Michael Nentwich

cyberscience

Research in the Age of the Internet

Austrian Academy of Sciences Press
Vienna 2003

CONTENT OVERVIEW

Foreword V

Content overview VII

Detailed list of contents VIII

List of tables XV

List of figures XVI

List of overviews XVII

Executive summary XIX

Part one: METHODS AND THEORY **1**

0 Project outline and background 3

1 Conceptual framework: definitions and a model 21

Part Two: TECHNOLOGICAL PERSPECTIVES AND STATUS QUO 65

2 Cyberscience: the new tools – the new working environment 67

3 Cyber-sciences – cyber-humanities – cyber-social-sciences 107

Part three: IMPACT ASSESSMENT 183

4 Cyberscience and the spatial dimension 185

5 Cyberscience and the distribution of roles in academia 235

6 Cyberscience and knowledge representation 257

7 Cyberscience and publishing 317

8 Cyberscience, quality control and crediting academic output 367

9 Cyberscience: economic and legal aspects 399

10 Cyberscience and the substance of research 439

Part four CONCLUSIONS 463

11 Cyberscience and politics 465

12 Overall conclusions 479

Annexes **A-l**

Abbreviations A-3

Annexes I-VI A-8

References A-41

Index A-67

DETAILED LIST OF CONTENTS

Foreword V

Content overview VII

Detailed list of contents VIII

List of tables XV

List of figures XVI

List of overviews XVII

Executive summary XIX

Part one: METHODS AND THEORY 1

0 Project outline and background 3
0.1 The status quo of research on cyberscience 4
0.1.1 Various research streams pertinent for the study of cyberscience 5
0.1.2 STS and cyberscience 7
0.1.3 TA and cyberscience 8
0.2 Innovative aspects of the project, aims and methodology 8
0.3 Outline of the empirical research 11
0.3.1 The selection of disciplines 11
0.3.2 The Internet enquiry 12
0.3.3 The interviews 13
0.3.3.1 In-depth interviews versus anonymous questionnaire 13
0.3.3.2 The interviewees 14
0.3.3.3 Practical aspects and transcription 15
0.3.4 Practical tests of E-tools 16
0.3.4.1 Interactive link collections 17
0.3.4.2 Hypertext writing 17
0.3.4.3 E-conferencing 18
0.3.4.4 Digital brainstorming at distance and chatting 19
0.3.4.5 E-publishing and E-archiving 19
0.4 Outline of the book 20

1 Conceptual framework: definitions and a model 21
1.1 What is cyberscience? 21
1.1.1 Cyberscience as an encompassing phenomenon 23
1.2 Modelling ICT impact on academia 26
1.2.1 The basic model 29
1.2.1.1 Information and communication technologies 30
1.2.2 Scholarly communication and its dimensions 31
1.2.2.1 Functions of the publication system 35
1.2.3 The diffusion of ICT in the scholarly communication system 37
1.2.3.1 Institutional factors 38
1.2.3.2 Technical and functional factors 41
1.2.3.3 Actor-related factors 43
1.2.3.4 The intervening factors at work 45
1.2.3.5 Qualitative trend extrapolation 47
1.2.4 Impact assessment 50
1.2.4.1 Changes of the scholarly communication system 52
1.2.4.2 General impact on academia 59
1.2.4.3 Impact on research substance 60
1.2.5 The complete model 63

Part Two: TECHNOLOGICAL PERSPECTIVES AND STATUS QUO 65

2 Cyberscience: the new tools – the new working environment 67
2.1 Basics 67
2.1.1 Computers: stand-alone and networked 67
2.1.2 Internetand WWW 68
2.1.2.1 The future of the Internet 71
2.2 Machine-to-machine communication 72
2.2.1 Distributed computing 72
2.2.2 Semi-autonomous information retrieval 73
2.2.2.1 On the path to the Semantic Web: meta-data, XML et al 73
2.2.2.2 Knowbots, infobots, software agents 75
2.2.2.3 Knowledge discovery: data mining, web mining, bibliomining 75
2.3 People-to-machine communication 75
2.3.1 Screen technology 76
2.3.2 Interactive electronic reading devices 77
2.3.3 Telework and ubiquitous computing 77
2.3.4 Databases: digital libraries, archives et al 78
2.3.4.1 Typology 78
2.3.4.2 Access to databases 81
2.3.4.3 Cross-linking 83
2.3.5 Web search-engines and directories 84
2.3.6 Web forms 84
2.3.7 Remote control et al 85
2.3.8 Speech recognition 85
2.4 People-to-people communication 86
2.4.1 E-mail 86
2.4.2 E-lists 87
2.4.3 Homepages et al 89
2.4.4 Academic E-publishing 91
2.4.4.1 Formats 91
2.4.4.2 Tools for editors and publishers 93
2.4.4.3 Tools for authors 94
2.4.4.4 Print-on-demand and electronic document delivery 95
2.4.5 E-conferencing 96
2.4.6 Content management systems 98
2.4.7 Groupware 99
2.4.8 E-teaching tools 100
2.4.9 Translation tools on the web 102
2.5 Archiving 103
2.6 Outlook 105

