

# Subject Index

- Aardvark, 24, 33, 34, 40
- Academic pages, 270
- Accuracy
  - definition of, 267
- Ad hoc retrieval, 36
- Advanced user models
  - novelty and diversity, 124–126
    - $\alpha$ -NDCG, 128
    - intent-aware ERR, 129
    - intent aware family, 127–128
    - subtopic recall and precision, 126–127
  - sessions, 129–130
    - expected global utility, 131–132
    - extending tradition to session measures, 132–133
    - normalized session discounted cumulative gain (nsDCG), 131
- Advocacy Web pages, 255
- Aesthetic theory, 298–300
- Alchemy, 156
- $\alpha$ -NDCG, 128
- Amazon, 36, 37, 232
- Amnesty International Web site, 296
- Anglo-American legal system, 290
- “Anomalous state of knowledge” (ASK), 22
- AOL, 232
- ARES (A Retrieval Engine Based on Sentiments), 150
- Aristotelian logic, 286
- Artworld, 299
- Ask, 232
- Average precision (AP), 113
  - interpretation of, 115–116
- Baidu, 175
- Bayesian mathematics, 297
- Bayesian methods, 297
- “Bell curve”, 219
- BibTip, 37
- Binary Preference (BPref), 217
  - definition of, 218
- Bing, 3, 61, 252, 256, 257
- BingMaps, 60, 64, 66
- Bing Streetside, 56
- Blekkio, 38, 39
- “Bundles”, 230
- Business Web pages, 255
- Cards, 233–234, 235–237
- Central limit theorem, 219
- Citation networks, 296
- Clewwa, 268
- Click popularity, 25
- Coefficients of probability, 287
- Collaborative filtering, 35–38
- Collaborative networks, 295
- Collaborative search system, 34–35
- Commodity market, 87–88
- Common law, 290
- Community pages, 270
- Computer-supported cooperative work (CSCW), 231
- Connotea, 22

- Corporate pages, 270
- CoSearch system, 35
- Country-code top level domain (ccTLD), 71
- Coverage, 167
- Cranfield paradigm, 106, 107
- Crawling, 297
- Creativity, 302
- Credibility assessments, influence of
  - search engine interfaces on, 251
  - alternative search engine interfaces, 262
  - Information Foraging Theory, 256–257
  - Information Quality and Credibility, 254
  - on Web, 259
  - Prominence-Interpretation Theory, 259–260
- Criteria-based judgments, 293
  
- Daily duplication frequency,
  - 178, 179
  - histograms, 179
  - patterns, 181–182
- Daily rank change frequency, 178, 180
  - histograms, 180
  - patterns, 183, 185
- Data type, 220
- Deduction works, 286
- Discounted cumulative gain (DCG), 119, 120, 152
- Discovery search, 54
- Diversity (Web search), 140–141
  - assessing, 145
  - comprehensive analysis of, 149–151
  - definition, 142
  - dimensions of, 144–146
  - evaluation measures for result, 151–153
  - future challenges, 159–160
  - in result sets and ranking, 147–149
- Diversity-aware search, medical web
  - content and, 153–158
  - architecture of search engine, 155–156
  - content collector, 156
  - content extractor, 156–157
  - diversity assessor, 157
  - relevance filter, 156
  - retrieval and ranking engine, 157
  - visualization component, 157–158
  - functionality of search engine, 153–154
- Document ranking, 212–213
- Dogear, 229
- Dragon Toolkit, 156
  
- eBay, 232
- Effectiveness evaluation, 106
  - advanced user modules of, 124–133
  - retrieval systems, 133–134
  - traditional measures of, 111–113
  - with user modules, 114–124
- English (British) law system, 290
- Entertainment Web pages, 255
- Escapism, 231
- Eurekster, 38
- Evaluation measures, traditional, 111
  - average precision, 113
  - precision and recall, 111–113
- Evaluation parameters, 169, 170
- Evolutionary epistemology, 300
- Evolutionary theory, 300
- Expected browsing utility (EBU), 122–124
- Expected global utility (EGU), 131–132

- Expected reciprocal rank (ERR), 121–122
- Expected search length (ESL), Cooper's, 114–115
- Explicit collaboration, 25
- Exploratory search, 54, 229
  
