0xdata. *H2O*, 2013. https://github.com/0xdata/h2o.

3 Round Stones. *Callimachus*, 2013. http://callimachusproject.org/.

451 Research. "Database Landscape Map," n.d. http://blogs.the451group.com/information_management/files/2013/02/db_Map_2_13.jpg.

Abelson, Harold. *Structure and Interpretation of Computer Programs*. 2nd ed. Cambridge, Mass. : New York: MIT Press ; McGraw-Hill, 1996.

Abu-Mostafa, Yaser S, Magdon-Ismail, and Hsuan-Tien Lin. *Learning from Data: a Short Course*. [United States]: AMLBook.com, 2012.

Agarwal, Rohit. "Sparse Matrix Multiplication Using SQL," October 5, 2013. http://notes.mindprince.in/2013/06/07/sparse-matrix-multiplication-using-sql.html.

AGBeat. "Visualizing the Differences Between UX and UI," October 5, 2013. http://agbeat.com/tech-news/visualizing-the-differences-between-ux-and-ui/.

Agrawal, Ritesh. "On Writing Python UDF for Pig: A Perspective." *Memento*, September 23, 2013. http://ragrawal.wordpress.com/2013/02/24/on-writing-python-udf-for-pig-a-perspective/.

Airbnb. *Chronos*, 2013. https://github.com/airbnb/chronos.

Alag, Satnam. *Collective Intelligence in Action*. Greenwich, Conn.: Manning, 2008.

AlchemyAPI. *AlchemyAPI*, 2013. http://www.alchemyapi.com/.

Alves, Alexandre, Lloyd Williams, and Robin J. Smith. "An Overview of Complex Event Processing | Packt Publishing," n.d. http://www.packtpub.com/article/overview-complex-event-processing.

Amazon Web Services. *Amazon DynamoDB*, 2013. http://aws.amazon.com/dynamodb/.

———. *Amazon Elastic MapReduce (EMR)*, n.d. http://aws.amazon.com/elasticmapreduce/.

———. *Amazon Redshift*. Amazon Web Services, Inc., 2013. http://aws.amazon.com/redshift/.

AMPLab at UC Berkeley. *Shark*, 2013. https://github.com/amplab/shark.

Anand, Sid. "LinkedIn's Segmentation & Targeting Platform." Business & Mgmt, June 26, 2013. http://www.slideshare.net/r39132/linkedins-segmentation-targeting-platform?ref=http://e ventifier.co/event/hadoopsummit2013/slides.

Anderson, Jesse. "How-to: Use Eclipse with MapReduce in Cloudera's QuickStart VM," n.d. http://blog.cloudera.com/blog/2013/08/how-to-use-eclipse-with-mapreduce-in-clouderas-quickstart-vm/.

Anil, Robin, Ted Dunning, and Ellen Friedman. *Mahout in Action*. Manning Publications Company, 2011.

Apache Software Foundation. *Apache Ambari*, 2013. http://incubator.apache.org/ambari/.

———. *Apache Avro*, 2013. http://avro.apache.org/.

———. *Apache Cassandra*, 2013. http://cassandra.apache.org/.

———. *Apache Crunch*, 2013. http://crunch.apache.org/.

———. *Apache Drill*, 2013. http://incubator.apache.org/drill/.

———. *Apache Falcon*, 2013. http://falcon.incubator.apache.org/.

———. *Apache Flume*, n.d. http://flume.apache.org/.

———. *Apache Giraph*, 2013. http://giraph.apache.org/.

———. *Apache Hadoop*, 2013. http://hadoop.apache.org/.

———. *Apache Hadoop YARN*, 2013. http://hadoop.apache.org/docs/current/hadoop-yarn/hadoop-yarn-site/YARN.html.

———. *Apache HBase*, 2013. http://hbase.apache.org/.

———. *Apache Hive*, n.d. http://hive.apache.org/.

———. *Apache Kafka*, n.d. http://kafka.apache.org.

———. *Apache Lucene*, 2013. http://lucene.apache.org/.

———. *Apache Mahout*, 2013. http://mahout.apache.org/.

———. *Apache Mesos*, n.d. http://mesos.apache.org/.

———. *Apache Oozie*, 2013. http://oozie.apache.org/.

———. *Apache Pig*, 2013. http://pig.apache.org/.

———. *Apache S4*, n.d. http://incubator.apache.org/s4/.

———. *Apache Samza*, 2013. http://samza.incubator.apache.org/.

———. *Apache Solr*, 2013. http://lucene.apache.org/solr/.

———. *Apache Spark*, n.d. http://spark-project.org/.

———. *Apache Sqoop*, n.d. http://sqoop.apache.org/.

———. *Apache Tez*, 2013. http://tez.incubator.apache.org/.

———. *Apache Thrift*, 2013. http://thrift.apache.org/.

———. *Apache ZooKeeper*, 2013. http://zookeeper.apache.org/.

———. "SolrEcosystem," October 5, 2013. http://wiki.apache.org/solr/SolrEcosystem.

———. "The Apache HBase Reference Guide," September 23, 2013.
    http://hbase.apache.org/book/book.html.

Apt, Krzysztof R., Evangelos Markakis, and Sunil Simon. *Paradoxes in Social Networks with
    Multiple Products*. ArXiv e-print, January 31, 2013. http://arxiv.org/abs/1301.7592.

Armstrong, Joe. "Solving the Wrong Problem," n.d.
    http://joearms.github.com/2013/03/28/solving-the-wrong-problem.html.

Arthur, David. "Pig and HBase with LucidWorks Big Data," September 23, 2013.
    http://info.lucidworks.com/blog/bid/272738/Pig-and-HBase-with-LucidWorks-Big-Data.

Atbrox. *Atbr*, 2013. https://github.com/atbrox/atbr.

Aurelius. *Faunus*, 2013. https://github.com/thinkaurelius/faunus.

———. *Titan*, 2013. https://github.com/thinkaurelius/titan.

Automic (UC4). *ONE Automation*, n.d. http://automic.com/the-automation-engine/.

Bahmani, Bahman. "Sketching Techniques for Real-time Big Data." September 23, 2013.
    http://strataconf.com/strata2013/public/schedule/detail/27311.

Ballabio, Luigi. *QuantLib*, 2013. https://github.com/lballabio/quantlib.

———. *QuantLib-Python*, 2013. https://pypi.python.org/pypi/QuantLib-Python.

Barber, David. *Bayesian Reasoning and Machine Learning*. Cambridge; New York: Cambridge
    University Press, 2011.

Barker, Tom. *Pro Data Visualization Using R and JavaScript*. Berkeley, CA; New York:
    Apress ; Distributed to the book trade worldwide by Springer, 2013.

Barlow, Mike. *Big Data Culture Gap: Technology Advancing More Quickly Than People and
    Processes - Strata*, October 6, 2013.
    http://strata.oreilly.com/2013/09/big-data-culture-gap-technology-advancing-more-quickl
    y-than-people-and-processes.html#more-59859.

Bartlett, Randy. *A Practitioner's Guide to Business Analytics: Using Data Analysis Tools to
    Improve Your Organization's Decision Making and Strategy*. New York: McGraw-Hill,
    2013.

Bauer, Stefan. *Getting Started With Amazon Redshift*. Packt Publishing, 2013.
    http://www.myilibrary.com?id=498652.

Bayer, Judy, and Marie Taillard. "Data Analysis Should Be a Social Event." *Harvard Business*

Review, October 2, 2013. http://blogs.hbr.org/2013/04/data-analysis-should-be-a-soci/.

Beazley, David M. *Python Cookbook*. Beijing: O'Reilly, 2013.

———. *Python Essential Reference*. 4th ed. Developer's Library. Upper Saddle River, NJ: Addison-Wesley, 2009.

Beckett, Dave. *Redland RDF Libraries*, n.d. http://librdf.org/.

Bejeck, Bill. *Getting Started with Google Guava Write Better, More Efficient Java, and Have Fun Doing So*. Olton Birmingham [England]: Packt Pub. Ltd., 2013. http://site.ebrary.com/id/10747042.

Bergstein, Brian, and 2013. "Review: The Problem with Our Data Obsession." *MIT Technology Review*, October 5, 2013. http://www.technologyreview.com/review/511176/the-problem-with-our-data-obsession/.

Bernard, Scott A. *Introduction to Enterprise Architecture: Third Edition.* [S.l.]: Authorhouse, 2012.

Bhattacharya, Dibyendu. "Using Coprocessors to Index Columns in an Elasticsea..." Technology, August 2, 2013. http://www.slideshare.net/cloudera/hbasecon-2013-apache-hadoop-and-apache-hbase-for-realtime-video-analytics.

Biddulph, Matt. *Pig RedisStorer*, 2013. https://github.com/mattb/pig-redis.

Bird, Steven. *Natural Language Processing with Python*. 1st ed. Beijing ; Cambridge [Mass.]: O'Reilly, 2009.

———. *NLTK*, n.d. http://nltk.org/.

Bishop, Christopher M. *Pattern Recognition and Machine Learning*. New York, NY: Springer, 2009.

bitly. *Dablooms*, 2013. https://github.com/bitly/dablooms.

———. *Forget-Table*, 2013. https://github.com/bitly/forgettable.

Bladt, Jeff, and Bob Filbin. "A Data Scientist's Real Job: Storytelling," September 26, 2013. http://blogs.hbr.org/2013/03/a-data-scientists-real-job-sto/.

BNOSAC. "Massive Online Data Stream Mining with R," October 4, 2013. http://www.bnosac.be/index.php/blog/26-massive-online-data-stream-mining-with-r.

Borasky, Edward. *Computational-Journalism-Publishers-Workbench*, 2013. https://github.com/znmeb/Computational-Journalism-Publishers-Workbench.

Bosteels, Klaas. *Dumbo*, 2013. https://github.com/klbostee/dumbo.

Bostock, Michael. "Bl.ocks.org - Mbostock," n.d. http://bl.ocks.org/mbostock.

———. *D3 Plugins*, 2013. https://github.com/d3/d3-plugins.

———. *D3.js*, 2013. https://github.com/mbostock/d3.

———. "How Selections Work," n.d. http://bost.ocks.org/mike/selection/.

———. "Let's Make a Map," n.d. http://bost.ocks.org/mike/map/.

———. *Topojson*, 2013. https://github.com/mbostock/topojson.