3 Cyber-sciences – cyber-humanities – cyber-social-sciences 107
3.1 Introduction 107
3.2 (Sub-)disciplinary case studies 109
3.2.1 Humanities and cultural studies 109
3.2.1.1 Language studies 110
3.2.1.2 Philosophy 113
3.2.1.3 History 114
3.2.1.4 Classical studies (papyrology) 116
3.2.1.5 Anthropology 117
3.2.2 Social sciences 118
3.2.2.1 Political science 118
3.2.2.2 Sociology 120
3.2.2.3 Economics 120
3.2.3 Natural sciences and applied/engineering sciences 122
3.2.3.1 Physics 122
3.2.3.2 Life sciences I: biology 123
3.2.3.3 Life sciences II: medicine 125

3.2.4 "Formal" sciences 126
3.2.4.1 Mathematics 126
3.2.4.2 Law 127
3.3 Cross-disciplinary comparison: the status quo 129
3.3.1 E-mail as standard communication channel 129
3.3.2 E-journals fighting for recognition 131
3.3.3 The early days of hyper/multimedia applications 135
3.3.4 E-conferencing only in exceptional cases 136
3.3.5 Importance of E-lists varies 137
3.3.6 E-(pre-)print servers established in some fields 139
3.3.7 Disciplinary databases as a standard tool 141
3.3.8 Digital libraries spreading 142
3.3.9 Groupware not yet well-known 143
3.3.10 Growing number of virtual institutes and extended research groups 144
3.3.11 Synopsis: how "cyber" is academia today? 145
3.4 Explaining the differences: a few hypotheses under scrutiny 148
3.4.1 General co-ordinates of disciplines 149
3.4.1.1 Size of the scientific community 149
3.4.1.2 Distribution of the relevant community around the globe 151
3.4.2 Functional explanations 152
3.4.2.1 "Embeddedness" in physical locales 152
3.4.2.2 Visual orientation 153
3.4.2.3 Data and model driven fields 155
3.4.2.4 Time pressure – rhythms – pace of discovery 156
3.4.2.5 Geographical focus of the subject 159
3.4.3 Economic factors 160
3.4.3.1 Closeness of the (sub-)discipline to economic applications 160
3.4.3.2 Competitiveness 162
3.4.3.3 Funding 163
3.4.4 The disciplinary culture 163
3.4.4.1 A general natural science versus social sciences/humanities cleavage? 164
3.4.4.2 Collaborative culture, multi-authorship and interconnectedness 168
3.4.4.3 "Cumulative" knowledge production 170
3.4.4.4 Publishing traditions in a discipline 172
3.4.4.5 Discipline-wide uniform method, style and paradigm 174
3.4.5 Agency 175
3.4.5.1 Cyber-entrepreneurs 175
3.4.5.2 Reputation 177
3.5 Summary 178

Part three: IMPACT ASSESSMENT 183

4 Cyberscience and the spatial dimension 185
4.1 Elements of a new spatial layout of academia 185
4.2 Key issues on the path to the new spatial layout of academia 189
4.2.1 On the suitability of E-mail for academic communication 189
4.2.2 Prospects of virtual seminars and conferences 192
4.2.2.1 Functions of scholarly conferences 192
4.2.2.2 Organisational and technical aspects of E-conferencing 194
4.2.3 Can multimedia eventually replace face-to-face? 196
4.2.4 Informal research activities: the importance of the "Cafe" 198
4.2.5 Establishing virtual contacts 201
4.3 Impact assessment of the new spatiality 202
4.3.1 Distant collaboration 203
4.3.1.1 Increase of collaboration 203
4.3.1.2 New collaboration patterns in cyberspace 205
4.3.1.3 More communication vs. increasing isolation 208

4.3.2 Enhanced efficiency 209
4.3.2.1 Productivity of the researchers 210
4.3.2.2 Productivity of the academic system 214
4.3.3 Written culture 217
4.3.4 Academic infrastructure in the future 220
4.3.4.1 The future of the university 220
4.3.4.2 New information infrastructure: digital libraries 222
4.3.4.3 Digital unity? Perspectives for peripheral research 224
4.3.5 The virtual re-constitution of scientific communities 227
4.3.5.1 Transforming invisible colleges 228
4.3.5.2 Building international scientific communities 229
4.3.5.3 Sub-field fragmentation? 229
4.3.5.4 Worldwide synchronisation/co-ordination of research? 231
4.4 On balance: De-materialisation of research? 232