- Facebook, 32, 33–34, 232
- Faceted search, 150–151, 229
- Falsification, 285
- Feature parameters, 169, 170–172, 198–201
  - database, 171
  - home page, 171
  - keyword entry options, 171
  - result, 171
  - search options, 171
  - user preferences, 170
- Field trial, 238–240
- Flickr, 22, 52
- F-measure, 112, 152
- FOAF-format (friend of a friend), 23
- Forex (foreign exchange/currency market), 87, 89–90
- Foursquare, 71
  
- Game studies, 301
- General Web search, 65
- Geographic Information Retrieval (GIR), 7, 48, 50, 53
- Geographic Information Systems (GIS), 50–51
- Geospatial search, 49
- Goodman and Kruskal's G, 217
- Google, 2, 40, 164, 232, 252, 256, 257
- Google's PageRank, 25
- Google + 1-button, 40, 41
- Google Answers, 33
- Google Co-op, 38, 39
- Google Earth, 51
- Googlefight, 292
- Google Local, 51
- Google Maps, 51, 60, 61, 62, 65, 69
- Google News, 266
- Google Plus, 41
- Google Realtime, 24
- Google SearchWiki, 40, 41
- Google SERP interface, 274
- Google Sidewiki, 40
- Google Social Search, 40, 41
- Google Streetview, 56
- Google Trends, 83
- Graded relevance, 116–117
- Graphical user interface (GUI), 230, 235
  
- Hakia, 268
- Hermann's theory, 300
- Histogram patterns
  - benchmark, 188
  - and classification, 180–182
  - expected value, 187–190
  - parameters for, 187
  - validation of, 190–193
- Human-computer interaction (HCI), 230
- Human motivations, 247
- Hunch, 37, 38
- Hypertext 2.0*, 301
  
- Implicit collaboration, 25
- Index as an aesthetic marker of presence, 298–299
- Induction works, 286
- Informational queries, 142
- Informational Web pages, 255
- Information Foraging Theory, 256–257
- Information networks, 296
- Information pages, 270
- Information quality
  - and credibility, 254
  - definitions of, 254
  - expertise, 254

- heterogeneity of information sources, 254–256
- trustworthiness, 254
- dimensions, 254
- multidimensional construct, 255
- Information Retrieval (IR)
  - community, 3, 4
  - history of, 21–22
- Information retrieval (IR) system, 106, 166
  - classical model of, 229
- Information retrieval, 231
- Internet, 80
- Internet Engineering Task Force (IETF), 166
- Internet Movie Database, 295
- Internet scholarship, 288
- IR, *see* Information retrieval
- iSearch, 23
  
- “Jerry and David’s Guide to the World Wide Web,” 26–27
- Judgmental truth claims, 292–294
  
- Kartoo, 262, 263f
- Kendall’s  $\tau$ , 210
  - definition of, 210
- Knowledge networks, *see* Information networks
- Kurrently, 24
  
- “Laboratory Model of Information Retrieval,” 22
- Last.fm, 37
- Latent Semantic Analysis (LSA), 269
- Law, truth claim in, 290–292
- Legal standards of proof, 291
- Library and Information Science (LIS), 3, 4
- LifeLines, 229
- Link topology, 25
  
- Local search engines, 7, 48–50, 71–72
  - current, 55
  - as entity search, 59–60
  - further research on, 74
  - history of, 50–51
  - impact, 72–73
  - map-based, 58–65
  - usability of, 56
- Location-based services (LBS), 50, 51, 52
- Logical empiricism, 285
- Logical positivism, 285
  
- Manual indexing, 26, 27
- Map and search engine, 55–56
- Map-based local web search, 58–65
- MapQuest, 232
- Market of search engine, 2–3
- “Mash-up”, 233
- Mean average precision (MAP), 214
- MedSearch, 150
- Mendeley, 37
- Metaphor, 233, 234
- Microsoft, 232
- Middlespot, 262
- Mobile local search, 73–74
- MovieLens, 37
- MSN, 257
  