Bostock, Michael, Vadim Ogievetsky, and Jeffrey Heer. "D 3: Data-Driven Documents" (September 28, 2013). http://vis.stanford.edu/files/2011-D3-InfoVis.pdf.

Botzum, Keys. "Basic Notes on Configuring Eclipse as a Hadoop Development Environment for MapR," n.d. http://www.mapr.com/blog/basic-notes-on-configuring-eclipse-as-a-hadoop-development-environment-for-mapr.

Bowen, Jonathan. *Getting Started with Talend Open Studio for Data Integration*. Packt Publishing Ltd, 2012.

Boyd, Danah, and Kate Crawford. *Six Provocations for Big Data*. SSRN Scholarly Paper. Rochester, NY: Social Science Research Network, September 21, 2011. http://papers.ssrn.com/abstract=1926431.

Bressert, Eli. *SciPy and NumPy*. Farnham: O'Reilly, 2013.

Brewer, Eric. "CAP Twelve Years Later: How the 'Rules' Have Changed," October 2, 2013. http://www.infoq.com/articles/cap-twelve-years-later-how-the-rules-have-changed.

Britcher, Robert N. *The Limits of Software: People, Projects, and Perspectives*. Reading, Mass.: Addison-Wesley, 1999.

Brown, Aaron. *Red-blooded Risk: The Secret History of Wall Street*. Hoboken, N.J: Wiley, 2012.

Bryer, Jason. *Sqlutils*, 2013. https://github.com/jbryer/sqlutils.

Bumgarner, Vincent. *Implementing Splunk: Big Data Reporting and Development for Operational Intelligence*. Birmingham: Packt Pub., 2013.

Cairo, Alberto. *The Functional Art: An Introduction to Information Graphics and Visualization*, 2013.

Capan, Gokhan. "Using Apache HBase for Large Matrices." Technology, July 9, 2013. http://www.slideshare.net/cloudera/hbasecon-2013-using-apache-hbase-for-large-matrices.

Capriolo, Edward. *Cassandra High Performance Cookbook*. Birmingham, UK: Packt Pub., 2011. http://proquest.safaribooksonline.com/?fpi=9781849515122.

———. *Clojure Data Analysis Cookbook*. [S.l.]: Packt Publishing Limited, 2013. http://lib.myilibrary.com?id=475335.

Capriolo, Edward, Dean Wampler, and Jason Rutherglen. *Programming Hive*. Sebastopol, CA: O'Reilly & Associates, 2012.

Carlsson, Gunnar. "Topology and Data." *Bulletin of the American Mathematical Society* 46, no. 2 (2009): 255–308.

Carroll, Jeremy. "Apache HBase Operations at Pinterest." Technology, July 9, 2013. http://www.slideshare.net/cloudera/operations-session-1.

Chambers, John M. *Software for Data Analysis Programming with R*. New York; London: Springer, 2008. http://dx.doi.org/10.1007/978-0-387-75936-4.

Chan, Ernest P. *Algorithmic Trading: Winning Strategies and Their Rationale*, 2013. http://search.ebscohost.com/login.aspx?direct=true&scope=site&db=nlebk&db=nlabk&AN=587988.

Chang, Sau Sheong. *Exploring Everythings with R and Ruby*. O'Reilly, 2012.

Chang, Winston. *R Graphics Cookbook*. Sebastopol, CA: O'Reilly Media, 2012. http://proquest.safaribooksonline.com/?fpi=9781449363086.

Chauhan, Avkash, Valentine Fontama, Michele Hart, Wee-Hyong Tok, and Denny Guang-Yeu Lee. *Simplifying Big Data With Microsoft Hdinsight*. Microsoft Press, 2013.

Chen, Songting. "Cheetah: a High Performance, Custom Data Warehouse on Top of MapReduce." *Proceedings of the VLDB Endowment* 3, no. 1–2 (2010): 1459–1468.

Chinchilla, Carlos. *Tales*, 2013. https://github.com/calufa/tales-core.

Ching, Avery. "Scaling Apache Giraph to a Trillion Edges," n.d. https://www.facebook.com/notes/facebook-engineering/scaling-apache-giraph-to-a-trillion-edges/10151617006153920.

Chu-Carroll, Mark C. *Good Math: a Geek's Guide to the Beauty of Numbers, Logic, and Computation*. Dallas, TX: Pragmatic Programmers, 2013.

Citus Data. *CitusDB*, 2013. http://www.citusdata.com/downloads.

Clarke, Daoud. *Seven Steps to Success: Machine Learning in Practice*, n.d.
    http://daoudclarke.github.io/guide.pdf.

Clayton, Rich. "Why Business Leaders Must Master Data And Analytics." *Forbes*, September
    27, 2013.
    http://www.forbes.com/sites/oracle/2013/04/22/why-business-leaders-must-master-data-a
    nd-analytics/.

Clegg, Andrew. "Pig-data-mining-talk." September 23, 2013.
    https://github.com/andrewclegg/pig-data-mining-talk.

Cloudera. *HTrace*, 2013. https://github.com/cloudera/htrace.

———. *Hue*, 2013. https://github.com/cloudera/hue.

———. *Impala*, 2013. http://www.cloudera.com/content/cloudera/en/products/cdh/impala.html.

———. *RecordBreaker*, n.d. http://cloudera.github.com/RecordBreaker/.

Collie, Rob. *DAX Formulas for PowerPivot*. Uniontown, Ohio: Holy Macro! Books, 2013.
    http://www.books24x7.com/marc.asp?bookid=50324.

Concurrent. *Cascading*, n.d. http://www.cascading.org/.

Condie, Tyson, Neil Conway, Peter Alvaro, Joseph M. Hellerstein, Khaled Elmeleegy, and
    Russell Sears. "MapReduce Online." 21–21. NSDI'10. Berkeley, CA, USA: USENIX
    Association, 2010. http://dl.acm.org/citation.cfm?id=1855711.1855732.

Congiu, Roberto. *Hive-JSON-Serde*, 2013. https://github.com/rcongiu/Hive-JSON-Serde.

Continuuity. *Continuuity Reactor*, 2013. http://continuuity.com/.

———. *Weave*, 2013. https://github.com/continuuity/weave.

Continuum Analytics. *Anaconda*. Continuum Analytics, 2013.
    https://store.continuum.io/cshop/anaconda/.

———. *Blaze*. Continuum Analytics, Inc., 2013. https://github.com/ContinuumIO/blaze.

———. *Bokeh*. Continuum Analytics, Inc., 2013. https://github.com/ContinuumIO/Bokeh.

———. *Wakari*, n.d. https://wakari.io/.

Conway, Drew, and John Myles White. *Machine Learning for Hackers*. Sebastopol, CA:
    O'Reilly, 2012.
    http://search.ebscohost.com/login.aspx?direct=true&scope=site&db=nlebk&db=nlabk&A
    N=436647.

Cook, John D. "Recognizing Numbers." *The Endeavour*, n.d.
    http://www.johndcook.com/blog/2013/04/30/recognizing-numbers/.

Corr, Lawrence, and Jim Stagnitto. *Agile Data Warehouse Design: Collaborative Dimensional
    Modeling, from Whiteboard to Star Schema*. Leeds: Decisionone Press, 2012.

Corum, Jonathan. "Storytelling with Data." n.d. http://style.org/tapestry/.

Craig, Terence, and Mary E Ludloff. *Privacy and Big Data*. Sebastopol, CA: O'Reilly, 2011.

Crawford, Kate. "The Hidden Biases in Big Data." *HBR Blog Network - Harvard Business
    Review*, September 26, 2013. http://blogs.hbr.org/2013/04/the-hidden-biases-in-big-data/.

Creasy, Jonathan, and Geoffrey Anderson. "OpenTSDB at Box." Technology, July 9, 2013.
    http://www.slideshare.net/cloudera/operations-session-3.

Croll, Alistair. "The Business Singularity: Why Software Means Cycle Time Trumps Scale."
    September 23, 2013. http://strataconf.com/strata2013/public/schedule/detail/27993.

CRS4. *Pydoop*, 2013. https://github.com/crs4/pydoop.

Damarla, Chanu. "Warren Buffett, The Human Big Data Engine." *Business 2 Community*,
    September 24, 2013.
    http://www.business2community.com/big-data/warren-buffett-the-human-big-data-engin

e-0484510.

Daniel Tunkelang. "Find and Be Found: Information Retrieval at LinkedIn." Technology, July 31, 2013. http://www.slideshare.net/dtunkelang/find-and-be-found-information-retrieval-at-linked-in.

Daniels, Doug. "Domesticating Pig with Lipstick." *Mortar Data Blog*, September 23, 2013. http://blog.mortardata.com/post/60765120319/domesticating-pig-with-lipstick.

Das, Devaraj, and Nicolas Liochon. "How to Get the MTTR Below 1 Minute and More." Technology, July 9, 2013. http://www.slideshare.net/cloudera/internals-session-2.

Datameer. *Datameer*, n.d. http://www.datameer.com/index.html.

Datasalt. *Pangool*, 2013. http://pangool.net/.

Davenport, Thomas H. *Enterprise Analytics: Optimize Performance, Process and Decisions through Big Data*. Upper Saddle River, N.J.: Financial Times/Prentice Hall, 2013.

Davenport, Thomas H, and Jinho Kim. *Keeping up with the Quants: Your Guide to Understanding and Using Analytics*, 2013.

Davenport, Tom. "How P&G Presents Data to Decision-Makers," n.d. http://blogs.hbr.org/cs/2013/04/how_p_and_g_presents_data.html.

Davidson-Pilon, Cameron. "Probabilistic-Programming-and-Bayesian-Methods-for-Hackers." *GitHub*, October 2, 2013. https://github.com/CamDavidsonPilon/Probabilistic-Programming-and-Bayesian-Methods-for-Hackers.

Davis, Kord, and Doug Patterson. *Ethics of Big Data*. Farnham: O'Reilly, 2012.

Day, Allen. "Data Platform Design Patterns." Education, September 17, 2013. http://www.slideshare.net/allenday54/20130905-singapore-big-datasg.

———. "NoSQL and SQL Work Side-by-Side to Tackle Real-time Big Data Needs." September 17, 2013. http://www.slideshare.net/allenday54/20130617-new-york-open-analytics-summit-no-sql-and-sql.

De Cadenza, Michele. "Online Available Data Services: a Primer." Education, July 3, 2013. http://www.slideshare.net/indecadenza/daa-s-xchange2013galassopiuntiv04.

