

## Index

---

- A/B testing, 73, 213
- abstract syntax tree (AST), 23, 90, 143, 153
- acceptable ads, 227, 232
- accounts, 86
- accuracy, 149
- ACLU, 11, 259
- active learning, 159
- ad blockers, 225
  - browser extensions, 226
  - CSS filtering, 227
  - DNS blocking, 225
  - EFF Privacy Badger, 261
  - effectiveness, 229
  - efficiency, 229
  - filter lists, 110, 138, 139, 227, 233, 235, 276
    - effectiveness, 230
  - interception proxy, 226
  - prevalence, 231
  - users of, 140
- ad campaigns
  - performance of, 93
- ad explanations, 53, 93, 195
- ad impressions, 93, 191, 194
  - detection of, 152
- ad preference managers, 40, 52, 195
- ad requests, 196, 231
- AdaBoost, 154
- AdChoices, 227
- ads
  - detection of, 138
  - political, 139
  - relevance, 192
  - topic detection, 152, 153
  - user exposure to, 95, 189, 274
  - visual properties, 92
- ads.txt, 52, 193
- advertising, 45
  - ad clicks, 93
  - ad creative, 53, 81, 138
  - ad delivery, 191
  - ad exchanges, 46, 193
  - ad networks, 46
  - ad platforms, 266
  - ad traffic, 139, 188
  - adversarial, 198
  - as a service, 198
  - demand-side platforms (DSP), 46, 49, 193
  - ecosystem, 188
  - history of, 45
  - information flows, 273
  - libraries, 53, 196, 219
  - malvertising, 47, 139, 189
  - mobile, 53, 196
  - pricing, 194
    - cost-per-click (CPC), 48
    - cost-per-mille (CPM), 48, 194
    - direct buy, 45
  - real-time bidding (RTB), *see* real-time bidding (RTB)
  - retargeting, 47, 140, 193
  - revenue, 4, 231
  - supply-side platforms (SSP), 46, 49
  - targeting, 47, 53, 80, 92, 190, 273
    - attributes, 92, 191, 195, 196
    - behavioral, 47, 194
    - contextual, 47
    - custom audiences, 47
    - detection of, 140
    - lookalike, 48
    - PII, 47, 190
    - prevalence, 190

advertising (cont.)  
   profile-based, 47, 140  
   value of categories, 194  
 Alexa list, *see* toplists  
 algorithmic pricing, 213  
 Amazon Mechanical Turk, 119  
 analytics, 44, 133, 188  
   libraries, 44, 188  
   on mobile platforms, 44  
 Android, 25  
   Android Package format (APK), 25  
 AndroZoo, 111  
 anonymity set size, 127, 181  
 anti-ad blockers, 232  
   detection of, 153, 234  
   filter rules, 234  
   prevalence, 234  
   undesirable effects, 233  
 Anti-Adblock Killer List, 110, 235  
 API calls, 97, 161  
   obfuscated, 269  
 app stores, 81, 219  
   characteristics, 219, 275  
 apps, *see* mobile apps  
 audience size estimate, 93  
 audit rights, 248  
 audit study, 65  
 auditability, 248  
 autocomplete suggestions, 95, 207, 209  
 automation, 267  
 autonomy, 9, 12, 39  
  
 background traffic, 117  
 bag-of-words, 150  
 Bayesian classifier, 154  
 betweenness centrality, 189  
 bi-grams, 151  
 bias, 271  
   location, 208  
   political, 208, 211  
   ranking, 99  
 black-box system, 65  
 black-hat search engine optimization, *see*  
   cloaking techniques  
 blocking (statistics), 170, 171  
 Bluetooth Low Energy (BLE), 34, 278  
 browser extensions, 24, 77, 91, 114, 121, 215  
   extension API, 24  
   list of, 181  
   security, 24  
 browsers, 77, 91, 180

browsing behavior, realistic, 113  
 browsing history, 157  
 business model, 5, 7, 232  
  
