>Case Study</td>
<td>151</td>
</tr>
<tr>
<td>10.6.1</td>
<td>Case Study – Information Search (Biology)</td>
<td>151</td>
</tr>
<tr>
<td>10.7</td>
<td>Arts and Humanities</td>
<td>152</td>
</tr>
</tbody>
</table>
Phase Three

Drivers for Change

Change can be complex

Financial and Administrative Drivers

Chapter 11 Business Models as Driver for Change

11.1 Opening up the Market?
11.1.1 Open Access Initiatives
11.1.2 Open Access Journals – the Gold Route to open access
11.1.3 Author self-depositing articles – the Green Route to open access
11.1.4 Harvesting the open access material
11.1.5 Open Access projects
11.1.6 Economics supporting open access
11.1.7 Impact of OA on Publishers
11.1.8 Trends favouring Open Access
11.1.9 Implications for Authors
11.1.10 Implications for Publishers
11.2 Online Advertising as a new business model
11.2.1 Online Advertising
11.2.2 Advertising in the scholarly area
11.3 Summary

Chapter 12 Funding Research as a Driver for Change

12.1 Political developments
12.2 Open Access Initiatives
12.3 Ranking countries by research output
12.4 National and International government initiatives
12.4.1 A model for a new electronic publishing paradigm
12.4.2 European Commission FP7 e-infrastructures
12.4.3 EU Study of Scientific Publishing (2006) .......................... 190
12.4.4 EU open access developments ..................................... 193
12.4.5 European Research Council ........................................... 195
12.5 Publisher Initiatives ................................................................. 195
  12.5.1 US publishers’ PR campaign ............................................. 195
  12.5.2 PRISM – Advocacy programme from the publishers ............ 197
  12.5.3 The European PEER Project .............................................. 198
  12.5.4 Publishers’ White paper on academic use of journal content ................................................ 199
12.6 Library Initiatives ................................................................. 200
  12.6.1 SPARC ................................................................. 200
  12.6.2 SHERPA/ OpenDOAR .................................................... 201
12.7 Global Research Trends ...................................................... 201
12.8 Research funding as a driver for change ................................. 203
  12.8.1 Public funded R&D in the UK .......................................... 203
  12.8.2 Structure of Research Funding in the UK ......................... 204
12.9 Research assessment .............................................................. 204
  12.9.1 A Dangerous Economy (RCUK) ........................................ 204
  12.9.2 The Death of Peer Review (RAE) .................................... 205
  12.9.3 The 2008 RAE .............................................................. 206
12.10 Other funding agencies ......................................................... 207
  12.10.1 JISC in UK ............................................................... 207
  12.10.2 MPS and DFG in Germany ............................................. 207
  12.10.3 Charities ................................................................. 208
12.11 Summary ............................................................................. 208

Technological Drivers

Chapter 13 Efficiency Improvements as a Driver for Change 211
  13.1 Industry Collaboration to achieve improved efficiency in EP .... 211
    13.1.1 Trade Associations ....................................................... 211
      Publisher Trade Associations ............................................. 211
      Library Trade Associations .............................................. 212
    13.1.2 Research Information Network (RIN) ............................. 213
    13.1.3 Publishers Research Consortium .................................. 214
    13.1.4 Publishing Cooperatives ............................................. 215
  13.2 Changes in Format .............................................................. 217
    13.2.1 Markup Languages ...................................................... 217
    13.2.2 Metadata .................................................................. 218
  13.3 Structural Efficiencies .......................................................... 219
    13.3.1 Mergers and Acquisitions ............................................ 219
    13.3.2 Economies of Scale .................................................... 220
    13.3.3 Why is market consolidation taking place in scholarly publishing? ............................................. 221
    13.3.4 Why are the larger publishers able to succeed where small publishers find it difficult? 222
  13.4 Standards and Protocols ....................................................... 223
    13.4.1 ONIX for Publisher Licences ........................................ 223
Chapter 14 Technology as a Driver for Change

14.1 Background ................................................. 229
14.2 Past impact of Technology ............................... 229
14.3 The technological infrastructure ........................ 230
  14.3.1 Digital Resource Management (DRM) ............. 231
  14.3.2 Athens .............................................. 231
  14.3.3 Shibboleth .......................................... 232
  14.3.4 UK Access Management Federation ................. 233
  14.3.5 OpenID ............................................. 233
14.4 Technology and Standards ............................... 234
  14.4.1 Digital Object Identifier (DOI) .................... 234
      Concerns .............................................. 234
  14.4.2 CrossRef ........................................... 235
  14.4.3 Other Identifiers .................................... 236
14.5 New Products and Services ................................ 236
14.6 Other Technical applications ............................ 237
  14.6.1 The “Cloud” ........................................ 238
14.7 Three predictions on scholarly communication technology ................................. 239

Chapter 15 Data and Datasets as a Driver for Change

15.1 Background ................................................. 243
15.2 Main data centres ........................................... 243
15.3 The Data Challenge ......................................... 244
15.4 Standards and Procedures ................................. 246
  15.4.1 Data Management Systems ......................... 246
  15.4.2 Metadata of Data ................................... 246
  15.4.3 Data Webs .......................................... 247
15.5 The Researcher’s wishes ................................... 247
15.6 Reproducible Results ..................................... 249
15.7 Integration between Data and Text ........................ 249
15.8 Business Model for Data ................................... 249
  15.8.1 NSF funds for data compilations ................. 250
  15.8.2 Google and Datasets ................................ 251
15.9 Impact of data on Libraries ................................ 251
  15.9.1 Data Curation ....................................... 252
15.10 Impact of data on Publishers ............................. 252
15.11 Impact on other institutions ............................. 253
  15.11.1 Research Councils .................................. 253
15.12 Summary .................................................... 254
Chapter 16 Mining of Text and Data ................................. 255
16.1 Background ............................................... 255
16.2 Implications .............................................. 256
16.3 The mechanism of Text Mining ............................. 256
16.4 Recent History ............................................ 257
16.5 Challenges facing Text and Data mining ................... 258
16.6 Practical Examples ........................................ 260
16.7 Implications in applying text mining ....................... 261
16.8 The Future ................................................ 261
16.9 Impact on Libraries ....................................... 262
16.10 Impact on Publishers ..................................... 262

Chapter 17 E-science and Cyberinfrastructure as Drivers for Change 263
17.1 Background ................................................ 263
17.2 The e-Science Challenge .................................. 263
17.3 Visions for e-Science ...................................... 264
17.4 Overall context of e-Science ............................... 264
17.4.1 Public Engagement .................................... 265
17.5 Future role of e-Science ................................... 265