5 Cyberscience and the distribution of roles in academia 235
5.1 Scholars: new skills, new roles 235
5.2 Researchers as teachers 238
5.3 From librarian to cybrarian 241
5.3.1 Information brokering, consultancy, information management 244
5.4 Role dynamics in the publishing sector 246
5.4.1 The age of outsourcing 247
5.4.2 Slow re-definition of institutional roles 247
5.5 The changing role of academic computer departments 249
5.6 Democratisation? Status and hierarchy in cyberscience 250
5.6.1 The case in favour of democratisation 251
5.6.2 The Matthew effect in contrast to democratisation 253
5.6.3 Empirical evidence and conclusions 254
5.7 Conclusions: the new distribution of roles in the age of cyberscience 255

6 Cyberscience and knowledge representation 257
6.1 Knowledge representation today 258
6.2 Cyberscience developments 259
6.2.1 Digitisation 260
6.2.2 Multimedia – hypermedia – new forms of visualisation 260
6.2.2.1 Examples of academic multimedia implementations 262
6.2.3 Hypertext 263
6.2.3.1 Modularity 264
6.2.3.2 Links 266
6.2.3.3 Paths 269
6.2.3.4 Examples of academic hypertexts 269
6.3 From text to knowledge base – one plus five scenarios 270
6.3.1 Scenario 1: "Layered" E-publications 272
6.3.2 Scenario 2: Field-wide thematic "hyperbases" 273
6.3.3 Scenario 3: Open disciplinary "hyperdiscussions" 274
6.3.4 Scenario 4: "Hyperbooks" 275
6.3.5 Scenario 5: Consolidated knowledge bases 276
6.4 Impact assessment of cyber-knowledge representation 277
6.4.1 Changing character and type of texts 278
6.4.1.1 Digital production 278
6.4.1.2 Length irrelevant? 279
6.4.1.3 Fluidity 280
6.4.1.4 Modularisation 286
6.4.1.5 Match of object and description 291
6.4.2 Effects on authors 292
6.4.2.1 New writing – new skills 292
6.4.2.2 Authorship blurred? 293
6.4.2.3 Merger of authoring, information seeking and analysis 295

6.4.3 Readability of digital media 295
6.4.3.1 Potential advantages of hypertext reading 297
6.4.3.2 Known problems of hypertext reading 298
6.4.3.3 New reading and "empowerment" of the reader? 301
6.4.4 Enhanced communication? 304
6.4.4.1 The communicative potential 304
6.4.4.2 Enhanced connectivity 306
6.4.4.3 New doors for the academic ivory tower 308
6.4.5 On balance 309
6.5 Analysis of factors influencing the development 310
6.5.1 Technical and functional factors 310
6.5.2 Institutional factors 312
6.5.3 Agency 313
6.5.4 The likelihood of the scenarios 313
6.6 Conclusions 314

7 Cyberscience and publishing 317
7.1 The status quo: P-publishing, E-publishing and hybrid forms 317
7.1.1 The increase of academic publishing, its reasons and its consequences 318
7.1.2 E-publishing today 319
7.2 What is new about E-publishing? 322
7.2.1 Speeding up publishing and dissemination 323
7.2.2 Advantages for searching 325
7.2.3 New modes of distribution and enhanced reach 326
7.2.4 New forms of academic publishing 327
7.2.4.1 Innovative journal formats 327
7.2.4.2 Special web publishing formats 330
7.2.4.3 Databases and archives 331
7.2.4.4 Scholarly "skywriting" 332
7.2.4.5 "Crossover publications" 333
7.2.5 Quoting E-publications 334
7.3 The end of scholarly publications on paper? 335
7.3.1 The future of print publications and their digital alternatives 337
7.3.1.1 The alternative technologies to paper 337
7.3.1.2 Academic publications are different 337
7.3.1.3 Academic books 338
7.3.1.4 Journals 345
7.3.1.5 Grey literature 348
7.3.2 Discussion 349
7.3.2.1 Technical and functional factors 349
7.3.2.2 Actor-related factors 351
7.3.2.3 Institutional factors 352
7.3.3 The P-to-E scenario matrix 356
7.3.4 Expectations within the academic community 358
7.4 Archiving scholarly E-publications 359
7.4.1 What is to be archived? 360
7.4.2 Responsibility 362
7.5 Assessment and outlook 365