 call chain, 163  
 captcha, 87  
 captive portals, 20, 179  
 Cascading Style Sheets (CSS), 21  
 causal effects, 169  
   experimental studies, 170  
   observational studies, 171  
 causal inference, 68, 268  
 CERT, 263  
 certificate pinning, 117  
 chilling effect, 10  
 civil rights, *see* human rights  
 classification, 148  
 click interception, 216  
 click traces, 44  
 clickjacking, 216  
 cloaking techniques, 156, 216  
 cloud services, 98, 222  
 clustering, 148  
 code audit, 66  
 codes of ethics, 102  
 collaboration, 259, 261  
 collateral data, 3  
 Common Crawl, 109  
 Computer Fraud and Abuse Act (CFAA),  
   67, 106  
 concentration of power, 126  
 confidence intervals, 168  
 confounding factors, 171  
 consent, 94, 275  
 consent management providers, 203, 265  
 contact forms, 216  
 Content Security Policy (CSP), 18  
 contextual integrity, 105  
 control flow graph, 98, 161, 163, 197  
 control tasks, 120  
 control treatment, 68  
 control variables, 171  
 convenience sample, 83, 121  
 convolutional neural network (CNN), 153, 158  
 cookie consent libraries, 94, 203  
 cookie linking graph, 88  
 cookie matching, *see* cookie synchronization  
 cookie notices, 94, 203  
   detection of, 144  
 cookie synchronization, 30, 180, 186, 193  
   detection of, 135, 151

- cookies, 28, 235
  - first-party, 29
  - Flash, 31
  - HTML5, 32
  - identifying, 89, 133
  - respawning deleted cookies, 32
  - same-origin policy, 30
  - third-party, 29
- corporate ownership, 60, 180, 271
- corporate surveillance, 3
  - characteristics of, 60
  - ecosystem, 5
  - effectiveness of, 62
  - effects of, 62
  - evolution of, 3
  - extent of, 61
  - perception of, 64
  - purpose of, 62
  - undesirable effects of, 9
- cosine similarity, 130
- countermeasures, 63, 225, 249, 276
  - ad blockers, *see* ad blockers
  - effectiveness of, 63
  - fingerprinting, *see* fingerprinting blockers
  - for IoT devices, 280
  - for mobile devices, 243, 276
  - obfuscation-based, 240
- crawls, 113
  - efficiency, 113
  - synchronized, 114
- CrowdFlower, 119
- crowdsourced audits, 68
- crowdsourcing, 118, 270
  - compensation, 121
  - fair compensation, 118
  - platforms, 68
  - quality control, 120
  - screening, 120
  - tasks, 120
- Curlie, 156
- custom audiences, 80
  
- Damerau–Levenshtein distance, 131
- dark patterns, 214, 267
  - mobile apps, 275
- data brokers, 43, 195
- data collection, 269
  - active, 112, 170
  - application layer, 112
  - automated, 114
  - from mobile apps, 118
  - mobile traffic, 116
  - parallel, 171
  - passive, 111, 171, 232
- data management platform (DMP), 49
- data minimization, 104
- data preservation, 123
- data protection
  - complaint, 260
- data storage, 123
- decision tree, 152, 236
- demographics, 84, 93, 99, 223
  - detection of, 144
- difference-in-differences method, 172
- differential analysis, 142
- diffusion of information, *see* information flows
- digital hegemony, 208
- discrimination, 39, 129, 191, 258, 271
  - ad audiences, 192
  - direct, 129
  - indirect, 129
  - price, 212
- disparate impact, 129
- disparate treatment, 129
- disparity, 129
- distributed measurement, 72
- diversification, 185
- DNS, 19
  - load balancing, 72
  - requests, 113
  - resolvers, 113
  - traffic, 278
- do not track, 94
- Document Object Model (DOM), 21, 90
- domain classification, 76, 135, 139, 155, 271
- Domain Name System, *see* DNS
- domain owner, 19, 89
  - detection of, 132, 271
- domains
  - domain names, 19
  - free live streaming, 156
  - parked, 155, 217
- DoubleClick, 180, 189
- dynamic analysis, 161
- dynamic instrumentation, 116
  
- EasyList, 110, 134, 227, 235
- EasyPrivacy, 110, 134
- ecological validity, 105
- economic impact, 231, 232, 274
- economic incentives, 191
- Electronic Frontier Foundation (EFF), 261

emotional contagion, 209  
 ensemble classifier, 154  
 entropy, 127, 181  
 ePrivacy directive, 185, 260  
 ethical considerations, 101  
   codes of ethics, 102  
   research guidelines, 102  
 execution trace, 90, 163  
 experiment design, 264  
   fractional factorial, 69, 265  
   latin hypercube, 265  
   latin square, 69  
   randomized block, 69  
 experimental factor, 68, 74  
 experimental treatment, 68, 166  
 experimental unit, 68  
 experiments, 59, 267  
   quasi, 171  
   semi-automated, 114  
 explanatory variable, 169  
 exploratory data analysis, 166  
 externalities, 97, 223, 224