De Goes, John. "Analytics Maturity Model." Technology, April 9, 2013. http://www.slideshare.net/jdegoes/analytics-maturity-model.

Dem\vsar, Janez, Toma\vz Curk, Ale\vs Erjavec, \vCrt Gorup, Toma\vz Ho\vcevar, Mitar Milutinovi\vc, Martin Mo\vzina, Matija Polajnar, Marko Toplak, and An\vze Stari\vc. "Orange: Data Mining Toolbox in Python." *Journal of Machine Learning Research* 14 (2013): 2349–2353.

Deming, W. Edwards. *Out of the Crisis*. Cambridge, Mass: Massachusetts Institute of Technology, Center for Advanced Engineering Study, 1986.

Denis, Frank. "Why We Love Apache Pig." *Umbrella Security Labs*, September 23, 2013. http://labs.umbrella.com/2013/04/08/pig-jruby/.

Devlin, Catherine. *Ipython-sql*, 2013. https://pypi.python.org/pypi/ipython-sql.

DigitalPebble. *Behemoth*, 2013. https://github.com/DigitalPebble/behemoth.

Dimiduk, Nick. "HBase Client APIs (for Webapps?)." March 28, 2013. http://www.slideshare.net/xefyr/hbase-client-apis-for-webapps.

———. "HBase Data Types (WIP)." June 25, 2013. http://www.slideshare.net/xefyr/hbase-data-types-wip.

———. *HBase in Action*. Shelter Island, NY: Manning, 2013.

Disco Project. *Disco*, 2013. https://github.com/discoproject/disco.

Downey, Allen. *Think Python*. Sebastopol, CA: O'Reilly, 2012.
http://www.contentreserve.com/TitleInfo.asp?ID={8BDDEF5E-F64A-4212-A868-9B60
E742EC40}&Format=50.

———. *Think Stats*. 1st ed. Probability and Statistics for Programmers. Sebastopol, CA:
O'Reilly Media, 2011.

Downey, Allen B. *Think Complexity*. Bejing [etc.]: O'Reilly, 2012.

Downey, Allen B. *Think Bayes.* Oreilly & Associates Inc, 2013.

Duarte, Nancy. "When Presenting Your Data, Get to the Point Fast." *HBR Blog Network -
Harvard Business Review*, September 27, 2013.
http://blogs.hbr.org/2013/03/when-presenting-your-data-get/.

DuCharme, Bob. *Learning SPARQL: Querying and Updating with SPARQL 1.1*. 1st ed.
Sebastopol, CA: O'Reilly Media, 2011.

Dunn, Roland, Ger Hobbelt, Andrew Thornton, Chris Viau, and Troy Mott. *Developing a D3.js
Edge*. Bleeding Edge Press, 2013.

Dunning, Ted. "Introduction to Mahout and How To Build a Recommender." Technology,
August 24, 2013.
http://www.slideshare.net/tdunning/dfw-big-data-talk-on-mahout-recommenders.

———. "Multi-Modal Recommendation." Technology, June 5, 2013.
http://www.slideshare.net/tdunning/buzz-wordsdunningmultimodalrecommendation.

———. "Using Mahout and a Search Engine for Recommendation." Technology, October 5,
2013.
http://www.slideshare.net/tdunning/using-mahout-and-a-search-engine-for-recommendati
on.

eBay Software Foundation. *Ql.io*, 2013. https://github.com/ql-io/ql.io.

Eckerson, Wayne. "Analytical Modeling Is Both Science and Art," September 26, 2013.
http://searchbusinessanalytics.techtarget.com/opinion/Analytical-modeling-is-both-scienc
e-and-art.

———. *Secrets of Analytical Leaders: Insights from Information Insiders*. 1st ed. Westfield, NJ:
Technics Publications, LLC, 2012.

———. *Visual Discovery Tools: Market Segmentation and Product Positioning*, September 23,
2013.
http://www.idgconnect.com/view_abstract/14386/visual-discovery-tools-market-segment
ation-product-positioning.

Elasticsearch. *ElasticSearch*, 2013. https://github.com/elasticsearch/elasticsearch.

———. *Elasticsearch Hadoop*, 2013. https://github.com/elasticsearch/elasticsearch-hadoop.

Elmahrek, Abraham. "How-to: Import a Pre-existing Oozie Workflow into Hue," n.d.
http://blog.cloudera.com/blog/2013/03/how-to-import-a-pre-existing-oozie-workflow-int
o-hue/.

EMRG. *NodeBox*, n.d. http://nodebox.net/.

Enthought. *EPD*, n.d. http://www.enthought.com/.

EPFL. *Scala*, 2013. http://www.scala-lang.org/.

Epstein, Charles, Gunnar Carlsson, and Herbert Edelsbrunner. "Topological Data Analysis."
Inverse Problems 27, no. 12 (December 1, 2011): 120201.
doi:10.1088/0266-5611/27/12/120201.

Erickson, Justin. "Impala: A Modern SQL Engine for Hadoop." September 23, 2013.
http://strataconf.com/strata2013/public/schedule/detail/27991.

EsperTech. *Esper*, n.d. http://esper.codehaus.org/.

Esri. *GIS Tools for Hadoop*, 2013. https://github.com/Esri/gis-tools-for-hadoop.

———. *Spatial Framework for Hadoop*, 2013.
https://github.com/Esri/spatial-framework-for-hadoop.

Etzion, Opher. *Event Processing in Action*. Greenwich, 74° w. long: Manning, 2011.

European Journalism Centre, and Open Knowledge Foundation. *The Data Journalism Handbook*. 1st ed. Sebastopol, CA: O'Reilly Media, 2012.

Facebook. *Scribe*, 2013. https://github.com/facebook/scribe.

Farina, Daniel. *WAL-E*, n.d. https://github.com/wal-e/wal-e#wal-e-disaster-recovery.

Farmer, F. Randall, and Bryce Glass. *Building Web Reputation Systems*. Sebastopol, CA: O'Reilly, 2010.

Faroult, Stéphane. *The Art of SQL*. 1st ed. Sebastopol, CA: O'Reilly, 2006.

Ferrera, P., I. de Prado, E. Palacios, J.L. Fernandez-Marquez, and G. Di Marzo Serugendo. "Tuple MapReduce." 260–269, 2012. doi:10.1109/ICDM.2012.141.

Few, Stephen. *Information Dashboard Design: Displaying Data for At-a-glance Monitoring*. Burlingame, CA: Analytics Press, 2013.

———. *Information Dashboard Design: The Effective Visual Communication of Data*. 1st ed. Beijing ; Cambride [MA]: O'Reilly, 2006.

———. *Now You See It: Simple Visualization Techniques for Quantitative Analysis*. Oakland, Calif: Analytics Press, 2009.

Flach, Peter. *Machine Learning*. Cambridge: CAMBRIDGE UNIVERSITY PRESS, 2012.

Flach, Peter A. *Machine Learning the Art and Science of Algorithms That Make Sense of Data*. Cambridge: Cambridge University Press, 2012.
http://dx.doi.org/10.1017/CBO9780511973000.

Fonnesbeck, Chris, Anand Patil, and John Salvatier. *PyMC*, 2013.
https://github.com/pymc-devs/pymc.

Ford, Daniel, Fran\ccois Labelle, Florentina I. Popovici, Murray Stokely, Van-Anh Truong, Luiz Barroso, Carrie Grimes, and Sean Quinlan. "Availability in Globally Distributed Storage Systems." 61–74, 2010. http://static.usenix.org/events/osdi10/tech/full_papers/Ford.pdf.

Franks, Bill. *Taming the Big Data Tidal Wave: Finding Opportunities in Huge Data Streams with Advanced Analytics*, 2012.

Fredheim, Rolf. "Quantifying Memory: Experiments in Python and D3 from R: GDELT Made Easy." *Quantifying Memory*, October 4, 2013.
http://quantifyingmemory.blogspot.com/2013/04/experiments-in-python-and-d3-from-r.html.

Fry, Ben. *Visualizing Data: Exploring and Explaining Data with the Processing Environment*. 1 edition. O'Reilly Media, 2008.

Fung, Kaiser. *Numbersense: How to Use Big Data to Your Advantage*, 2013.

Furuhashi, Sadayuki. *MessagePack*, 2013. https://github.com/msgpack/msgpack.

———. *MessagePack-Hadoop*, 2013. https://github.com/msgpack/msgpack-hadoop.

Galloway, Alexander R. *The Interface Effect*. Cambridge, UK ; Malden, MA: Polity, 2012.

Ganglia Development Team. *Ganglia Monitoring System*, 2013. http://ganglia.info/.

García, Miguel, and Barry Harmsen. *QlikView 11 for Developers: Develop Business Intelligence Applications with QlikView 11*. Birmingham: Packt Pub., 2012.

Garnaat, Mitch. *Python and AWS Cookbook*. Sebastopol, CA: O'Reilly, 2012.
http://proquest.safaribooksonline.com/9781449308100.

Gates, Alan. "How to Make Apache Hive 100X Faster? Apache Hive Stinger." *Hortonworks*,
September 23, 2013. http://hortonworks.com/blog/100x-faster-hive/.

———. "Pig and Hive at Yahoo!" *Yahoo*, September 23, 2013.
http://developer.yahoo.com/blogs/hadoop/pig-hive-yahoo-464.html.

———. *Programming Pig*. Sebastopol, CA: O'Reilly, 2011.
http://proquest.safaribooksonline.com/?fpi=9781449317881.

George, Lars. *HBase the Definitive Guide*. Sebastopol, CA: O'Reilly, 2011.
http://proquest.safaribooksonline.com/9781449314682.

Gitelman, Lisa. *"Raw Data" Is an Oxymoron*, 2013.
http://books.google.com/books?isbn=0262518287.

Godinez, Mario. *The Art of Enterprise Information Architecture: a Systems-based Approach for
Unlocking Business Insight*. Upper Saddle River, NJ: IBM Press, Pearson, 2010.

Goodhope, Ken, Joel Koshy, Jay Kreps, Neha Narkhede, Richard Park, Jun Rao, and Victor
Yang Ye. "Building LinkedIn's Real-time Activity Data Pipeline." *IEEE Data Eng. Bull.*
35, no. 2 (2012): 33–45.