Chapter 18 Workflow Processes and Virtual Research Environments 267
18.1 Integration into Work Flow Process ......................... 267
18.2 The research process ...................................... 268
18.3 Examples of a work bench approach ....................... 269
18.3.1 Virtual Research Environments (VRE) ............... 269
18.4 Summary .................................................. 270

Chapter 19 The Semantic Web as a Driver for Change .......... 271
19.1 The Challenge of the Semantic Web ....................... 271
19.2 Critiques .................................................. 271
19.3 Web Science Research Initiative .......................... 272
19.4 Examples of Semantic Web in scholarly publishing ....... 272
19.4.1 Knowlets .............................................. 273
19.5 Implications of Semantic Web for EP ..................... 274

Chapter 20 Mobile Devices as Driver for Change ................. 275
20.1 Background ................................................ 275
20.2 The wireless economy ..................................... 275
20.3 Intelligent spectacles? ..................................... 276
20.4 Amazon’s ‘Kindle’ ......................................... 276
### Chapter 21 Archiving and Preservation as Drivers for Change

21.1 The Challenge of Archiving and Preservation .............................................. 279
21.2 Preservation and Access ................................................................................. 279
21.3 Archive requirements ..................................................................................... 281
21.4 International Collaboration ........................................................................... 282
   21.4.1 US-based Task Force on sustainable digital preservation and access .......... 282
   21.4.2 The European Alliance for Permanent Access ......................................... 283

### Social Drivers

#### The Google Generation

### Chapter 22 Findability as a Driver for change

22.1 The rise of Search Engines ................................................................. 289
22.2 Resource Discovery and Navigation ...................................................... 289
22.3 How users find information ......................................................................... 290
22.4 The Findability Challenge ......................................................................... 292
22.5 Case Study: The Google mantra .............................................................. 293
22.6 Other search engines ................................................................................... 296
22.7 Impact of search engines on publishers ..................................................... 297
22.8 Book digitisation and the Copyright issue .................................................. 298
22.9 What of the Future? ...................................................................................... 299

### Chapter 23 Web 2.0 and Social Collaboration as Drivers for Change

23.1 Wisdom of the crowds .................................................................................. 301
23.2 The Challenge of Web 2.0 ........................................................................... 302
23.3 Critiques of the Web 2.0 movement ............................................................... 303
23.4 Case Study – O’Reilly ................................................................................. 305
23.5 The Web 2.0 business model ...................................................................... 307
   23.5.1 Blogs and Wikis ......................................................................................... 309
   23.5.2 Mash-ups ................................................................................................. 309
23.6 Drive towards Consumer-based Collaborative systems ................................ 309
23.7 Communication ............................................................................................. 311
23.8 Case Study – Wikipedia and online encyclopedias ...................................... 311
23.9 Wikinomics .................................................................................................. 313
23.10 Case Study: InnoCentive ........................................................................... 315
23.11 Summary ..................................................................................................... 317

### Chapter 24 Trust

24.1 Trust ............................................................................................................. 319
24.2 Fraud and Plagiarism .................................................................................... 321

### Chapter 25 Timeline – Emergence of Electronic Publishing

25.1 Where we come from .................................................................................... 323
25.2 Users of scholarly communication ............................................................... 324
Abstract

The traditional publishing model is being challenged as it adapts from a print-only to a hybrid print and digital mode of communication, and in some instances to a fully electronic publication system. There is no single issue that is driving this adaptation – it comes in a variety of forms with different drivers, though technology-related stimuli, emerging social trends and business model changes are the most prominent. This book reviews these various drivers for change and assesses their impact on an industry sector in transformation. An underlying theme is the effect which such changes is having on user behaviour – on whether the emerging structure is being led by changes in information habits by the user population or whether end users are being buffeted and driven into new ways of doing things by developments and trends outside their control.

The jury is still out on the crucial question of the power of users and their behaviour patterns as the ultimate determinant of how electronic publishing will emerge and what this means for existing stakeholders such as publishers and librarians. It is a jury without solid evidence, however. The data which could be used to assess the current social aspects of scholarly communication have in the main been under-analysed and under-reported. We know more, but not much more, about user behaviour in a print dominated information system. We know much less about the way the various user communities have acclimatised themselves to information as it embraces the digital age. The Google Generation in particular remains something of an enigma.

However, there is a new environment which is opening up for those involved in electronic publishing. In the past, in the print paradigm, a cottage industry mentality prevailed as many of the costs – notably typesetting and printing – were outsourced. In the electronic publishing world, to be a key player, content providers need to invest in platforms, in retro-digitisation, in acquiring ICT skills all of which have implications on economies of scale. Large conglomerates are now emerging, fuelled in some cases by venture capital money, which have been active in acquiring other publishers to achieve the new economy of scale. Other publishers have sought protection through aggregation and sharing some of their operating functions. New large players have entered this business sector, notably search engines which have become sizeable elephants in the emerging EP room. The culture of publishing is changing.

Combined with the change in industry structure there have been the Drivers, twelve of which have been identified and will be described later in this book. Some of the key features of these drivers are that ‘openess’ or freedom to access and use the results of the publication process has become high on the industry agenda. In
addition, a collaborative, community-based approach is emerging in some areas where ‘correspondence’, ‘communication’ and ‘computation’ – the three ‘C’s – have become the by-words as a counter to the traditional formalised science, technology and medical publishing. This is interacting with another major force – alternative information sources or formats to the publication of record. These are emerging as highly desirable aspects of the new communication system. Raw data and datasets, the ability to mine these to identify new relationships, and incorporating data and publications into an integrated work flow system which closely matches the way researchers work, has become a key topic – these are new information services which now sit on top of traditional content and are changing the nature of electronic publishing into a service rather than product-orientated business. In addition there is a growing demand for ‘reproducible results’. If you can’t reproduce the original results described in the published article through the reader’s own desk top there is no result.

Supporting all these broad trends is public investment in the networks which allow the new information to flow between author and reader, networks which are huge in their technical capacity and are part of national and global e-Science and cyberinfrastructural investments. Electronic publishing is changing faster now than it was ten years ago. In the next ten years the changes are likely to escalate rather than diminish.