F1 score, 149  
 fair compensation, 118  
 feature engineering, 148  
 Federal Trade Commission (FTC), 259  
 feedback mechanisms, 63  
 fingerprintability, 181  
 fingerprinting, 33, 92, 273  
   accelerometer, 34  
   attributes, 35  
   BatteryStatus API, 35  
   Bluetooth Low Energy (BLE), 34  
   browser, 35  
   canvas, 36, 136, 181  
   canvas font, 36, 136, 181  
   clock skew, 34  
   countermeasures, 92, 136, *see also*  
     fingerprinting blockers  
   detection of, 136  
   device, 34  
   effectiveness, 183  
   evolution, 182  
   fingerprint2.js, 36  
   gyroscope, 34  
   HTML5 AudioContext, 37  
   inconsistent attributes, 137, 239  
   IoT devices, 278  
   known scripts, 240  
   linking of fingerprints, 152

magnetometer, 34  
 microphone, 34  
 mobile app list, 243  
 Panoptick, 35  
 prevalence, 181  
 speaker, 34  
 wireless network interface cards (NICs), 34

fingerprinting blockers, 237  
   anti-fingerprinting blockers, 239  
   client-side diversification, 239  
   effectiveness, 239  
   randomization, 237  
   standardization, 238, 239

Flesch Reading Ease, 94  
 Flesch Reading Ease (FRE), 204  
 fragmentation, 268  
 free live streaming domain, 156, 216  
 Friedman test, 167  
 function call graph, 163  
 functionality, 227, 237, 271  
   page reload rate, 237

GDPR, 103, 203, 259, 260  
   influence on tracking, 185  
 gender, 215  
   detection of, 145  
 geo-blocking, *see* server-side blocking  
 geolocation, 81, 97, 203, 265  
   influence on tracking, 184  
 goodness of fit test, 165  
 Google Play, 25, 219  
 group fairness, 129, 215

harmful effects  
   limiting, 107  
   on individuals, 103  
   on organizations, 106

header bidding, 51, 92, 180, 194  
   detection of, 141

Herfindahl–Hirschman Index (HHI), 126, 184  
 heuristics, 132, 228, 271  
 Hidden Markov Model (HMM), 157  
 Holdout method, 148  
 Holm–Bonferroni correction, 165, 166, 168  
 homogeneity test, 166  
 HTML, 21  
   Ambient Light Sensor API, 261  
   BatteryStatus API, 35  
   canvas, 21  
   nested elements, 138

- HTTP, 16
  - headers, 17
    - Cookie, 17, 29, 88
    - DNT, 17
    - ETag, 18
    - Last-Modified, 18
    - Location, 18, 88
    - Referer, 17, 88
    - Set-Cookie, 18, 29
    - Tk (tracking status), 18
    - User-Agent, 17
  - redirect, 30
  - requests, 17, 111, 177, 197, 202
    - detection of PII in, 154
  - responses, 18, 111, 177, 197, 202
  - state management, 29
  - status codes, 18, 88
- HTTP Strict Transport Security (HSTS), 18
- HTTPS, 16, 18, 112, 227, 269
- human rights, 107, 247
  - equality, 10
  - freedom of expression, 10
  - nondiscrimination, 10
  - privacy, 9
  - read anonymously, 10
  - risks, 9
- Hypertext Markup Language, *see* HTML
- Hypertext Transfer Protocol, *see* HTTP
- hypothesis test, 164, 168
  - multiple hypotheses, 165
- IAB (Interactive Advertising Bureau), 50, 273
- ICO, 260
- identifiability, 128
- impact, 247
  - pathways, 249
  - planning for, 249
  - types of, 250
- inclusion graph, 90
- inclusion tree, 90, 114
- independent samples, 167
- individual fairness, 129, 215
- inducing interests, 85
- information diet, 211
- information flows, 273
- information leakage
  - advertisers, 197
  - browser extensions, 163
  - browsing history, 157
  - mobile apps, 220
  - obfuscated, 142
  - PII, 216, 274, 277
- informed consent, 104
- input variables, 68, 265
- institutional review board (IRB), 101
- interaction, 270
  - automated, 162, 270
  - with IoT devices, 280
- interdependent privacy, 272
- International Mobile Equipment Identity (IMEI), 26, 99
- International Mobile Subscriber Identity (IMSI), 99
- Internet Archive, 108
- internet of things (IoT), 276, 280
- internet protocol suite, 14
- Internet service provider (ISP), 15
- iOS, 26
  - permissions, 26
- IP address, 15, 33, 99, 184
  - mapping to geolocation, 15
  - residential, 79
- Jaccard index, 131
- JavaScript, 21, 23, 90, 150, 177, 235, 272
  - API calls, 90, 114
  - detection of similar, 143
  - engine, 23, 163
  - obfuscated, 143, 234
  - runtime objects, 163
- k-fold cross validation, 149
- Kendall Tau distance, 131
- Kolmogorov–Smirnov test, 166
- Kruskal–Wallis test, 167
- latency, 92, 194
- legal implications, 258
- legislative process, 256–258
- link preview, 212
- local VPN, 116
- localness, 208
- location, *see* geolocation
- location bias, 208
- longitudinal studies, 70, 268
- MAC address, 15, 99
- machine learning, 148, 228, 271, 272
  - hardening against, 272
  - retraining, 272