Goodman, Sasha C. *RSeek.org*, n.d. http://rseek.org/.

Google. *Google BigQuery*, n.d. https://developers.google.com/bigquery/docs/overview.

———. *Google Fusion Tables*, 2013. https://developers.google.com/fusiontables/.

———. *Protocol Buffers*, 2013. https://code.google.com/p/protobuf/.

Gopalakrishna, Kishore. "Untangling Cluster Management with Helix." January 29, 2013.
http://www.slideshare.net/amywtang/helix-socc-v10final-16247096.

Goyal, Sonal. *Crux*, 2013. https://github.com/sonalgoyal/crux.

———. *HIHO*, 2013. https://github.com/sonalgoyal/hiho.

Graham, Bill. "Hadoop and Vertica: The Data Analytics Platform at Twitter." Business & Mgmt,
June 21, 2012.
http://www.slideshare.net/billonahill/hadoop-summit-2012-hadoop-and-vertica-the-data-a
nalytics-platform-at-twitter.

Gravity. *HPaste*, 2013. https://github.com/GravityLabs/HPaste.

Gregg, Brendan. "Linux Performance Analysis and Tools." March 1, 2013.
http://www.slideshare.net/brendangregg/linux-performance-analysis-and-tools.

Gregg, Forest, and Derek Eder. *Dedupe*, 2013. https://github.com/open-city/dedupe.

Groskopf, Christopher. *Csvkit*, 2013. https://github.com/onyxfish/csvkit.

Guha, Saptarshi. *RHIPE*, 2013. http://www.datadr.org/.

Guha, Saptarshi, Ryan Hafen, Jeremiah Rounds, Jin Xia, Jianfu Li, Bowei Xi, and William S.
Cleveland. "Large Complex Data: Divide and Recombine (D&R) with RHIPE." *Stat* 1,
no. 1 (2012): 53–67.

Guha, Saptarshi, Paul Kidwell, Ryan Hafen, and William S. Cleveland. "Visualization Databases
for the Analysis of Large Complex Datasets." 193–200, 2009.
http://machinelearning.wustl.edu/mlpapers/paper_files/AISTATS09_GuhaKHC.pdf.

Guo, Shumin. *Hadoop Operations and Cluster Management Cookbook*. Birmingham, UK: Packt
Publishing Ltd., 2013. http://site.ebrary.com/id/10742618.

Hafen, Ryan. *Datadr*, 2013. https://github.com/hafen/datadr.

———. *Trelliscope*, 2013. https://github.com/hafen/trelliscope.

Hall, Alexander, Olaf Bachmann, Robert Büssow, Silviu G\uanceanu, and Marc Nunkesser.

"Processing a Trillion Cells Per Mouse Click." *Proceedings of the VLDB Endowment* 5, no. 11 (2012): 1436–1446.

Hamstra, Mark, and Matei Zaharia. *Learning Spark: Lightning-fast Big Data Analytics*, 2013.

Hanborq. *PigML*, 2013. https://github.com/hanborq/pigml.

———. *RockStor*, 2013. https://github.com/hanborq/rockstor.

Harrington, Peter. *Machine Learning in Action*. Shelter Island, N.Y.: Manning Publications Co., 2012.

Hastie, Trevor, Robert Tibshirani, and J. H Friedman. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. New York: Springer, 2009.

Hayes, Matthew. "A Brief Tour of DataFu." Design, June 25, 2013. http://www.slideshare.net/matthewterencehayes/datafu.

———. "Building Data Products at LinkedIn with DataFu." Technology, August 26, 2013. http://www.slideshare.net/matthewterencehayes/building-data-products-at-linkedin-with-datafu.

———. "Hourglass: a Library for Incremental Processing on Hadoop." Education, October 9, 2013. http://www.slideshare.net/matthewterencehayes/hourglass-a-library-for-incremental-processing-on-hadoop.

Hayes, Matthew, and Sam Shah. "Hourglass: a Library for Incremental Processing on Hadoop." Santa Clara, CA, USA, n.d.

Heer, Jeffrey, Michael Bostock, and Vadim Ogievetsky. "A Tour Through the Visualization Zoo," September 26, 2013. http://hci.stanford.edu/jheer/files/zoo/.

Heer, Jeffrey, and Danah Boyd. "Vizster: Visualizing Online Social Networks." 32–39, 2005. http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=1532126.

Helmling, Gary. "A Bird's-Eye View of Pig and Scalding with hRaven." Business & Mgmt, June 28, 2013. http://www.slideshare.net/ghelmling/helmling-june27-1100amhall1?ref=http://eventifier.co/event/hadoopsummit2013/slides.

Hemann, Chuck, and Ken Burbary. *Digital Marketing Analytics Making Sense of Consumer Data in a Digital World*. Indianapolis, Ind.: Que, 2013. http://proquest.safaribooksonline.com/?fpi=9780133150933.

Herodotou, Herodotos, Harold Lim, Gang Luo, Nedyalko Borisov, Liang Dong, Fatma Bilgen Cetin, and Shivnath Babu. "Starfish: A Self-tuning System for Big Data Analytics." 11:261–272, 2011. http://x86.cs.duke.edu/~gang/documents/CIDR11_Paper36.pdf.

Herodotou, Herodotos, Harold Lim, Gang Luo, Nedyalko Borisov, Liang Dong, Fatma Bilgen Cetin, Shivnath Babu, Jie Li, Kunal Agarwal, and Fei Dong. *Starfish*, 2013. http://www.cs.duke.edu/starfish/.

Heule, Stefan, Marc Nunkesser, and Alexander Hall. "HyperLogLog in Practice: Algorithmic Engineering of a State of the Art Cardinality Estimation Algorithm." 683–692, 2013. http://dl.acm.org/citation.cfm?id=2452456.

Hewitt, Eben. *Cassandra: The Definitive Guide*. Beijing: O'Reilly, 2011.

Hewlett-Packard. *Vertica*, 2013. http://www.vertica.com/.

Hey, Tony, and Kristin Tolle. *The Fourth Paradigm Data-intensive Scientific Discovery*. Redmond, Wash.: Microsoft Research, 2009.

Hintjens, Pieter. *ZeroMQ*. Beijing: O'Reilly, 2013.

Hirst, Tony. "Mapping Corporate Networks With OpenCorporates." Technology, April 23, 2013.

http://www.slideshare.net/psychemedia/mapping-corporate-networks-with-opencorporate s?ref=http://blog.ouseful.info/2013/04/23/mapping-corporate-networks-with-opencorpora tes/.

Hoffman, Steve. *Apache Flume: Distributed Log Collection for Hadoop*. Packt Publishing, 2013. http://lib.myilibrary.com?id=504765.

Højsgaard, Søren, David Edwards, and Steffen L Lauritzen. *Graphical Models with R*. New York: Springer, 2012. http://dx.doi.org/10.1007/978-1-4614-2299-0.

Holmes, Alex. *Hadoop in Practice*. Shelter Island, NY: Manning, 2012.

Hornick, Mark, and Tom Plunkett. *Using R to Unlock the Value of Big Data: Big Data Anylytics with Oracle R Enterprise and Oracle R Connector for Hadoop*. New York: Mcgraw-Hill, 2013.

Hortonworks. *Apache Hadoop Patterns of Use: Refine, Enrich and Explore*, September 23, 2013. http://hortonworks.com/blog/apache-hadoop-patterns-of-use-refine-enrich-and-explore/.

———. "Hadoop Icons," n.d. http://hortonworks.com/blog/a-set-of-hadoop-related-icons/?mkt_tok=3RkMMJWWfF9 wsRoiua7OZKXonjHpfsX54%2BgvXKG%2FlMI%2F0ER3fOvrPUfGjI4CTsdjI%2BSL DwEYGJlv6SgFT7TMMbFh1rgNUxc%3D.

———. "Hadoop Patterns of Use." Business & Mgmt, April 10, 2013. http://www.slideshare.net/hortonworks/hadoop-patterns-of-use.

———. *Hoya*, 2013. https://github.com/hortonworks/hoya.

———. *The Business Value of Hadoop as Seen through the Big Data*, September 23, 2013. http://hortonworks.com/blog/the-business-value-of-hadoop-as-seen-through-the-big-data/ .

Howard, Jeremy, Margit Zwemer, and Mike Loukides. *Designing Great Data Products*. O'Reilly Media, Inc., 2012.

Hsieh, Jonathan, and Himanshu Vashishtha. "Using HBase Effectively - What You Need to Know as an Application Developer." September 23, 2013. http://strataconf.com/strata2013/public/schedule/detail/27109.

Hubbard, Douglas W. *How to Measure Anything Finding the Value of "Intangibles" in Business*. Hoboken, N.J.: Wiley, 2010. http://site.ebrary.com/id/10381009.

Huber, Peter J. *Data Analysis: What Can Be Learned From the Past 50 Years*. John Wiley & Sons, 2012.

Huff, Darrell. *How to Lie with Statistics*. New York: Norton, 1993.

Hultberg, Theo. *Piglet*, 2013. https://github.com/iconara/piglet.

———. *Rubydoop*, 2013. https://github.com/iconara/rubydoop.

Huß, Roland. *Jolokia*, 2013. https://github.com/rhuss/jolokia.

IBM. *Many Eyes*, n.d. http://www-958.ibm.com/software/data/cognos/manyeyes/.

Idris, Ivan. *NumPy Cookbook*. Birmingham: Packt Pub., 2012. http://proquest.safaribooksonline.com/?fpi=9781849518925.

iMatix. *ZeroMQ*, 2013. http://zeromq.org/.

Infochimps. *How to Do a Big Data Project*, n.d. http://bigdata.infochimps.com/Portals/174427/docs/bigdataprojecthowto.pdf.

———. *Pigsy*, 2013. https://github.com/infochimps-labs/pigsy.

———. *Wonderdog*, 2013. https://github.com/infochimps-labs/wonderdog.

———. *Wukong*, 2013. https://github.com/infochimps/wukong.

Ingersoll, Grant S. "Large-Scale Search Discovery Analytics with Hadoop, Mahout, Solr."
Technology, June 20, 2012.
http://www.slideshare.net/Hadoop_Summit/large-scale-search-discover-analytics-mahoot
-solr.

———. *Taming Text: How to Find, Organise, and Manipulate It*. Shelter Island, NY : London:
Manning ; Pearson Education [distributor], 2013.