There are however many issues which remain unresolved. Whether open access as a business model is viable and sustainable, and whether it offers up publicly funded research results to a traditionally disenfranchised community. Or whether the so-called ‘data deluge’ will change the way scholarly research is conducted and reported. What role will e-Science and the Grid take in communicating the results of research effort, and what impact will grassroots social collaborative tools, often referred to as Web 2.0 developments, have on the creation and assessment of research? This is a confusing mixture of trends, drivers and developments with no clear indication of how each will determine the shape of industry in years to come. An attempt will be made later to approximate a timeline for some of these drivers for change.

There is currently uncertainty among the main stakeholders in the industry. Given the changes, will ‘publishers’ survive with their alleged fixation over ownership of content and fail to provide new value-added services in a digital world? Will ‘librarians’, worried about the storage of material in physical buildings and the dwindling number of visits to access their collections, be bypassed in future by new information gatekeepers with new skill sets? Will the existing intermediaries – booksellers and subscription agents – who are losing ground to direct transactions between publishers and end users – be totally disintermediated? And if these traditional key players lose ground, which new agencies will emerge to take over the role of communicators of information in a truly digital world? These are some of the key issues which will be raised in the following chapters. Given the volatile nature of the industry at present it is by no means clear that anyone has all the answers.

This is a snapshot description of an electronic publishing system in 2007/8, a system which is on the move, and the move can follow a number of routes. An audit is therefore attempted of available information and opinions which is related to this emerging situation, and this audit is culled from a wide variety of sources.
Keywords

Electronic publishing, journal publishing, scholarly and STM publishers, research libraries, intermediaries, science policy, technology, user behaviour, authors, collaborations, open access, institutional repositories, e-Science, e-Research, Grids, data and text mining, e-books, mobile devices. social networking, social collaboration, blogs, wikis, podcasts, Web 2.0, semantic web, folksonomies, ontologies.
Acknowledgements

This book rests heavily on the shoulders of giants. Without access to some of the legendary work undertaken by experts such as professors Carol Tenopir and Donald King, from the universities of Tennessee and North Carolina respectively, the first phase of this book which explores some of the background and genesis of the electronic publishing industry, would have been all the poorer. The second phase, that of assessing the current structures and drivers rests on the shoulders of other giants such as professor David Nicholas and Dr Ian Rowlands from University College London whose work through UCL’s Centre for Publishing and CIBER has been both pioneering and illuminating. Also the many research studies commissioned by and undertaken by Michael Mabe whilst at Elsevier and more recently as chief executive of the STM Association has been an important source of industry insight as change takes place. Many of the scenarios which are painted in the third phase, the assessment of future trends and drivers, have relied on the thoughts and vision of David Worlock, senior research fellow at Outsell Inc, amongst others. Research commissioned in the UK by agencies such as Research Information Network (RIN) and the Joint Information Services Council (JISC) have also figured prominently in compiling this book. The support and help from colleagues at the British Library has been exceptional, and grateful thanks are extended to all concerned.

This book is truly a collective effort in pulling together developments, ideas and strands from a diverse group in an attempt to describe the current confusion that is electronic publishing.

Notwithstanding the eminence of the sources used, any omissions, mistakes or false interpretations made are entirely those of the authors. After all, it is a diverse and diffuse industry segment, one which is undergoing dramatic change which is itself a breeding ground for inaccurate subjective assessments and misrepresentations. And this is the by-word of the book – that Change is dominating the scene and all players involved in electronic publishing need to come to terms with Change in its many guises. To make the assessment more manageable, a carefully circumscribed aspect of electronic publishing has been made – electronic publishing as it affects libraries and publishers is a broad church. This book is focused on that one part which is at the forefront of adaptation to the changing environment – the scholarly information area dealing with research, particularly in the hard sciences. It is contended that work being done in these areas will spill over into other publishing areas in due course. But even this circumscription on science, technology and medicine does not guarantee that mistakes and misinterpretations will not be made.

Electronic publishing is an important topic. It is a function that lubricates the wheels of the scientific research effort in particular, providing essential informa-
tion to ensure that new research findings are part of a consistent corpus of activity, and not undertaken in isolation. Over the decades and centuries mechanisms have been established by the academic and industrial research communities to ensure that project findings are credible and that recognition is given to those who merit opprobrium from their peers. It is this mechanism that the new electronic publishing system is testing, and this book builds on the suggestions, advice and information supplied by many in the industry in looking at how robust the current mechanism is.
Chapter 1
Background

1.1 Introduction

This book is about the mechanisms which facilitate the exchange of ideas and information within those sectors of the modern economy which advance the cause of innovation, new ideas and cultural exchange. Electronic publishing is that mechanism. The book describes the changes which are taking place in the way individuals and organisations communicate, and how they adapt to the new information systems which electronic publishing is creating.

Electronic publishing is often the hidden process at the heart of academic and corporate research, of professional and learned society communication, of collaboration between research communities. It fuels the research effort, both nationally and internationally. It is a harmonising activity, and in its most effective form has a democratising role as it facilitates dissemination both widely and rapidly.

However, it is in a state of flux. As indicated in the Abstract, there are many drivers for change impacting on electronic publishing. They come from different quarters, and their confluence is producing turmoil. As will be emphasised during the book there is concern that the traditional stakeholders in the information sector are at risk. Their survival is threatened. Talk of dinosaurs and ‘heads-in-sand’ have frequented meetings at which one stakeholder takes pot shots at another. As a consequence, at stake is the health of the very social and industrial processes which are at the heart of a vibrant and growing economy.

As much as possible, evidence will be substituted for emotion – emotion has tended to enter into much of the dialogue between the separate parties involved in the communication process. This unfortunately often obscures the real facts and emerging trends. Where evidence and hard facts exist these will be highlighted, though their paucity is frequently lamented upon. But it is a critical conclusion from this book that more attention be given to extracting facts from the current situation in scholarly communication and isolate these facts from the hype and rhetoric which unfortunately dominates the listservs, blogs and wikis. These latter have emerged to provide sometimes unhelpful commentary on what has become the ‘scholarly communication problem’.

1.2 The Book

of DJB Associates on behalf of the British Library. It painted a scenario whereby
the Internet was just beginning to grapple with the scholarly communications pro-
cess, and CD-ROMs were felt to be the most likely digital message carrier of the
future.

As has become evident, the past ten years have seen a dramatic change in the
way electronic publishing has emerged and transformed traditional print-based
publishing systems. These electronic publishing systems have also affected the
relative importance of different electronic media in communicating scholarly in-
formation. The importance of CD-ROM technology as a support mechanism was
over-exaggerated in the original book, whereas the effect of the Internet and Web
developments were greatly underestimated. New business opportunities have
been awakened by open access developments, e-Science trends and Web 2.0 based
social collaborative tools, all of which were barely evident in 1996.