malicious websites, 77  
     classifiers for, 156  
 manifest, 25  
 Mann-Whitney test, 167  
 mapping, 218  
 market concentration, 184  
 matching methods, 171  
 maximum entropy algorithm, 160  
 McNemar's test, 167  
 measurement  
     based on ads, 198  
     distributed, 72  
     lock-step, 73  
 measures, 125, 270  
     combinations of, 128  
 media coverage, 247, 254  
 media extraction, 142  
 message box, 255  
 minimum wage, 119  
 mitmproxy, 114, 116  
 mobile advertising identifier (MAID), 26, 99  
 mobile apps, 81, 122  
     automation of, 116  
     characteristics, 276  
     cloned, 219  
     comparison with websites, 221  
     detection of owner, 133  
     dynamic analysis, 161  
     exclusion, 82  
     free vs. paid, 196, 220  
     historical versions, 111  
     incentivized installs, 221  
     instrumenting, 116  
     malicious, 219  
     detection of, 160  
     obfuscated libraries, 147  
     privacy leaks, 220  
     purpose of network access, 160  
     sensitive, 243  
     social libraries, 219  
     static analysis, 161  
     updates, 219  
 monkey tools, 116  
 monopoly, 8  
 multi-party privacy, 272  
 Naive Bayes classifier, 153  
     multinomial, 156  
 names, 80  
 natural language processing, 149  
     generative, 154

Needleman–Wunsch algorithm, 143  
 NetFlow, 15, 84  
 network traffic  
     detection of PII in, 162  
 news curation, 207  
 noise factors, 68, 72, 117, 267  
 nondeterminism, 268, 269  
 nongovernment organizations, 258  
 noninvasive user audit, 67  
 normalized entropy, 128  
 noyb, 259  
 null hypothesis, 68, 164, 165, 168  
 Oaxaca–Blinder decomposition, 169  
 obfuscation  
     API calls, 269  
     as a countermeasure, 240  
     de-obfuscation, 142  
     deterministic, 142  
     effectiveness, 241  
     for advertising, 241  
     JavaScript, 143  
     of mobile app code, 98, 147  
     search terms, 241  
     traffic, 241  
 one-sample test, 165  
 online status indicators, 96  
 Open Observatory of Network Interference (OONI), 109  
 openRTB, 50  
 OpenWPM, 109, 123, 261  
 opt out, 43, 94, 205, 237  
     detection of, 159  
 overlap coefficient, 131  
 p-value, 164  
 paired test, 165  
 parallel bidding, 51  
 parent-subsidiary relationships,  
     *see* corporate ownership  
 parked domain, 155, 217  
 participants  
     human, 104  
     paid, 119  
     recruitment, 121  
 paywalls, 217–218  
 Pearson's chi-squared test, 165  
 penetration, 127  
 permissions, 25, 53, 82, 97, 118, 160, 196,  
     219, 220  
     dangerous, 26, 196