Inmon, William H, and Krish Krishnan. *Building the Unstructured Data Warehouse*. Bradley
Beach, NJ: Technics Publications, 2011.
http://proquest.safaribooksonline.com/?fpi=9781935504047.

Internet Archive. *Waimea*, 2013. https://github.com/internetarchive/waimea.

IPython Development Team. *IPython*, 2013. https://github.com/ipython/ipython.

Isson, Jean Paul, and Jesse Harriott. *Win with Advanced Business Analytics Creating Business
Value from Your Data*. Hoboken, N.J.: John Wiley & Sons, 2013.
http://proxy.uqtr.ca/login.cgi?action=login&u=uqtr&db=books24x7&ezproxy=1&ezurl=
http://library.books24x7.com/library.asp?%5EB&bookid=46732.

Jacomy, Alexis. *Sigma.js*, n.d. http://sigmajs.org/.

Janert, Philipp K. *Data Analysis with Open Source Tools*. Sebastopol, CA: O'Reilly, 2011.

———. "Feedback Control for Programmers and Other Strangers." September 23, 2013.
http://strataconf.com/strata2013/public/schedule/detail/27667.

Jannach, Dietmar. *Recommender Systems: An Introduction*. New York: Cambridge University
Press, 2011.

Jian, Li. *Large Scale Data Processing in Hadoop MapReduce Scenario: Time Estimation and
Computation Models*. LAP LAMBERT Academic Publishing, 2012.

Jiang, Y. *HBase Administration Cookbook*. Birmingham: Packt Publishing, Limited, 2012.
https://ezproxy.siast.sk.ca:443/login?url=http://proquest.safaribooksonline.com/97818495
17140.

Johnson, Ian. *Tributary*, 2013. https://github.com/enjalot/tributary.

Jonas, Jeff. "Streaming Analytics Vs. Perpetual Analytics (Advantages of Windowless
Thinking)," October 4, 2013.
http://jeffjonas.typepad.com/jeff_jonas/2007/04/streaming_analy.html.

Journey, Russell. "HOWTO Use Hive to SQLize Your Own Tweets - Part Two: Loading Hive,
SQL Queries." *Hortonworks*, September 23, 2013.
http://hortonworks.com/blog/howto-use-hive-to-sqlize-your-own-tweets-part-two-loading
-hive-sql-queries/.

———. "LinkedIn, Apache Pig, and Open Source," September 23, 2013.
http://blog.linkedin.com/2010/07/01/linkedin-apache-pig/.

———. "Pig, ToJson, and Redis to Publish Data with Flask." *Hortonworks*, September 23, 2013.
http://hortonworks.com/blog/pig-tojson-and-redis-to-publish-data-with-flask/.

JSON-LD. *JSON-LD*, 2013. https://github.com/json-ld/json-ld.org.

Junqueira, Flavio, and Benjamin Reed. *Zookeeper: Distributed Process Coordination*. O'Reilly
Media, Incorporated, 2013.

Jurka, Tim. *RTextTools*, 2013. https://github.com/timjurka/RTextTools.

Jurney, Russell. *Agile Data Science: Building Data Analytics Applications with Hadoop*. [S.l.]:
O'Reilly Media, 2012.

Kabacoff, Robert. *R in Action: Data Analysis and Graphics with R*. Shelter Island, NY:
Manning, 2011.

Kane, Andrew. *Chartkick*, 2013. https://github.com/ankane/chartkick.

Karambelkar, H. *Scaling Big Data with Hadoop and Solr.* [S.l.]: Packt Publishing Limited, 2013.

Karmasphere. *Karmasphere*, 2013. http://www.karmasphere.com/.

Karnezos, Timon. "AWS Redshift: How Amazon Changed The Game," n.d.
    http://blog.aggregateknowledge.com/2013/05/16/aws-redshift-how-amazon-changed-the-game/.

Katz, Joel. *Designing Information: Human Factors and Common Sense in Information Design*.
    Hoboken, N.J: Wiley, 2012.

Kelly, Jeff. "Big Data Vendor Revenue and Market Forecast 2012-2017," October 5, 2013.
    http://wikibon.org/w/index.php?title=Big_Data_Vendor_Revenue_and_Market_Forecast
    _2012-2017&printable=yes.

Kent, William. *Data and Reality*. Bloomington, Ind.: 1st Books Library, 2000.

Khanwalkar, Manoj, and Govind Asawa. "Evolving a First-Generation Apache HBase
    Deployment..." Technology, August 2, 2013.
    http://www.slideshare.net/cloudera/hbasecon-2013-evolving-a-firstgeneration-apache-hba
    se-deployment-to-second-generation-and-beyond.

Kiji Project. *Kiji*, 2013. http://www.kiji.org/.

Kimball, Ralph, and Margy Ross. *The Data Warehouse Toolkit The Definitive Guide to
    Dimensional Modeling.* Hoboken: Wiley, 2013.
    http://public.eblib.com/EBLPublic/PublicView.do?ptiID=1313513.

Kirk, Andy. *Data Visualization a Successful Design Process*. Birmingham, UK: Packt Pub,
    2013.

———. "Essential Collection of Visualisation Resources," October 8, 2013.
    http://www.visualisingdata.com/index.php/resources/.

———. "The 8 Hats of Data Visualisation." Technology, June 23, 2012.
    http://www.slideshare.net/visualisingdata/the-8-hats-of-data-visualisation.

———. "Understanding Learning in Order to Implement Efficient Visualisation Methods."
    Technology, September 21, 2012.
    http://www.slideshare.net/visualisingdata/andy-kirk-talk-at-big-data-world-europe-septe
    mber-2012.

———. "Visualisation Workflow: Finding Stories and Telling Stories." January 10, 2013.
    http://www.slideshare.net/visualisingdata/andy-kirks-facebook-talk.

Klr. "D3 Lifeline from Vega and Clickme." *Timely Portfolio*, April 4, 2013.
    http://timelyportfolio.blogspot.com/2013/04/d3-lifeline-from-vega-and-clickme.html.

KNIME.com AG. *KNIME*, 2013. http://www.knime.org/.

Kohei, KaiGai. *Pg_strom*, 2013. https://github.com/kaigai/pg_strom.

Koomey, Jon. *Turning Numbers into Knowledge: Mastering the Art of Problem Solving*.
    Oakland, CA: Analytics Press, 2008.

Kosara, Robert, and Jock Mackinlay. "Storytelling: The Next Step for Visualization." *Computer*
    46, no. 5 (2013): 44–50.

Koshy, Joel. "Building a Real-Time Data Pipeline: Apache Kafka at LinkedIn." Technology,
    July 2, 2013.
    http://www.slideshare.net/amywtang/building-a-realtime-data-pipeline-apache-kafka-at-li
    nked-in.

Krishnan, Krish. *Data Warehousing in the Age of Big Data*, 2013.
    http://lib.myilibrary.com?id=490441.

Krishnan, Sriram. "Genie - Hadoop Platform as a Service at Netflix." Technology, June 25, 2013. http://www.slideshare.net/krishflix/genie-hadoop-platform-as-a-service-at-netflix.

Kromer, Philip. *Big Data for Chimps.* [S.l.]: O'Reilly Media, 2013.

Kuć, Rafał. *Apache Solr 4 Cookbook*. Birmingham, UK: Packt Pub., 2013. http://proquest.safaribooksonline.com/?fpi=9781782161325.

Kuć, Rafał, and Marek Rogoziński. *Elasticsearch Server: Create a Fast, Scalable, and Flexible Search Solution with the Emerging Open Source Search Server, ElasticSearch*. Birmingham [u.a.]: Packt Publ., 2013.

Kuhn, Max, and Kjell Johnson. *Applied Predictive Modeling*. New York, NY: Springer, 2013.

Kumar, Arun, Feng Niu, and Christopher Ré. "Hazy: Making It Easier to Build and Maintain Big-data Analytics." *Queue* 11, no. 1 (January 2013): 30:30–30:46. doi:10.1145/2428616.2431055.

Kumar, Hari. "Ericsson Labs | HBase: Performance Tuners," September 23, 2013. http://labs.ericsson.com/blog/hbase-performance-tuners.

Kyrola, Aapo, Guy Blelloch, and Carlos Guestrin. "GraphChi: Large-scale Graph Computation on Just a PC." 31–46, 2012. https://www.usenix.org/system/files/conference/osdi12/osdi12-final-126.pdf.

Kyrola, Aapo, and GraphLab. *GraphChi*, n.d. http://graphchi.org.

Laberge, Robert. *The Data Warehouse Mentor: Practical Data Warehouse and Business Intelligence Insights*. New York: McGraw-Hill, 2011.

Ladley, John. *Data Governance How to Design, Deploy, and Sustain an Effective Data Governance Program*. Waltham, Mass.: Morgan Kaufmann, 2012.

Lam, Chuck. *Hadoop in Action*. Greenwich, Conn.: Manning Publications, 2011.

Lamb, Andrew, Matt Fuller, Ramakrishna Varadarajan, Nga Tran, Ben Vandiver, Lyric Doshi, and Chuck Bear. "The Vertica Analytic Database: C-store 7 Years Later." *Proceedings of the VLDB Endowment* 5, no. 12 (2012): 1790–1801.

Lans, Rick F. van der. *Data Virtualization for Business Intelligence Systems: Revolutionizing Data Integration for Data Warehouses*. Amsterdam: Elsevier, Morgan Kaufmann, 2012.

Le Dem, Julien, and Nong Li. "Parquet: Columnar Storage for the People." Technology, June 27, 2013. http://www.slideshare.net/julienledem/parquet-hadoop-summit-2013.

Ledolter, Johannes. *Data Mining and Business Analytics with R*. Hoboken, N.J.: Wiley, 2013. http://UCM.eblib.com/patron/FullRecord.aspx?p=1204741.

Lee, George, Jimmy Lin, Chuang Liu, Andrew Lorek, and Dmitriy Ryaboy. "The Unified Logging Infrastructure for Data Analytics at Twitter." *Proceedings of the VLDB Endowment* 5, no. 12 (2012): 1771–1780.

Lee, Kyungyong, Roy Indrajit, and Vanish Talwar. "Comparing Pattern Mining on a Billion Records with HP Vertica and Hadoop," n.d. http://www.vertica.com/2013/04/08/comparing-pattern-mining-on-a-billion-records-with-hp-vertica-and-hadoop/.