As such the publisher, Saur and its patent company de Gruyter, has commis-
sioned a follow-up report to address these unforeseen and unforeseeable changes
and to assess their impact on the current publishing and library sectors in par-
ticular. The approach adopted in this book, as in 1996, is to take a snap-shot of
electronic publishing developments and make an assessment of them based on
currently available knowledge and evidence. It offers no ‘final solution’ for how
the electronic publishing industry will actually emerge from the buffeting it is tak-
ing from many external events. However, it will attempt to put the many changes
which are taking place into some semblance of order and describe their potential
impact on the overall scholarly communications process.

The journey is still underway – given the experience of Google, for example,
we can expect a rocky ride. New challenges and opportunities are likely to appear
in some cases from entirely unforeseen quarters. It has meant that this update of
the first attempt at investigating ‘Electronic Publishing and Libraries’ (1996) will
not follow the earlier practice of forecasting the future in a precise and quanti-
tative way. Change is currently too volatile and the risks of getting it wrong are
too great. Instead this book follows a more qualitative assessment of trends and
developments, underpinned by quantitative data where available, in the hope of
enabling the reader to come to his or her own judgement about the current and
future state of Electronic Publishing.

1.3 Purpose of the Book

So why write a book now about electronic publishing when little is settled and
change is being claimed by us to be so rampant? It is evident that the rules of
engagement are altering dramatically in electronic publishing (EP) from a situation
where print publication and dissemination dominated, as recently as two decades
ago, to a point where digital publication and dissemination plays an increasingly
important role. This raises questions about the very direction and key players in
the scholarly communications sector. Is this critical communications system stable
or will it fragment into many separate pieces?

Except for a few specialised exceptions, we are not yet in a situation where
digital information resources totally dominate the information-seeking behaviour
of researchers. We are still in a hybrid system, though there is undoubtedly a rapid
migration from print to digital as the IT infrastructure becomes more reliable,
ubiquitous and secure. This book identifies some of the key issues which have arisen as this hybrid model envelops the scholarly communication process. As such it is intended to become a platform from which the debate about a digital communication system can take place. This book aims to bring together in one place those important features of the changes that are currently in evidence.

It also describes the tensions that currently exist between the main players in the digital publishing sector. Some of these tensions are inherited from the past, and often revolve around the legal ownership of content and its use. But new tensions are also evident as they are driven by the enabling power of technology. Some of these tensions arise from the ease with which content, in a digital world, can be disseminated and replicated without recourse to the content owners. It is a significant feature and debates about the future development of the research sector’s information requirements have become quickly polarised. There is nonetheless the need for mutual cooperation and collaboration to ensure that a smooth transition to the new digital information order is effective, efficient and viable. There is a risk that we focus on short-term issues at the expense of moving towards a long-term vision about which there is currently limited consensus. This is where emotion has taken hold, and conflict arises as each part of the information sector tries to protect its own short-term patch without any real understanding of the other sectors’ positions and how the whole fits together. There is the lack of a consistent, holistic vision of the future which all stakeholders can buy into.

A fundamental issue which this book has to face is whether publishers and librarians as we know them have outlived their purposes in the new digital environment. Are they hanging on to functions that have a diminished or have no place whatsoever in the new Millennium? Are they dinosaurs, with the impact of digital trends having the same consequence on publishing and librarianship as meteors did on dinosaurs? Are they clinging on to out-dated concepts, particularly of ownership when openness is increasingly the name of the game? Are they showing signs of failure to adapt given new circumstances and environmental change?

Changes can be evolutionary – a gradual change over decades – or alternatively they can be revolutionary, making substantial inroads into the way communication takes place. Both types of changes are seen within scholarly communications. In the traditional aspects of communication, the way people work with information resources, the changes have been evolutionary in nature as most people have slowly adapted their habits to the new (digital) information formats and technical infrastructure. More recently, the rapid development of the Internet and associated technologies has had a major impact on the way in which information can be obtained, consumed and created. This in turn has opened up new avenues of opportunity for the users of these information resources as long as they are prepared to adapt to technological advances. The younger generation, and scholars operating in high-technology areas have become particularly influenced by such developments. As such there is a mix of both evolution and revolution in the emergence of scholarly communication as we see it.

One striking aspect of this confluence is that it has produced a new social phenomenon, it has created the ‘digital native’ or the millennium generation. These are individuals, usually young, who have been brought up on a diet of video games, interactive online services and technological gizmos which have dampened their enthusiasm to gain information and knowledge from the highly structured printed publication system, either on paper or its digital version. Instead they expect to
find their information through comprehensive and powerful search engines, such as Google or Yahoo, or increasingly that information will be delivered to them in anticipation of their demand for it by new specialised and targeted information services. Some are moving into a social networking form of mass collaboration, potentially eschewing the traditional highly structured refereeing systems that have been the bulwark of traditional publishing. The increased participation in services such as MySpace, FaceBook and SecondLife – with now millions of adherents – point the way to a new form of collaboration in content creation. Intelligent information systems are also offering enticing but as yet ill-defined opportunities for the emerging digital natives.

This book will look at the background to these Changes, where they originated, what the implications are and how things may develop in the future. The central purpose is to highlight the complexity of the current situation and the need to be aware of the holistic nature of the scholarly communication process in coming up with solutions which meet the external conditions and criteria.

### 1.4 ‘Change management’

Fundamentally, there has been a growing disconnect between the supply and demand forces for published information supporting research in recent decades. These forces have been rooted in old systems which served print-based book and journal publishing reasonably well – that authors submitted works to publishers who created product which they sold, through intermediaries, to the library who paid for them from budgets allotted by their institution or company. All this was in anticipation of a demand for information and knowledge being met by the library for their constituency.

However, in a major part of the digital publishing system – in scholarly and research publishing – there has been a growing tension caused by the imbalance between the forces (R&D funding) which fund the creation of the traditional product (books and journals) and those forces which are responsible for buying and using the product (libraries within their respective institutions and companies). The two are financed from separate agencies and budgets, with little or no structural interconnectivity. This lack of balance has created the so-called ‘frustration gap’ between supply and demand. A practical illustration of this can be seen from the different growth rates of ARL research budgets in the US and the funds identified by the NSF in undertaking research (see Chapter Two). During the 1990’s in particular there was little correlation between the supply and demand curves for research-focused information.