- iOS, 26
- over-privileged, 219
- quarantine requests, 243
- permutation test, 170
- personalization, 207, 209, 213, 218
- personally identifiable information (PII), 142
- plug-ins, 24
  - Flash, 24
- polycymaking process, 256–258
- political action committees (PACs), 258
- political bias, 208, 211
- political leaning, 145
- political mobilization, 210
- Porter–Stemmer algorithm, 149
- precision, 149
- press release, 254
- prevalence, 125, 177
- price discrimination, 212
- price parity clauses, 214
- prices, 96, 212
  - algorithmic, 213
- Principal Component Analysis (PCA), 143, 160
- privacy by design, 102
- Privacy International, 261
- privacy policies, 94, 110, 203, 260, 275
  - classifiers for, 158
  - compliance, 205
  - contradictory statements, 204
  - detection of, 144, 158
  - disclosure of data flows, 204
  - flow-to-policy consistency, 275
  - labeling of practices, 159
  - linguistic properties, 94, 204
  - mobile apps, 220
  - opt out, *see* opt out
  - reading ease, 94, 204
- privacy regulatory model, 185, 258
  - co-regulatory, 185
  - comprehensive, 185
  - sectoral, 185
- privacy settings, 99
- product taxonomy, 77
- profile contamination, 73, 85, 87
- profiling, 39, 186
  - anonymity, 42
  - attributes, 40, 93, 187
    - accuracy of, 187
    - sensitive, 153, 190, 192
  - big-5 personality traits, 41
  - demographics, 40
  - emotional state, 41
  - singling-out, 42, 283
- Prolific Academic, 119
- prominence, 126
- propensity score, 171
- ProPrivacy, 261
- protected characteristics, 10
- public good, 107
- public suffix list, 89
- public-private partnerships, 282
- publication of code, 248
  
- quantitative measures, *see* measures
  
- radicalization pathway, 211
- random forest, 152, 153, 228
- random sampling, 171
- randomization, 170
- randomized controlled trials (RCT), 65, 68
- rank-biased overlap, 132
- ranking bias, 99
- rankings, 214
- Ratcliff-Obershelp algorithm, 134
- ratings, 96, 214
- reidentification risk, 42, 187
- reach, 126, 177, 180, 186
- reachability analysis, 161, 163, 197
- readability, *see* Flesch Reading Ease (FRE)
- real-time bidding (RTB), 48, 192, 194
  - detection of, 141
  - generalized second-price auctions, 48
  - prices, 152
  - Vickrey-Clarke-Groves auctions, 48
  - waterfall model, 51
- recall, 149
- recommendation algorithms, 211
- regression, 148, 168
  - coefficients, 168
  - linear, 168
  - logistic, 171
  - negative binomial, 169
  - propensity-score stratified, 172
  - rank-ordered logit, 169
- regulations, 63, 184
  - competition, 184, 249
  - data protection, 103, 184, 249
  - enforcement, 249
  - evidence of violation, 275
- regulatory bodies, 259
- regulatory environment, 265
- replication, 69, 120, 167, 269

328

representation ratio, 129  
 representative samples, 83, 119, 121  
 reproducibility, 115, 269  
 request tree, 88, 139  
 research questions, 59–64  
 residential internet connection, 105  
 response variables, 68, 87  
 responsible disclosure, 108, 247, 248, 262  
   process, 263  
 reverse DNS (rDNS), 19  
 ride-sharing, 97, 222  
   availability, 223  
 route complexity, 224  
 route planning, 97, 224  
  
 same-origin policy, 30  
 scale, 59, 266  
 scope, 59, 70, 266  
 scraping audit, 67  
 search  
   biased rankings, 209  
   localness, 208  
   personalization, 207  
   result rankings, 10, 79, 95  
   terms, 78, 275  
 search engine manipulation effect, 209  
 search engine result pages (SERPs), 95, 205  
 search engines, 205  
   autocomplete suggestions, 207  
   for source code, 110  
   on mobile devices, 276  
   résumés, 215  
   text snippet algorithm, 205  
   top stories algorithm, 206  
   user-generated content, 206  
 server name indication (SNI), 16  
 server-side blocking, 200  
   detection of, 146  
   geo-blocking, 201  
   prevalence, 200  
   reasons for, 200  
 significance level, 164, 169  
 smart cities, 280  
   accessibility, 283  
   applications, 280  
   characteristics, 283  
   technologies, 280  
 smart speakers, 278  
   adversarial activation, 279  
   misactivation, 279  
   skills, 279