Leetaru, Kalev. "GDELT Data," n.d. http://eventdata.psu.edu/data.dir/GDELT.html#.UWMC3-UnquE.twitter.

Lemoine, Nathan. "Python Complements R's Shortcomings." *R-bloggers*, October 4, 2013. http://www.r-bloggers.com/python-compliments-rs-shortcomings/.

Lengler, Ralph, and Martin Eppler. "A Periodic Table of Visualization Methods," September 26, 2013. http://www.visual-literacy.org/periodic_table/periodic_table.html.

Leo, Simone, and Gianluigi Zanetti. "Pydoop: a Python MapReduce and HDFS API for

Hadoop." 819–825. HPDC '10. New York, NY, USA: ACM, 2010. doi:10.1145/1851476.1851594.

Leonhardi, David. "Soft Skills: The 'Killer App' for Analytics," October 4, 2013. http://www.analytics-magazine.org/november-december-2011/475-soft-skills-the-killer-app-for-analytics?tmpl=component&print=1&page=.

Liebowitz, Jay. *Big Data and Business Analytics*. Boca Raton, FL: CRC Press, 2013. http://search.ebscohost.com/login.aspx?direct=true&scope=site&db=nlebk&db=nlabk&AN=562870.

Lin, Jimmy. *Cloud9*, 2013. https://github.com/lintool/Cloud9.

———. *Ivory*, 2013. https://github.com/lintool/Ivory.

———. *MapReduce Is Good Enough? If All You Have Is a Hammer, Throw Away Everything That's Not a Nail!* ArXiv e-print, September 10, 2012. http://arxiv.org/abs/1209.2191.

Lin, Jimmy, and Chris Dyer. *Data-intensive Text Processing with MapReduce*. [San Rafael, Calif.]: Morgan & Claypool Publishers, 2010. http://www.morganclaypool.com/doi/abs/10.2200/S00274ED1V01Y201006HLT007.

———. "Data-intensive Text Processing with MapReduce," September 28, 2013. http://lintool.github.io/MapReduceAlgorithms/index.html.

Lin, Jimmy, and Alek Kolcz. "Large-scale Machine Learning at Twitter." 793–804, 2012. http://dl.acm.org/citation.cfm?id=2213958.

Lin, Jimmy, and Dmitriy Ryaboy. "Scaling Big Data Mining Infrastructure: The Twitter Experience." *ACM SIGKDD Explorations Newsletter* 14, no. 2 (2013): 6–19.

Lin, Liang, Vera Lychagina, Weiran Liu, Younghee Kwon, Sagar Mittal, and Michael Wong. "Tenzing a Sql Implementation on the Mapreduce Framework" (2011). http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.226.772.

Linked Data, 2013. http://linkeddata.org/.

LinkedIn. *Azkaban*, 2013. http://azkaban.github.io/azkaban2/.

———. *Camus*. Accessed October 12, 2013. https://github.com/linkedin/camus.

———. *Databus*, 2013. https://github.com/linkedin/databus.

———. *DataFu*, 2013. https://github.com/linkedin/datafu.

———. *White Elephant*. Accessed October 12, 2013. https://github.com/linkedin/white-elephant.

Linoff, Gordon. *Data Analysis Using SQL and Excel*. Indianapolis, IN: Wiley Pub, 2008.

Liu, Bing. *Sentiment Analysis and Opinion Mining*. San Rafael, Calif: Morgan & Claypool, 2012.

Longo, Chris. *Pig for Sublime Text*. Accessed October 12, 2013. https://github.com/chrislongo/Pig.

Loo, Mark P.J. van der. *Learning RStudio for R Statistical Computing: Learn to Effectively Perform R Development, Statistical Analysis, and Reporting with the Most Popular R IDE*. Birmingham: Packt Publishing, 2012.

Lorica, Ben. "Data Science Tools: Are You 'All in' or Do You 'Mix and Match'?," September 26, 2013. http://strata.oreilly.com/2013/03/data-science-tools-all-in-or-mix-and-match.html.

———. "Near Realtime, Streaming, and Perpetual Analytics," September 26, 2013. http://strata.oreilly.com/2013/07/near-realtime-streaming-and-perpetual-analytics.html?imm_mid=0adcff&cmp=em-strata-na-na-newsltr_20130807_elist.

———. "Simpler Workflow Tools Enable the Rapid Deployment of Models," September 26,

2013.
http://strata.oreilly.com/2013/04/workflow-tools-enable-the-rapid-deployment-of-models.html.

Loshin, David. *Big Data Analytics From Strategic Planning to Enterprise Integration With Tools, Techniques, Nosql, and Graph.* Morgan Kaufmann Pub, 2013.

———. "New Paradigms for High Performance Analytical Computing" (2009).
https://www.ndm.net/datawarehouse/pdf/High_Performance_Paradigms.pdf.

Loughran, Steve. "Hoya (HBase on YARN) : Application Architecture." *Hortonworks*, September 23, 2013.
http://hortonworks.com/blog/hoya-hbase-on-yarn-application-architecture/.

Loukides, Mike. *The Evolution of Data Products*. O'Reilly Media, Inc., 2011.

Lublinsky, Boris, Kevin T Smith, and Yakubovich. *Professional Hadoop Solutions*, 2013.

LucidWorks. *LucidWorks Big Data*, 2013.
http://docs.lucidworks.com/display/bigdata/System+Architecture.

Lum, P. Y., G. Singh, A. Lehman, T. Ishkanov, M. Vejdemo-Johansson, M. Alagappan, J. Carlsson, and G. Carlsson. "Extracting Insights from the Shape of Complex Data Using Topology." *Scientific Reports* 3 (February 7, 2013). doi:10.1038/srep01236.

MADlib Project. *MADlib*, 2013. https://github.com/madlib/madlib.

Malewicz, Grzegorz, Matthew H. Austern, Aart J.C Bik, James C. Dehnert, Ilan Horn, Naty Leiser, and Grzegorz Czajkowski. "Pregel: a System for Large-scale Graph Processing." 135–146. SIGMOD '10. New York, NY, USA: ACM, 2010.
doi:10.1145/1807167.1807184.

Malik, Jaibeer. "ElasticSearch: Text Analysis for Content Enrichment." *Jai's Weblog - Tech, Security & Fun...*, September 26, 2013.
http://jaibeermalik.wordpress.com/2013/03/26/elasticsearch-text-analysis-for-content-enrichment/.

Manning, Christopher D, Prabhakar Raghavan, and Hinrich Schütze. *Introduction to Information Retrieval*. New York: Cambridge University Press, 2008.

Manoochehri, Michael. *Data Just Right: Practical Big Data Analytics*. Harlow: Addison-Wesley, 2013.

Manyika, James, Michael Chui, Brad Brown, Jacques Bughin, Richard Dobbs, Charles Roxburgh, and Angela Hung Byers. *Big Data: The Next Frontier for Innovation, Competition, and Productivity*. McKinsey Global Institute, 2011.

Marchand, Donald A., and Joe Peppard. "Why IT Fumbles Analytics," October 2, 2013.
http://hbr.org/2013/01/why-it-fumbles-analytics/ar/1?imm_mid=0a4eb2&cmp=em-strata-newsletters-stratasc-video-20130403-elist.

Marmanis, Haralambos, and Dmitry Babenko. *Algorithms of the Intelligent Web*. Greenwich, Conn.: Manning, 2009.

Martin, Patrick. *Dex*, n.d. http://dexvis.com/.

———. *DexCharts*, 2013. https://github.com/PatMartin/DexCharts.

Marz, Nathan. *Big Data: Principles and Best Practices of Scalable Realtime Data Systems.* [S.l.]: O'Reilly Media, 2013.

———. *Cascalog*, 2013. https://github.com/nathanmarz/cascalog.

———. *ElephantDB*, 2013. https://github.com/nathanmarz/elephantdb.

———. "Human Fault-tolerance." September 23, 2013.
http://strataconf.com/strata2013/public/schedule/detail/27610.

———. *Storm*, 2013. https://github.com/nathanmarz/storm.

Mashape. "List of 40+ Machine Learning APIs." *Mashape's Voice*, September 27, 2013. http://blog.mashape.com/post/48074869493/list-of-40-machine-learning-apis.

Massie, Matt. "A Powerful Big Data Trio: Spark, Parquet and Avro," September 26, 2013. http://zenfractal.com/2013/08/21/a-powerful-big-data-trio/.

———. *Monitoring with Ganglia*. Sebastopol, CA: O'Reilly Media, 2012. http://proquest.safaribooksonline.com/?fpi=9781449330637.

Matloff, Norman S. *The Art of R Programming: Tour of Statistical Software Design*. San Francisco: No Starch Press, 2011.

Matsumoto, Yukihiro. *Ruby*, 2013. https://www.ruby-lang.org.

Mayer-Schönberger, Viktor, and Kenneth Cukier. *Big Data: a Revolution That Will Transform How We Live, Work, and Think*. Boston: Houghton Mifflin Harcourt, 2013.

McCallum, Q. Ethan. *Bad Data Handbook*. Beijing; Sebastopol, CA: O'Reilly, 2012.

McDaniel, Eileen, and Stephen McDaniel. *The Accidental Analyst: Show Your Data Who's Boss*. Seattle, WA: Freakalytics, LLC, 2012.

McDonald, Ryan, Keith Hall, and Gideon Mann. "Distributed Training Strategies for the Structured Perceptron." 456–464. HLT '10. Stroudsburg, PA, USA: Association for Computational Linguistics, 2010. http://dl.acm.org/citation.cfm?id=1857999.1858068.

McKinney, Wes. "Intro to Python for Financial Data Analysis." Technology, June 22, 2012. http://www.slideshare.net/wesm/intro-to-python-for-financial-data-analysis.

———. *Python for Data Analysis*. Sebastopol, CA: O'Reilly Media, 2012. http://proquest.safaribooksonline.com/?fpi=9781449323592.

Melnik, Sergey, Andrey Gubarev, Jing Jing Long, Geoffrey Romer, Shiva Shivakumar, Matt Tolton, and Theo Vassilakis. "Dremel: Interactive Analysis of Web-scale Datasets." *Proceedings of the VLDB Endowment* 3, no. 1–2 (2010): 330–339.

Menguy, Charles. *Hadoop Programming - Pushing the Limit (Pushing the Limits)*. 1 edition. Wiley, 2014.