8 Cyberscience, quality control and crediting academic output 367
8.1 The traditional quality control system 367
8.1.1 Quality criteria and prestige 368
8.1.2 Traditional quality control systems 369
8.2 Cyberscience developments and opportunities 371
8.2.1 New forms of ex ante quality control 371
8.2.1.1 Open peer review/commentary 371
8.2.1.2 Mixed systems 373
8.2.1.3 Differentiated refereeing standards 374
8.2.1.4 Selection of reviewers 375

8.2.2 New forms of ex post quality control 376
8.2.2.1 Online (ex post) commenting/annotation 376
8.2.2.2 Rating/scoring 378
8.2.2.3 New perspectives for scientometrics: access and citation counts, use tracking 379
8.2.3 Quality labelling, selective databases, levelled archives and the "control zone" 381
8.2.4 Miscellaneous new aspects 383
8.2.4.1 Speeding-up 384
8.2.4.2 Increase of non-reviewed publications 384
8.2.4.3 Obfuscation of origin? 384
8.2.4.4 Plagiarism and priority claims 385
8.2.5 Outlook 386
8.3 The path to a cyberscience certification system 386
8.3.1 Does E-publishing lead to less quality? 387
8.3.2 Re-establishing and reforming peer-review 389
8.3.3 Implementation issues 390
8.4 Academic credentials for online publications 392
8.4.1 Publication record and E-publications 393
8.4.2 Credentials for new forms of cyber-scholarliness 394
8.5 Conclusions 396

9 Cyberscience: economic and legal aspects 399
9.1 Economics of E-publishing 399
9.1.1 Cost comparisons between E- and P-publishing 399
9.1.1.1 E-journals 402
9.1.1.2 E-books 403
9.1.1.3 E-pre-print archives 404
9.1.1.4 Cybraries and long-term archives 404
9.1.2 Charging mechanisms for E-publications 406
9.1.3 (Re-)De-commodification of academic publishing? 407
9.1.3.1 Three phases of scholarly communication 408
9.1.3.2 The crisis of phase II 410
9.1.3.3 The developments towards phase III 411
9.1.3.4 De-commodification? 422
9.1.3.5 Phase III in the making 426
9.2 Legal aspects of academic cyber-publishing 427
9.2.1 Intellectual property law 428
9.2.1.1 General questions 430
9.2.1.2 Fair use and applicability of exemptions for libraries and academic research 431
9.2.1.3 IP law enforcement issues 434
9.2.2 Academic providers: link setting, liability and data protection 435
9.2.3 Legal rules for archiving E-publications 436
9.3 Conclusions 437

10 Cyberscience and the substance of research 439
10.1 Introductory remarks 439
10.2 Effects due to changes in methodology 441
10.2.1 Outcomes, otherwise impossible 441
10.2.2 Initial input changes 442
10.2.3 Impact on choice of topic 444
10.2.4 Creative potential of the new media 445
10.2.5 Favouring inter- and transdisciplinarity 447
10.3 Effects due to changes in work modes 449
10.3.1 Collaboration effects 449
10.3.2 Time effects 451
10.4 Effects due to representational changes 452
10.4.1 The shadow of the format 453
10.4.2 Standardisation 456

10.4.3 Preliminarily 457
10.4.4 Transparency 458
10.4.5 Connectivity/interconnectedness 459
10.5 Does ICT influence thinking and analysis? 460

Part four: CONCLUSIONS 463

11 Cyberscience and politics 465
11.1 The politics of cyberscience: arenas, actors and issues 465
11.2 Salient policy issues 467
11.2.1 Cyberscience infrastructure and access 467
11.2.1.1 Technical needs 467
11.2.1.2 Organisational needs 469
11.2.1.3 Universal access 470
11.2.2 E-publishing environment 471
11.2.2.1 Intellectual property 471
11.2.2.2 Integrity, quality control, credentialing and incentives 472
11.2.2.3 Archiving and digitisation 474
11.2.2.4 De-commodification 474
11.2.3 Information management 475
11.2.3.1 Systems 475
11.2.3.2 Training 476
11.3 Policy recommendations 477

12 Overall conclusions 479
12.1 The cyberscience scenario(s) 480
12.2 Meta-level conclusions 486
12.3 Coda 489

Annexes A-1

Abbreviations A-3
Annex I: Data-sets A-8
Annex II: Interviewees A-11
Annex III: Interview questionnaire and guide A-13
Annex IV: Note to the interviewees in advance of the interview A-18
Annex V: Definitions of key conceptual terms (glossary) A-19
Annex VI: CYBERLINKS database: linear, static transcript A-21

References A-41

Index A-67