In addition, any assessment of the scholarly communication process should be mindful of the changes which are occurring as a result of the impact which the Internet, the web, the semantic web, cyber-infrastructural developments and social publishing is having on user behaviour. Changes in interpretation by some sectors of the role of copyright, intellectual property rights, legal deposit, interlibrary lending, etc, are also adding confusion. Some of the many specific areas of ‘change’ include:

- **Changes in Technology.** We are in transition between a print and a digital publishing market. Increased power of desktop computers and speed of communica-
tion is having a powerful influence on the form which scholarly information is delivered. In this respect the e-Science and e-Research movements are becoming significant.

- **Changes in the Business Model.** Coinciding with the technology change, there are also pressures on developing new ‘free’ ways to access information in line with social imperatives. Open access takes many forms, most of which are still unproven in a commercial sense, but nonetheless is still dominating the debate about the current publishing landscape and is central to any realistic discussion about Change.

- **Changes in the Product/Service concept.** Instead of a total reliance on formally published mainly textual research articles there is a growing requirement to provide access to supplementary material, grey literature, datasets, manipulation software, video and sound clips etc, which are increasingly part of the research output but are yet to establish widely accepted mechanisms for certification, dissemination and preservation.

- **Changes in User Behaviour.** Partly as a consequence of the available cheap and ubiquitous technology, authors and readers are changing their patterns of accessing and assimilating information and adopting digital and often informal communication methods.

- **Changes in Scientific Disciplines.** It is no longer possible to consider scholarly or research information services as a ‘one size fits all’ – each research area has its own culture, tradition, scale and format needs. Even more apparent is that they are each moving in different directions in their adoption of digital media. Some are embracing the new information technologies much faster than others.

- **Changes in Funding Agencies.** The approach that research funders are adopting to evaluate the results of their fund allocations (for example, through research assessment exercises) is resulting in a more metric-focused and evidenced-based evaluation of their grant programmes.

- **Changes in Copyright.** Whilst academic institutions are demanding new rights over work undertaken within their campuses, the issue of deposit mandates to ensure broad access to research outputs has become a growing trend. This has been fuelled by agencies such as Creative Commons and Science Commons that make the transfer of such rights away from publishers, the traditional holders of published copyright, easier.

- **Changes in Demography.** The rapid rise in social collaboration tools and user-generated media is breaking down the traditional formal scholarly publication systems producing an increasingly hybrid environment.

The combination and interaction of these changes creates an urgent need to understand and tackle the problems facing the scholarly communication industry. The danger of ignoring the effect of these is to foster an unstable and unreliable dissemination of research outputs.

Changes need to be made in an incremental way that is viable, sustainable and optimal in terms of providing users with the information they need. Revolutionary changes may destabilise this communication system causing loss of information, insight, time and money.
1.5 Target Audiences

This book is therefore aimed at all the current and future key stakeholders in the scholarly communication process, but particularly focused on publishers and libraries. It endeavours to highlight the underlying, and in some instances the unreported, trends that are creating Change, and the significance of these trends. The book speculates on the impact that these changes will have on the existing stakeholders. Whilst it will not be dogmatic in describing such impacts, it will draw on best evidence to produce scenarios within which existing players can assess their own strengths or weaknesses.

A basic fear – the thing that keeps many pundit awake at night – is that the simultaneous adoption of all that is claimed to be good for electronic publishing in principle could in effect throw the baby out with the bathwater and leave a future electronic publishing (EP) structure with no content. A viable mechanism must be found to ensure that authorship is rewarded, and reward in this instance may take a variety of financial or social forms. Ignoring this central tenant of the scholarly communications process would mean that the main purpose of electronic publishing – to keep people informed in a timely, efficient and cost-effective way – would not be achieved. There has to be value and purpose which attracts authors to disseminate the results of their ‘sweat of the brow’.

This book is addressed to all those sectors of the scholarly communication industry that are instrumental in bringing a new EP system into place. To create a common understanding of the issues among these partners. To ensure that sound bytes do not dominate the debate. Where possible use evidence and metrics to inform decisions; and that appreciation of the needs of the links in the information chain also be made.

1.6 Definitions

Before delving into the substance of the challenges facing electronic publishing, a few brief definitions of intent need to be made.

This is a book primarily about the scholarly information sector, the area that is in support of the scientific, technical, engineering and mathematical (STM, or STEM) research activities. This sector has particular needs that are often at the cutting edge of commercial, social and technological change. Innovation, leading to improved industrial products and services of benefit to society, is often at the basis of such needs, though basic research to understand the world we live in is also a key driver in academia. Such information is needed to push back the boundaries of knowledge, each research result building on others, which has led to the concept that advances are made by ‘standing on the shoulders of giants’. (a quote inaccurately attributed to Sir Isaac Newton in the first instance – it predates him by several centuries, ostensibly to Bernard of Chartres in the 12th century).

In addition, the book is focused on research, not education, nor entertainment, as the main stimulus for information demand.

The tradition over the past four centuries has been for the slow evolution of the journal to become the leading format for delivering scholarly information. For reasons that lie within the current budgetary mechanisms, such serials as journals have gradually replaced books as the main format in the more popular STM
sectors. The so-called ‘article economy’ was flaunted as a more recent way to un-pick the relevant items from a journal and deliver just what is needed on demand, but the inroads which such article delivery has had has been at best spotty. In effect, for centuries, the printed book and journal formed the mainstay of scholarly communication. During the past twenty years this is showing signs of change.

For example, a significant and recent development has been the emergence of data and datasets as the primary research resource that some specialised researchers need to conduct their work. The data itself is not encumbered with interpretations by other researchers about how it can be applied in other situations for other purposes. In most cases the data is neutral and as such a valuable new asset for the STM research community. This was rarely available in the pre-digital era as a scholarly resource. So in terms of product format for electronic publishing the net is being cast wide to include such new developments.

1.7 Objectives

This book attempts to pull all these strands together to allow the reader a broader perspective of all aspects of the electronic publishing industry into which we are being driven by events over which we do not necessarily have control. Whilst we may not have total control, it behoves those of us who are active in the industry to understand all the issues facing the parties involved in the sector and exert influence.

This book comes with a plea – that we remove the blinkers from our respective sectoral eyes and understand that in the volatile information world into which we are moving we all need to stand together, to work together, through a common understanding of each other’s needs. To use the perhaps over-worked and inappropriate recommendation from Benjamin Franklin at the US Declaration of Independence, 1776, “We must all hang together, or assuredly we shall all hang separately”. More importantly, to ignore legitimate roles for libraries, intermediaries, data providers, publishers and researchers all for the sake of protecting a particular narrow vested interest could hasten the demise of the current publishing system and lead to something totally new, something which could be alien to the best interests of the industry and its users.