Meru, Varad. "Step-by-step MapReduce Programming," October 5, 2013. http://www.orzota.com/step-by-step-mapreduce-programming/.

Mesosphere. *Marathon*, 2013. https://github.com/mesosphere/marathon.

Metamarkets. *Druid*, n.d. http://druid.io/.

Microsoft. *HDInsight*, n.d. http://www.microsoft.com/en-us/sqlserver/solutions-technologies/business-intelligence/big-data.aspx.

———. *StreamInsight*, n.d. http://www.microsoft.com/sqlserver/en/us/solutions-technologies/business-intelligence/streaming-data.aspx.

Microsoft, Yahoo, and John Langford. *Vowpal_wabbit*, 2013. https://github.com/JohnLangford/vowpal_wabbit.

Miller, Cat. "Pig Eye for the SQL Guy." *Hortonworks*, September 26, 2013. http://hortonworks.com/blog/pig-eye-for-the-sql-guy/.

Miller, Thomas W. *Modeling Techniques in Predictive Analytics: Business Problems and Solutions with R*, 2014.

Milne, Adrian. "Dispatches from the Codeface: Complex Event Processing Made Easy (using Esper)," October 3, 2013. http://corsoftlimited.blogspot.co.uk/2013/02/complex-event-processing-made-easy.html.

Minelli, Michael, Michele Chambers, and Ambiga Dhiraj. *Big Data, Big Analytics: Emerging Business Intelligence and Analytic Trends for Today's Businesses*. Hoboken, N.J: John Wiley, 2013.

Miner, Donald, and Adam Shook. *MapReduce Design Patterns*. Sebastopol, CA: Oreilly, 2013.

MIT. *STAR: Cluster*, n.d. http://star.mit.edu/cluster/.

Mohanty, Soumendra, Madhu Jagadeesh, and Harsha Srivatsa. *Big Data Imperatives: Enterprise Big Data Warehouse, BI Implementations and Analytics*, 2013. http://dx.doi.org/10.1007/978-1-4302-4873-6.

Mohri, Mehryar, Afshin Rostamizadeh, and Ameet Talwalkar. *Foundations of Machine Learning*. Cambridge, MA: MIT Press, 2012. http://search.ebscohost.com/login.aspx?direct=true&scope=site&db=nlebk&db=nlabk&AN=478737.

MongoDB. *MongoDB*, n.d. http://www.mongodb.org/.

Mortar Data. *Mortar Data*, 2013. http://mortardata.com/.

Movsisyan, Mher. *Chartkick.py*, 2013. https://github.com/mher/chartkick.py.

Mozilla Foundation. *Akela*, 2013. https://github.com/mozilla-metrics/akela.

———. *Socorro*, 2013. https://github.com/mozilla/socorro.

Mueller, Andreas. *Scikit-learn*, 2013. https://github.com/scikit-learn/scikit-learn.

Murphy, Kevin P. *Machine Learning a Probabilistic Perspective*. Cambridge, Mass.: MIT Press, 2012. http://search.ebscohost.com/login.aspx?direct=true&scope=site&db=nlebk&db=nlabk&AN=480968.

Murphy, Maxwell. "CFOs Ignore Big Data at Their Peril." *Wall Street Journal*, July 18, 2013. http://online.wsj.com/article/SB10001424127887323419604578571684135006800.html.

Murray, Scott. *Interactive Data Visualization for the Web*, 2013. http://search.ebscohost.com/login.aspx?direct=true&scope=site&db=nlebk&db=nlabk&AN=548862.

Murthy, Arun. *Apache Hadoop Yarn: Moving Beyond Mapreduce and Batch Processing with Apache*. [S.l.]: Addison-Wesley, 2013. http://my.safaribooksonline.com/9780133441925.

———. "Re: About the Combiner Execution," September 23, 2013. http://mail-archives.apache.org/mod_mbox/hadoop-hdfs-user/201107.mbox/%3C374D8F3F-B8B1-499F-BEDB-BFEE3219010C@hortonworks.com%3E.

Nahum, Dotan J. *Graphene*, 2013. https://github.com/jondot/graphene.

Nandeshwar, Ashutosh. *Tableau Data Visualization Cookbook*. Packt Publishing, Limited, 2013.

Nathan, Paco. *Enterprise Data Workflows with Cascading*. [S.l.]: O'Reilly Media, 2013.

———. "Pattern – an Open Source Project for Migrating Predi..." Technology, July 29, 2013. http://www.slideshare.net/pacoid/pattern-an-open-source-project-for-migrating-predictive-models-from-sas-etc-onto-hadoop.

Natkins, Jon. "How-to: Develop CDH Applications with Maven and Eclipse," n.d. http://blog.cloudera.com/blog/2012/08/developing-cdh-applications-with-maven-and-eclipse/.

Neo Technology. *Neo4j*, 2013. http://www.neo4j.org/.

Netflix. *Asgard*, 2013. https://github.com/Netflix/asgard.

———. *Exhibitor*, 2013. https://github.com/Netflix/exhibitor.

———. *Genie*, 2013. https://github.com/Netflix/genie.

———. *Lipstick*, 2013. https://github.com/Netflix/Lipstick.

———. *Priam*, 2013. https://github.com/Netflix/Priam.

———. *SimianArmy*, 2013. https://github.com/Netflix/SimianArmy.

Neumeyer, Leonardo, Bruce Robbins, Anish Nair, and Anand Kesari. "S4: Distributed Stream Computing Platform." 170–177, 2010. http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=5693297.

Noll, Machael. "Reading and Writing Avro Files from the Command Line," n.d. http://www.michael-noll.com/blog/2013/03/17/reading-and-writing-avro-files-from-the-command-line/.

Noris, Basilio. *MLDemos*, 2013. https://gitorious.org/mldemos/.

Nullege. "Nullege: A Search Engine for Python Source Code," n.d. http://nullege.com/.

Numenta. *NuPIC*, 2013. https://github.com/numenta/nupic.

Nussbaum, Bruce. *Creative Intelligence: Harnessing the Power to Create, Connect, and Inspire*. New York: HarperBusiness, 2013.

O'Malley, Owen. "ORC File and Vectorization - Hadoop Summit 2013." June 28, 2013. http://www.slideshare.net/oom65/orc-andvectorizationhadoopsummit?ref=http://www.slideshare.net/slideshow/embed_code/23625246.

O'Neil, Cathy. *Doing Data Science.* [S.l.]: O'Reilly Media, 2013.

O'Reilly, Tim. *How Data Science Is Transforming Health Care*. Sebastopol, CA: O'Reilly Media, 2012. http://search.ebscohost.com/login.aspx?direct=true&scope=site&db=nlebk&db=nlabk&AN=537815.

Ogievetsky, Vadim. "Facet: The Recursive Approach to Visualization." September 23, 2013. http://strataconf.com/strata2013/public/schedule/detail/27485.

Ohlhorst, Frank. *Big Data Analytics Turning Big Data into Big Money*. Hoboken, N.J.: Wiley, 2013. http://search.ebscohost.com/login.aspx?direct=true&scope=site&db=nlebk&db=nlabk&AN=503336.

Open Data Group. *Augustus*, n.d. https://code.google.com/p/augustus/.

OpenRefine Development Team. *OpenRefine*, n.d. https://github.com/OpenRefine/OpenRefine.

OpenSCG. *BigSQL*, n.d. http://www.bigsql.org/se/.

Ottenheimer, Davi. *Big Data Security.* John Wiley & Sons Inc, 2014.

Owens, Jonathan R, Jon Lentz, and Brian Femiano. *Hadoop Real-world Solutions Cookbook*. Birmingham: Packt Pub., 2013.

Pace, Larry. *Beginning R: An Introduction to Statistical Programming*. New York: Apress : Distributed to the book trade worldwide by Springer Science+Business Media, 2012.

Paradigm4. *SciDB*, n.d. http://www.scidb.org/.

Partner, Jonas. *Neo4j in Action.* [S.l.]: O'Reilly Media, 2013.

Patterson, Josh. *Knitting Boar*, 2013. https://github.com/jpatanooga/KnittingBoar.

———. "Linear Regression and Metronome." March 24, 2013. http://www.slideshare.net/jpatanooga/hadoop-summit-eu-2013-parallel-linear-regression-iterativereduce-and-yarn.

Patterson, Pat. *Database.com-FDW*, n.d. https://github.com/metadaddy-sfdc/Database.com-FDW-for-PostgreSQL.

Pentaho. *Pentaho Big Data Plugin*, 2013. https://github.com/pentaho/big-data-plugin.

Perera, Srinath, and Thilina Gunarathne. *Hadoop MapReduce Cookbook*. Birminham: Packt Pub., 2013.

Perez, Antonio Santiago. *Openlayers Cookbook*. [S.l.]: Packt Publishing Limited, 2012.

Perez, Fernando. "Literate Computing and Computational Reproducibility: IPython in the Age of Data-driven Journalism," n.d. http://blog.fperez.org/2013/04/literate-computing-and-computational.html.

Perkins, Jacob. "Data Recipes: Using Hadoop to Explore Chaos." *Data Recipes*, August 12, 2013. http://thedatachef.blogspot.com/2013/08/using-hadoop-to-explore-chaos.html.

———. *NLTK Trainer*, 2013. https://github.com/japerk/nltk-trainer.

———. *Python Text Processing with NLTK 2.0 Cookbook: Over 80 Practical Recipes for Using Python's NLTK Suite of Libraries to Maximize Your Natural Language Processing Capabilities*. Birmingham [u.a.]: Packt Publ., 2010.

Perkins, Jacob A. *Sounder*, 2013. https://github.com/thedatachef/sounder.

———. *SwineHerd*, 2013. https://github.com/thedatachef/swineherd.

———. *Varaha*, 2013. https://github.com/thedatachef/varaha.

Peterson, Mike. "Using Hadoop to Expand Data Warehousing." September 23, 2013. http://strataconf.com/strata2013/public/schedule/detail/28638.

Phillips, Judah. *Building a Digital Analytics Organization Create Value by Integrating Analytical Processes, Technology, and People into Business Operations*, 2013. http://proquest.safaribooksonline.com/?fpi=9780133372823.

Phillips, Lee, and Inc ebrary. *Gnuplot Cookbook*. Birmingham: Packt Publishing, Limited, 2012.