- Navigational queries, 142
- NetBeans Integrated Development Environment, 98
- NetBeans Platform, 96, 97, 98
- “Network of networks”, 294
- Networks
  - types and behaviors of, 294–296
- News Web pages, 255
- Newton, Isaac, 285
- Nexplore, 262
- Nonprofit pages, 271
- “Normal distribution”, 219

- Normalized discounted cumulative gain (NDCG), 118–121
- Normalized Discounted Cumulative Gain (NDCG), 219
- Normalized session discounted cumulative gain (nsDCG), 131
- “Normal science”, 285
- Null hypothesis, 287
  
- Objectivity
  - definition of, 267
- Ockham’s razor, 287
- Offline evaluation, 133, 134
- 123people, 23
- Online evaluation, 133, 134
- OpenStreetMap, 71–72
- Operational mechanism, 297–298
  
- PageRank, 253, 266, 267
- Pearson correlation coefficient, 215
- Pebbles, 234, 237, 238
- Peer review, 294
- People powered search engines, 23
- Performance measurement metrics, 165, 177–180
  - frequency measures, 179–180
  - stability, 178
  - timeliness, 177
- Performance parameters, 165, 169, 172–173, 202
  - quality of results, 172
  - response time, 172
  - total number of results, 172
- Performativity, 301–303
- Period duplication frequency, 178, 179
  - patterns, 183–186
- Personalised search engine
  - evaluation, 163–193
- Personalized social search engines, 38–40
- Personal pages, 271
  
- Personal Web pages, 255
- Philosophic perspectives, on nature of search, 298
- Pipl, 23
- Place Search, 69
- Popper’s theories, 285–286, 300
- Precision @  $N$ , 173, 174
- Precision, 166, 173
- Precision and recall as evaluation measures, 111–113, 152
- Preference networks, 295
- Probabilities of occurrences, 287
- Prominence-Interpretation Theory, 259–260, 272
- Psychological basis of search, 300
  
- Quality metrics, 166, 169
- Quantitative versus qualitative research methodologies, 289t
- Qype, 71
  
- Random networks, 296
- Rank-biased precision (RBP) models, 117–118
- Rank correlation and search engines, 203
  - alternative RCCs, 217–219
  - discussion, 215–217
  - document ranking, 212–213
  - examples, 207, 208–209
  - investigating, 219–220
  - motivations and concepts, 205–211
  - system architecture, 214–215
  - system ranking, 213–214
- Rank correlation coefficient (RCC), 204
  - linearity, 220
  - ties, 220
- RankSpeed, 150
- “Reasonable man” test, 292
- Recall, 166, 173
- Recall @  $M$ , 173

- Reciprocal rank (RR), 121–122
- Recommender networks, 295
- Recovery search, 54
- Redz, 262
- Relevance, 166
  - evaluation, 173–174
  - feedback, 21
  - metrics, 173, 174
  - scoring algorithms, 169
- Research inquiries, 283–284
- Research on search engines, 4–6
- Research tradition, 283
- Result diversification, 147
- Retrieval effectiveness, 165, 167, 168
- RMS (Root mean squared errors), 213
- Rollyo, 38, 39
  
- Scale-free network, 296
- Science, truth claim in, 285–287
- Scientific method, 285, 286–287
- SDART Ltd, 94
- Search Cube, 262
- Search engine comparison
  - case study, 176–177
  - data collection methodology, 175
  - qualitative evaluation, 175–176
  - quantitative evaluation, 175–176
- Search engine evaluation criteria, 165–168
  - qualitative evaluation, 165
  - quantitative evaluation, 165
  - user-oriented, 167–168
- Search engine evaluation models, 168–174
  - feature parameters, 170–172
  - performance parameters, 172–173
  - quality issues, 174
  - relevance evaluation, 173–174
  - weighted parameters and summary score, 169–170
- Search engine interfaces, influence
  - of, on credibility assessments, 251
- Search engine mechanisms, 297–298
- Search engine optimization (SEO), 253
- Search engine results pages (SERPs), 7, 204, 253
- Search index metric, 168
- Searchrolls, 39
- Search strategy, 168
- SearchTogether, 35
- SERPs, Quality-Related Cues on, 265–269
- Session test collection, 130
- 70 Features and performance parameters, 169
  - Six degrees of separation*, 295
- Slashtags, 39
- “Small world phenomenon,” 23
- Social collaboration, 25
- Social data, 23
- Social graph, 22
- Social media search engines, 24
- Socialmention, 24
- Social networks, 295
- SocialPageRank, 29
- Social question-answering (QA) community-based systems, 32
  - examples of, 31t
  - expert-based systems, 32–34
  - fundamentals of, 29–30
- Social science
  - triangulation in, 288
  - truth claim in, 288–289
- Social search, 20
  - definition, 21, 23
  - taxonomy of, 26f
    - see also* Search engines
- SocialSlimRank, 29
- Social software, 24
- Social tagging, 26
  - definition, 27