This is not a plea for a Luddite approach or support for a no-change scenario. There are undercurrents that already exist driving a different pattern of relationships between the present links in the information chain. What is intended from this book is that there are consequences that arise from our respective responses to change, and these actions should be as much positive as defensive, in harmony rather than self-serving. Ultimately the market will decide. The next chapter of this book looks at some of the issues that will determine this emerging market.

1.8 Approach adopted

This report deals with three main phases in the way electronic publishing has impacted on publishers and libraries.

Phase 1 – This is the period up until the early 1990’s when publishers and libraries were faced with a largely print-based information system. Though electronic me-
dia existed it was nevertheless early days and dominance was with publishing and curating the printed page. It was the period covered by the earlier book on “Electronic Publishing and Libraries – planning for the impact and growth to 2003”. It was a simple period in media terms, with the main problems encountered being economic rather than technical or social.

Phase 2 – From the mid 1990’s to the early 2000’s was a period of confusion. The arrival of the Internet created a new dimension to the information industry, bringing with it a whole set of new legal, business and technical challenges many of which remain unresolved.

Phase 3 – From the early 2000’s there has been a strong electronic publishing drive, with digital versions of information outselling and outperforming their analogue equivalents. Adjusting to this new all-digital environment has created new strains, with mergers and acquisitions dominating the publishing scene, massive investments in infrastructure being put in place by policy makers, and a period of ‘openess’ overtaking the traditional ‘toll-based’ access to information as a business model. In this phase three the dozen or so main drivers for change will be identified and described.

Throughout these separate phases the user needs have tended to be ignored, under-researched or marginalised. Nevertheless, whilst it may be somewhat of a cliché to claim that it is through an understanding of users that one achieves insight into the new electronic publishing system, this will be the undercurrent which runs through this book. Understanding the information habits of those who are being provided with information is fundamental.

Cliché or not we will begin this tour by identifying some of the macro trends which have impacted on the social trends within the industry and do this by drawing on parallels from other industries and situations. Examples of those which relate to the traditional print-based publication system, Phase One, will be used as a starting point. Such examples paint a clearer picture of where we have come from.
Phase One
We can look back in history to see whether there are lessons to be learnt about how to assess and cope with the current situation facing the publication system. A number of conceptual models have been proposed in the past which have some relevance to our ability to understand aspects of the current position. Whilst it is fair to say that we have never before trodden the path we are now following, some conceptual models from history may help us place things in perspective.

These conceptual models are abstracts. They provide a framework within which we can measure the significance of past events and place them in relation one with another. The first of these concepts addresses the key departure point for many aspects of electronic publishing – the financial stresses which lay within the printed publication system and why these were important in setting in train a series of developments which culminate in discussions over the appropriateness of prevailing business models.

2.1 Tragedy of the Commons

The concept of the Tragedy of the Commons involves a conflict over resources between individual interests and the common good. It comments on the relationship between free access to, and unrestricted demand for, a finite resource. The term derived originally from a comparison identified by William Forster Lloyd in his 1833 book on population, and was then popularised and extended by Garrett Hardin in his classic 1968 essay published in *Science* entitled “The Tragedy of the Commons.”

Under the Tragedy of the Commons concept, the common land would be over-grazed in medieval times until such a time as one extra beast tipped the scales and made the ‘common land’ or good totally useless for all. It would reach a stage where nothing could survive on the common land. All the herdsmen would suffer not just the last one. This collapse would happen quickly and was irreversible.

What is its relevance to the scholarly communication process? The often unwritten assumption by critics of the scholarly communication process as it existed up until the 1980’s was that scholarly publishing was headed in this same direction – that at some stage the collective library budget, the source for purchasing most of scholarly communications, would be insufficient to cope with the ever-expanding individually produced research output. The system would self-destruct dramatically and quickly under the strain. As new media and new versions of existing publications emerged the stresses would be ever greater and the collapse of the
system more imminent and catastrophic. The Tragedy conceptualises, in a way unintended by Hardin, the problem facing the pre-digital publication system in which publishers produced books and journals on an uncoordinated basis. This took no account of the fixed resource – the library budget – available from which to purchase their output. As such the extra book or journal subscription could become an unbearable burden for the library budget and new ways of serving their clients would have to be found by the librarian to justify their existence.

The economics of publishing is driven by the research output produced from an expanding R&D effort by society. This bears little relationship to the budgets being allocated by individual institutions to their libraries.

The Tragedy of the Commons did not happen. The switch from a print-based publication system to a hybrid and increasingly digital one has produced solutions which have given flexibility to the buying system, and enabled more information to be absorbed without causing the budgetary system to collapse. But it does indicate that there was something inherently flawed with the traditional mainly serial-based publishing system. There was a disaster waiting to happen. Any suggestion that there were halcyon days of scholarly publishing is pure myth.

### 2.2 Frustration Gap

Another way of describing this imbalance between supply and demand forces is to look at the expenditure by a country in its national research and development budget as compared with the expenditure on research libraries during the same period. For the United States such a comparison is possible and is illustrated below. The growing gap between the R&D infrastructure and the information support system is evident.

![Growth in Research vs. Library Spending](image)

**Figure 2.1** US Academic R&D Expenditure and ARL Library Budgets 1982 Constant Dollars

Sources: ARL and NSF
The frustration gap is the difference between the supply of publications arising from the R&D spend, compared with the ability of the library to buy these publications (the demand) as reflected in their overall budget increases.

Another significant metric showing how librarians have had difficulty coping with the inexorable growth in publishing output during this period is the falling share of the library budget, in comparison with the overall institutional budget, during the past decade. Between 1994 and 1999 the percentage share of the institutional budget of the leading higher education (university) establishments in the UK was approximately 3.1%. Since then there has been a gradual decline each year to 2.7% in 2005. Whether this reflects increased competition for the institutional budget, or whether it indicates the declining relevance being attached to the library and its role in the institution, is difficult to say. But there is a real concern that libraries cannot use the same metrics to advance their financial cause as can other departments within the same institution. It may be easier for a research centre to demonstrate tangible financial payback from additional investment in terms of additional students, extra postdocs, more research grants, etc. Payback which comes from an expansion in the library has less direct and immediate financial consequences. It is an intangible infrastructural item – in many cases a necessary rather than desirable entity.