Index

smart TVs  
   channels, 82, 277  
   countermeasures, 279  
   detection of, 278  
 SMOTE, 160  
 social media strategy, 256  
 social networks, 209  
   accounts, 86, 267  
   friend lists, 95  
   news feed, 95, 211  
   sockpuppet audit, 67  
 Spearman's Rank Correlation, 213  
 SQLite, 123  
 stakeholders, 251, 255  
 standardization bodies, 261  
 static analysis, 161  
 statistics, 164, *see also* hypothesis test  
   assumptions, 72  
   central tendency, 164  
   descriptive, 164  
   dispersion, 164  
   spatial analyses, 172  
 stop words, 160  
 stratified sampling, 75  
 study designs  
   code audit, 66  
   crowdsourced audit, 68  
   longitudinal, 70  
   noninvasive user audit, 67  
   scraping audit, 67  
   sockpuppet audit, 67  
 Subject access requests (SAR), 205  
 support vector machine (SVM), 150  
 surveys, 67  
 synthetic data, 106  
 synthetic user profiles, *see* virtual personas  
 system under study, 65  
  
 tagging services, *see* domain classification  
 taint tracking, 161  
 term frequency–inverse document frequency,  
   *see* tf-idf  
 terms of service, 67, 102, 107  
 test statistic, 164, 170  
 tf-idf, 130, 143, 150, 159  
 The Conversation, 254  
 third-party content inclusion, 157  
 third-party graph, 89  
 third-party libraries, 97, 220  
   detection of, 147  
   release dates, 147



- updating, 220
- versions, 147
- timestamps, 99
- TLS, 114
  - certificate pinning, 117, 280
  - certificates, 89, 133
  - handshake, 16
  - server name indication, 16
  - session resumption, 16, 89
- topic modeling, 155
- toplists
  - Alexa, 75
  - Hispar, 76
  - Tranco, 75
- Tor, 77, 110, 146, 201, 238, 240
- traceroute, 88, 146
- tracker blockers, 235
  - combinations of, 240
  - effectiveness, 235
  - on mobile devices, 236
- tracking, 177
  - cache-based, 32
  - classifiers for, 150
  - cross-device, 37, 137, 183, 273
  - detection of, 132, 134
  - email, 39, 137, 184, 273
  - geolocation, 273
  - impact of regulations, 273
  - internet of things, 277
  - invisible pixels, 135, 231
  - on captive portals, 202
  - on mobile devices, 38, 183
  - on smart TVs, 279
  - redirect, 30, 137
  - stateful, 28, 177
  - stateless, *see* fingerprinting
- tracking graph, 127
- tracking range, 126
- Tranco list, *see* toplist
- transience, 268
- Transmission Control Protocol (TCP), 15, 113
  - connection state, 88
- Transparency and Consent Framework (TCF), 203
- transparency research
  - blocking of, 189
  - challenges, 264
  - need for, 264
- transparency tools, 12
- Transport Layer Security (TLS), *see* TLS
- two-sample test, 165
- typosquatting, 217
- unicity, 42, 128, 187, 239
- uniform resource locator (URL), 17
- unique identifiers, 43, 99, 183
- uniqueness, *see* unicity
- usability, 267
- user awareness, 242, 267
- User Datagram Protocol (UDP), 15
- user engagement, 233
- user profiles, 40
  - consistency of, 187
- user recruitment, 267
- user registration, 217
- user reviews, 160, 222
- user studies, 267
- UUID, 278
- vantage points, 74, 79, 178, 265, 283
  - on mobile devices, 80
  - residential, 105
- virtual personas, 67, 84, 165, 192, 207, 265
  - on mobile devices, 85
- visualization, 164, 242
- voice assistants, *see* smart speakers
- voice commands, 82
- Wayback Machine, 108
- wealth transfer, 274
- web push notifications (WPN), 233
- WebGL, 182
- WebRTC, 33
- website breakage, *see* functionality
- websites, *see also* toplist, 265
  - categories, 76, 179, 234
  - control, 86
  - country-specific, 76
  - detection of topics, 155
  - e-commerce, 77, 96, 192
  - fast-changing, 132
  - malicious, 77, 156
  - online services, 77
  - selection of, 265
  - sensitive, 156, 266
  - subsites, 76, 179, 265
  - training, 85
  - travel, 78

330

WebSockets, 92, 233  
whois, 19  
  privacy solutions, 20, 132  
Whotracks.me, 110  
Wi-Fi, 20, 202, 277  
  captive portals, 20, 202  
  tracking on, 202  
  encryption, 20

Index

Wilcoxon signed-rank test, 167  
word embeddings, 150  
World Wide Web Consortium (W3C), 261  
  
Xposed framework, 118  
  
Zigbee, 278  
ZMap, 113