- direct usage of, 28
  - indexing and, 27–28
  - indirect usage of, 29
- SparTag, 229
- “Spatial Search and Browsing,” 51
- Spearman’s footrule, 178
- Spearman’s rank correlation, 210
- Stock market, 87, 88–89
- System architecture, 214–215
- System ranking, 213–214
  
- Targets, absence of, 247
- Taste graph, 37
- Technological networks, 295
- Technology probe investigation, 227
  - field trial, 238–240
  - findings, 240
    - confounding properties of the experience, 240–242
    - grasping the possibilities of pebbles, 242–245
    - reflections on cards, 245–246
  - implications for future research, 246–247
  - web interaction probes,
    - designing, 233–235
    - cards, 235–237
    - pebbles, 237, 238
  - web use, typologies of, 228–233
- Technorati, 24
- Test collection methodology, 107
- Text REtrieval Conference (TREC), 107–108, 153, 213
  - document collection, assembling, 108
  - documents relevant to topics,
    - identifying, 108, 109–110
  - information needs, creating, 108
- Theory of learning, 300
- Toponym extraction, 67
- Transactional queries, 141–142
- TREC Web Track, 159
- Truth claim, 282
  - in law, 290–292
  - in science, 285–287
  - in social science, 288–289
  - judgmental, 292–294
- Twitter, 32
- Typologies, of web use, 228–233
  
- Uclue, 33
- UMLS, 156, 157
- Universal search results, 68
- Usenet, 94, 95
- User-generated content (UGC)
  - tools, 231
- User models, effectiveness
  - evaluation with, 114
    - Cooper’s expected search length, 114–115
    - expected browsing utility (EBU), 122–124
    - expected reciprocal rank (RR), 121–122
    - graded relevance, 116–117
    - normalized discounted cumulative gain (NDCG), 118–121
    - rank-biased precision (RBP), 117–118
    - Robertson’s interpretation of average precision, 115–116
      - see also* Advanced user models
- User-oriented evaluation, 167–168
- Uses and gratifications theory (U&G), 230–231, 232
- UX paradigm, 230
  
- Viewzi, 263, 264f
  
- Web-based hubs, 296
- Web browsing, 231
- Web interaction probes, designing, 233–235
  - cards, 235–237

- pebbles, 237, 238
- Web page
  - publication of, 255
- Webpage, 205
- Web pages, 271
- WebPerceiver, 94–96
  - additional modules of, 99–100
  - future developments, 99
  - package diagram for, 98
  - use case for, 99
- Web search, 283, 284
- Web search engines statistics
  - analysis, 81
- Web search queries statistics, 82
  - advantages of, 82–83
  - associative analysis of, 93–94
  - computational processing and, 91–93
  - examples of, 87–90
  - issues of, 84
  - prognosis mechanism of, 85–87
- Web tools, 193
- Web trends, manual gathering and, 96
- Web use, typologies of, 228–233
- Webworld, 299
- Weighted Kendall's  $\tau_{c-w}$ , 218
  - definition of, 218
- Whostalking, 24
- Wiki Answers, 32
- Wikipedia, 295
- Wink, 23
- WordNet, 158
- World Wide Web (WWW), 252, 296
- World Wide Web Consortium (W3C), 166
- Yahoo, 3, 26, 175, 232, 252, 256, 257
- Yahoo! Answers, 32, 33
- Yahoo! FireEagle, 71
- Yahoo! Maps, 60, 63, 66
- Yahoo Search BOSS, 38, 39
- Yasni, 23
- Yellow pages, 52, 57–58
- Yelp, 71
- YouTube, 22, 232