The following chart illustrates more vividly the decline in the library share of all types of higher education establishments in the UK.

![Figure 2.2 Library as % of total institutional spend, 1992–2005](image)

UK Higher Educational establishments

### 2.3 Publisher and Library dissentions

Some prominent librarians have complained that publishers, notably the large commercial publishers who have become the dominant players in this sector, were sucking out the lifeblood of Science by not only overcharging for its publications, but also by restricting access to most of the corpus of relevant published information. Furthermore, the surpluses, the profits, were taken from the public science
effort and used to line the pockets of the rich capitalists who owned shares in the major European commercial publishers. There was no loop back into investment by and within the publicly-funded scientific effort.

As a consequence, it has also been claimed, libraries have been compelled to cut back on acquisitions and other operating costs on a regular and sustained basis. Cancellation has become a keyword, and it has become more difficult to justify expenditure on new information products. In the face of what may be perceived to be an erosion of the library as a core function or service within research institutions, some library advocates have recommended a much more radical approach to adapting the publication paradigm in favour of one which takes a higher moral ground, one which is more favourable to society’s overall needs. This has involved questioning the dominant position which the publishers have in the flow of information from creator or author to user. This is particularly the case in science where it is pointed out that that only 1–2% of the total research effort is in the communication process yet it is the control which publishers have in this which is stultifying the whole information dissemination system.

On the other hand, publishers have claimed that the so-called ‘Serials Crisis’ was largely of the making of librarians – that they had failed in their public relations and organisational politicking to convince their financial paymasters of the need for additional funds to buy essential publications. The importance of the library as the centre for knowledge exchange within an institution was being weakened and librarians had failed to come up with satisfactory arguments to protect themselves and their budgets.

This antagonistic atmosphere has existed for many years, and was enhanced more recently by bitter wrangling over specific areas of disagreement such as:

- Document Delivery particularly in the 1980’s and early 1990’s
- Interlibrary Loans particularly in the 2000’s
- Open Access as a business model (ongoing)
- Support for Institutional Repositories (IRs) just beginning
- Derivative works as an acceptable research procedure by third parties (under discussion)

It has furthermore spilled over into questions about the validity of legally sanctioned activities (in some countries) such as e-Legal Deposit. It also now threatens to undermine efforts to make effective and appropriate the application of new technology, new business models, adopting new paradigms and to generally ensure that the scholarly communication process in an electronic/digital environment is made efficient. There is a tradition of mistrust between some sectors.

Within this ongoing debate, the sociology of scholarly communication is emerging as a discipline in its own right. It embraces a number of motivational changes that are impacting the way individuals, kindred groups, and disciplines come to terms with the many external forces affecting them as a result of the change from a print to a print plus digital paradigm in the scholarly sector. But because these changes are mainly recent, the full impact of them on the scholarly communication process is difficult to measure.
2.4 Big Deals

Towards the end of the 1990’s a new business model emerged which toned down some of the conflicts which were threatening to engulf publisher/library relations. It arised out of a proposal first put forward in 1995 by Academic Press, now part of Reed-Elsevier, that libraries could be offered access to a complete package of Academic Press journals for a small additional charge over their current subscriptions with the publisher. The IDEAL (International Digital Electronic Access Library) and APPEAL (AP Print and Electronic Access Licence) was borne out of a new way to offer a consortial licence. It was a business model driven programme, not one which relied particularly on either end user feedback of technological development for its conception. A key point was reached when the Higher Education Funding Councils in the UK supported the idea of Big Deals and financed an experiment with a small group of significant journals publishers to see how the principle would work in practice within all UK higher education institutions. This included APPEAL as its basic template.

It was evident that libraries rarely subscribed to all journals from any one publisher, so in many cases the small additional charge which the publisher would levy would increase the holdings of the library dramatically. Albeit that many of these added titles were not always central to the needs of the users of the library in question.

Nevertheless, from 1998/9 Big Deals a new business relationship took off. The net effect has been that the average costs per title purchased fell whilst those publishers able to offer Big Deals have increased their revenues through the marginal increase in the price for accessing their total package. Publishers have been able to show that the price index for journals has fallen and that there are now more titles being subscribed to by libraries in comparison with the pre-Big Deal period. This process was made even more visible as these so-called Big Deals were extended beyond a deal done between a library and publisher to a consortium of libraries and a publisher, and in some cases consortia of publishers.

There are unforeseen consequences as a result of adopting Big Deals which will be explored in more depth later. In brief, the main adverse impacts have been:

- Selection of materials for the library collection are being taken out of the hands of the library collection development staff and shared with publishers
- It has been contended that there is a large amount of unwanted, unused material being put into the library domain
- Only those (large) publishers that have an extensive journal publication list stand to benefit, whereas the smaller often specialised learned society publishers lose out.

Whilst Big Deals have been seen as a godsend to the library community at a time when the serials crisis was about to bite hard, some five years later we are seeing a return to a similar set of concerns facing the librarian. It is recognised that the library budget is finite; that those who are able to offer Big Deals have made their offers; and that there is a growing impression that Big Deals have distorted the publishing system in favour of the large publishers’ lists irrespective of the quality of individual titles. The Big Deal itself is now coming under closer scrutiny.
2.5 The Tipping Point

But at this stage we need to consider the mechanisms that lead to a change occurring in the adoption of electronic material within the library and society in general. Will other new media formats have to follow a particular route in order to become established? Can adoption of more efficient means of scholarly communication be brought about quickly and effectively, reducing the cost to society of perpetuating old, inefficient systems? The concept of the ‘tipping point’ may help here.

According to Malcolm Gladwell, in his book entitled “The Tipping Point – How little things can make a big difference”, (2000), new innovations do not necessarily occur or take hold for logical or apparent reasons. Changes or tipping points can be stimulated by a number of events. Some of these are informal, subjective factors which defy the norms of efficiency. In fact, according to Gladwell, successful ideas, products and messages behave as epidemics and viruses. They become contagious. Even the smallest of factors can have a profound effect on creating a virus, and this can happen quickly. After the tipping point is reached subsequent progress as a result of the virus taking hold occurs geometrically.

Gladwell claims there are three rules which set off an epidemic.

1. The first is the Law of the Few – which means that a few key individuals can have a significant influence on creating change. They are connectors, mavens and salesmen. Connectors know lots of people – the average personal contact network is claimed to be twenty-one people but connectors tend to know far more than this. They are creators and sustainers of a wide network. They are on first name terms with movers and shakers in the industry. Mavens, however, are people who are very well informed and share this knowledge willingly. Maven is someone who accumulates knowledge. They keep the marketplace honest. They are not persuaders – they are teachers. Salesmen have the skills of persuasion. They tend to be subtle. They cannot be resisted.

These different ‘people types’ can have a profound influence in effecting a tipping point, in making a new product or service successful. Some combination of the three ‘types’ are required to get things moving, to effect a change or an innovation which succeeds spectacularly. There are some obvious candidates who we see as connectors, mavens and salesmen in the current controversies over aspects of scholarly communication – notably in the area of open access (OA) adoption.

2. The second rule of the epidemic is to create ‘stickiness’. There has to be quality in the message. The elements of stickiness in the message may be very small. There is usually a simple way to package information that can make it irresistible. For electronic publishing this can be tied to inherent technological progress within society leading to a ‘better’ information service – a key stickiness factor.

3. The final epidemic creator is the power of context. Epidemics are sensitive to conditions prevailing at the time. Getting an epidemic going involves a different set of human profiles – innovators, early adopters, early majority, late majority and finally the laggards. The first two are visionaries and risk takers, whereas the early majority avoid risks and are pragmatists. There is a chasm between the two. This is where the connectors, mavens and salesmen have a role in generating the epidemic. They translate the message from the first group to the second.

The point is that there are a variety of social mechanisms behind a change in acceptances and attitudes within society. This is as relevant in electronic publishing
as elsewhere. It means that technological efficiency by itself is not enough. Some of the Tipping Point ingredients, the social mechanisms, are necessary. Have the more significant aspects of electronic publishing achieved Tipping Point? Some have, some still have some way to go.

The next sections of this book will look at the evidence to see whether the above has made a substantive impact on the way researchers and academics communicate, or whether the preference remains with the printed word. Several studies have been undertaken which suggest there is indeed a radical change in user behaviour some ten years after the commencement of the electronic publishing revolution. But it is not an evenly distributed progress. Some parts of the electronic publishing still need the connectors, mavens and salesmen to be more active – for example many of the recent author studies show that, despite the claimed advantages for authors in having their articles published in open access format, as many as 90% of authors are still unconvinced. ‘Tipping point’ issues have not yet taken hold on the research author community.

A key feature of the scholarly communication process, and its adoption of electronic publishing, is that the numbers of individuals involved are large. It is also very diffuse, with huge differences between subject areas in how they adapt to EP. There are strong geographical differences, with the western world separating itself from the developing world in its ability to adopt and leap-frog innovative communication systems. There is no easy compartmentalisation of the scholarly communication process – it is a spectrum of interests, motives, drives and ambitions. Each has its own tipping point.

2.6 Open Access

The tipping point is frequently applied to a new business model which has arisen in electronic publishing. It is suggested that the failure of this model, the ‘open access’ model, to make significant inroads into the traditional subscription-based publication system is that the key elements of tipping point have not yet been reached.

In effect the tipping point concept has to compete with a well-established process in publication of research results, honed to a level of acceptability by the community over centuries. Despite the publisher/librarian tensions referred to above, there is a clear set of activities undertaken by journal publishers, who produced more and more articles (growing at 3.5% to 4% per annum) in response to the greater investment in scientific research since World War II. This has been balanced by librarians who bought annual subscriptions to these journals largely on a title-by-title basis, though with a growing number of cancellations as their largely static real materials budgets bit.

Superimposed on this structure which showed signs of some dysfunction a new movement appeared on the scene. This ‘open access’ model has as its dominant feature the provision of ‘free’ published information to all those within the community who might want to see and make use of the research results. Supporting this concept of ‘freedom’ is the much bigger open source movement which was sweeping through parts of the computer, IT and telecommunications industries. This has begun to spill-over into scholarly communication.

Many experts questioned why publishers have in the past increased their prices way over comparable price indices – publishers replied that the output of articles
was growing inexorably and that due to cancellations of journal subscriptions, they were forced to increase prices to compensate. Was this an indication of a market mechanism which was unstable? Both the Scientific and Technical Committee of the House of Commons in the UK, and the European Commission, have investigated the issue, with the former coming down strongly in favour of changing the market mechanism to correct its inadequacies (and the Commission still ruminating over the issue). There is a sense that instability is rife and growing. Some point the finger at publishers as reactionaries, protective over their profit margins and potential dinosaurs as the market mechanisms become more efficient.

The alternative market mechanism being proposed is a switch from the library bearing the brunt of journal subscription pricing to either the author paying to have his article refereed and copy-edited by the publisher, or for the institution to load the author’s final article (before publication) onto their server. This is the so-called Gold Route to open access (author pays) or the Green Route (whereby articles are hosted on the local or subject-based institutional repository). There are other variants around this theme, including the important source whereby authors make their articles available through their own personal web site. In all cases the articles are then accessible for free. Anyone – fellow researchers, peers both home and abroad, professionals, amateur scientists or laymen – all can access and download the article, in theory.

The problem comes with copyright. Copyright is what protects the author from seeing his or her work misused and misappropriated. The publisher has traditionally taken over the responsibility for ensuring that such works are protected by inviting authors to sign a copyright transfer form on acceptance by the publisher for publishing the article, and then policing this to see that no infringements occur. This has led to complaints that publishers have restricted the access to the fruits of society’s investment in R&D, and that such a restriction on the final stage of the research process by organisations that had no role in the rest of the research cycle was unacceptable.

In response, copyright forms have been made available online through a Stanford University based operation headed by Professor Lawrence Lessig and entitled the Creative Commons. This gave power to the author to decide what aspects of copyright should be exerted over access to their article.

Crucial to the future structure of STM is how much of a stranglehold the copyright and intellectual property rights (IPR) issue will continue to have over published output. There are signs that this is weakening as Creative Commons licences, and licences to publish from authors, gain pace, but as long as STM publishers prevent the reuse of their final published version (the Record of Science) many of the new STM information services will be stymied.

Nevertheless, the legal changes being experienced in this area are causing concerns among traditional journal publishers who see their business models being challenged and their power base eroded. The larger publishers in particular are looking at redefining their business and moving from a content-focus to a service orientation. More of this later.

The whole edifice in favour of the open access school of thought is built around the theory that there are more people who stand to benefit from access to the published research results than those who have bought access through a subscription or licence. In this respect the concept of the Long Tail is intrinsic to establishing whether there is a bigger ‘latent’ market from the one being reached traditionally.