Pidcock, Woody, and Michael Uschold. "What Are the Differences Between a Vocabulary, a Taxonomy, a Thesaurus, an Ontology, and a Meta-model?," October 4, 2013. http://infogrid.org/trac/wiki/Reference/PidcockArticle.

Pike, Rob, Sean Dorward, Robert Griesemer, and Sean Quinlan. "Interpreting the Data: Parallel Analysis with Sawzall" (October 2, 2013). http://lambda.csail.mit.edu/~chet/papers/others/p/pike/sawzall-sciprog.pdf.

Pinte, Didrik. "Interactive Python Graphics/visualisation with Excel," n.d. http://dpinte.wordpress.com/2010/03/12/interactive-python-graphicsvisualisation-with-excel/.

Plunkett, Tom, Brian Macdonald, Bruce Nelson, Helen Sun, Mark F Hornick, Laker, Khader Mohiuddin, et al. *Oracle Big Data Handbook*, 2014.

Preferred Infrastructure, and NTT Software Innovation Center. *Jubatus*, 2013. https://github.com/jubatus/jubatus.

Press, William H. *Numerical Recipes: The Art of Scientific Computing*. Cambridge [u.a.: Cambridge Univ. Press, 2007.

Project Blacklight. *Blacklight*, 2013. https://github.com/projectblacklight/blacklight.

Provost, Foster, and Tom Fawcett. *Data Science for Business: [what You Need to Know About Data Mining and Data-analytic Thinking]*. Sebastopol, Calif.: O'Reilly, 2013.

Python Software Foundation. *Python*, 2013. http://www.python.org/.

QlikTech. *QlikView*, 2013. http://www.qlikview.com/.

Quandl. *Quandl*, n.d. http://www.quandl.com/.

Quantcast. *QFS*, 2013. https://github.com/quantcast/qfs.

Rabinowitz, Nick, and Nate Agrin. "Seven Dirty Secrets of Data Visualisation," n.d. http://www.netmagazine.com/features/seven-dirty-secrets-data-visualisation.

Raghunathan, Balaji. *The Complete Book of Data Anonymization: From Planning to Implementation*, 2013.

Rajaraman, Anand, and Jeffrey D Ullman. *Mining of Massive Datasets*. New York, N.Y.;

Cambridge: Cambridge University Press, 2012.

Ramanujam, Srivatsan. *Pymadlib*, 2013. https://pypi.python.org/pypi/pymadlib/0.1.4.

Rao, Jun, and Sam Shah. "Data Infrastructure at LinkedIn." Technology, July 8, 2013. http://www.slideshare.net/amywtang/webscale-analyticsfb-24031929.

Raper, Simon. "Mahout for R Users." *R-bloggers*. Accessed October 10, 2013. http://www.r-bloggers.com/mahout-for-r-users/.

Rapid-I. *RapidMiner*, 2013. http://rapid-i.com/content/view/181/190/.

Redhat. *Drools*, 2013. http://www.jboss.org/drools/.

———. *JBoss Teiid*, n.d. http://www.jboss.org/teiid/.

Redmond, Stephen. *QlikView for Developers Cookbook*. Birmingham: Packt Publishing, 2013. http://public.eblib.com/EBLPublic/PublicView.do?ptiID=1192682.

Reeve, April. *Managing Data in Motion: Data Integration, Best Practice Techniques and Technologies*, 2013.

Reich, Justin. "Small Data and Big Data Should Be Best Friends." *Education Week - EdTech Researcher*, September 26, 2013. http://blogs.edweek.org/edweek/edtechresearcher/2013/03/small_data_and_big_data_should_be_best_friends.html?cmp=SOC-SHR-FB.

Reitz, Kenneth. *Elephant*, 2013. https://github.com/kennethreitz/elephant.

———. *Httpbin*, 2013. https://github.com/kennethreitz/httpbin.

———. *Requests*, 2013. https://github.com/kennethreitz/requests.

Rendgen, Sandra, Julius Wiedemann, Paolo Ciuccarelli, Richard Saul Wurman, Simon Rogers, and Nigel Holmes. *Information graphics*. Köln: Taschen, 2012.

Reshef, David, and Yakir Reshef. *MINE*, n.d. http://www.exploredata.net/.

Revolution Analytics. *RHadoop*, 2013. https://github.com/RevolutionAnalytics/RHadoop.

———. *Rhbase*, 2013. https://github.com/RevolutionAnalytics/rhbase.

———. *Rhdfs*, 2013. https://github.com/RevolutionAnalytics/rhdfs.

———. *Rmr2*, 2013. https://github.com/RevolutionAnalytics/rmr2.

Richardson, Leonard. *Beautiful Soup*, 2013. https://pypi.python.org/pypi/beautifulsoup4.

Richert, Willi, and Luis Pedro Coelho. *Building Machine Learning Systems with Python*. Birmingham, UK: Packt Pub., 2013. http://site.ebrary.com/id/10742638.

Robinson, Ian, James Webber, and Emil Efrem. *Graph Databases*. Sebastopol, Calif.: O'Reilly Media, 2013. http://search.ebscohost.com/login.aspx?direct=true&scope=site&db=nlebk&db=nlabk&AN=611703.

Roch, Eric. "Drive Smarter Decisions with Big Data Using Complex Event Processing." Technology, July 17, 2013. http://www.slideshare.net/perficientinc/drive-smarter-decisions-with-big-data-using-complex-event-processing.

Rodriguez, Marko A., and Peter Neubauer. *Constructions from Dots and Lines*. ArXiv e-print, June 11, 2010. http://arxiv.org/abs/1006.2361.

Ross, Noam. "FasteR! HigheR! StongeR! -- A Guide to Speeding Up R Code for Busy People," n.d. http://www.r-bloggers.com/faster-higher-stonger-a-guide-to-speeding-up-r-code-for-busy-people/.

Rossant, Cyrille. *Learning IPython for Interactive Computing and Data Visualization*, 2013.

Roy, Indrajit. "Distributed R for Big Data," September 23, 2013.

http://www.vertica.com/2013/02/21/presto-distributed-r-for-big-data/.

Roztočil, Jakub. *HTTPie*, 2013. https://github.com/jkbr/httpie.

RStudio. *RStudio*, 2013. https://github.com/rstudio/rstudio.

———. *Shiny*, 2013. https://github.com/rstudio/shiny.

Russell, Matthew A. *21 Recipes for Mining Twitter*. Sebastopol, Calif.: O'Reilly Media, 2011. http://proquest.safaribooksonline.com/9781449303624.

———. *Mining the Social Web*. Sebastopol, CA: O'Reilly, 2011.

Salesforce.com. *Phoenix*, 2013. https://github.com/forcedotcom/phoenix.

Sammer, Eric. *Hadoop Operations*. Sebastopol, CA: O'Reilly, 2012.

Sanfilippo, Salvatore. *Redis*, 2013. http://redis.io/.

Sarkar, D. *Pro Micosoft Hdinsight: Hadoop on Windows.* Berkeley: Apress, 2013.

Sarkar, Debarchan. *Microsoft SQL Server 2012 with Hadoop*. Packt Publishing, 2013.

Sathi, Arvind. *Big Data Analytics*. Boise: MC Press, 2012. http://site.ebrary.com/id/10621581.

Savage, Sam L. *The Flaw of Averages: Why We Underestimate Risk in the Face of Uncertainty*. Hoboken, N.J: Wiley, 2009.

Schmarzo, Bill. *Big Data Understanding How Data Powers Big Business.* John Wiley & Sons Inc, 2013.

Schmidt, Kevin. *Programming Elastic Mapreduce: Using Aws Services to Build an End-to-end Application.* [S.l.]: O'Reilly Media, 2013.

Schrage, Michael. "The Question All Smart Visualizations Should Ask." *HBR Blog Network - Harvard Business Review*, September 27, 2013. http://blogs.hbr.org/2013/03/the-question-all-smart-visualizations/.

ScraperWiki. *Custard*, 2013. https://github.com/scraperwiki/custard.

———. *ScraperWiki*, 2013. https://github.com/scraperwiki/scraperwiki-python.

Scrapy project. *Scrapely*, 2013. https://github.com/scrapy/scrapely.

SeatGeek. *FuzzyWuzzy*, 2013. https://github.com/seatgeek/fuzzywuzzy.

Sebastian-Coleman, Laura. *Measuring Data Quality for Ongoing Improvement: a Data Quality Assessment Framework*. Waltham, MA: Morgan Kaufmann, 2013.

Sedgewick, Robert, and Kevin Daniel Wayne. *Algorithms*. Upper Saddle River, NJ: Addison-Wesley, 2011.

Segaran, Toby. *Programming Collective Intelligence: Building Smart Web 2.0 Applications*. Beijing; Sebastapol [Calif.]: O'Reilly, 2007.

Segaran, Toby, and Jamie Taylor. *Programming the Semantic Web*. Beijing; Sebastopol, CA: O'Reilly, 2009.

Segel, E. "How to Tell Stories with Data (really)." Technology, November 5, 2011. http://www.slideshare.net/kris77chan/edward-segel-interactivestorytelling.

Segel, E., and J. Heer. "Narrative Visualization: Telling Stories with Data." *IEEE Transactions on Visualization and Computer Graphics* 16, no. 6 (2010): 1139–1148. doi:10.1109/TVCG.2010.179.

Sematext Group. *HBaseHUT*, 2013. https://github.com/sematext/HBaseHUT.

———. *HBaseWD*, 2013. https://github.com/sematext/HBaseWD.

SenseiDB. *SenseiDB*. Accessed October 12, 2013. https://github.com/senseidb/sensei.

Services, Amazon Web. *Amazon Elastic MapReduce Developer Guide*, n.d. http://amzn.com/B007US6CIO.

Shah, Sam, Craig AN Soules, Gregory R. Ganger, and Brian D. Noble. "Using Provenance to Aid in Personal File Search." In *USENIX Annual Technical Conference*, 171–184, 2007.

http://static.usenix.org/events/usenix07/tech/full_papers/shah/shah_html/.

Shanklin, Carter. "How To Perform Spatial Analytics with Hive and Hadoop." *Hortonworks*,
    September 23, 2013.
    http://hortonworks.com/blog/how-to-perform-spatial-analytics-with-hive-and-hadoop/.

Sharethrough. *Elasticity*, 2013. https://github.com/rslifka/elasticity.

Sheikh, Nauman. *Implementing Analytics: a Blueprint for Design, Development, and Adoption.*
    [S.l.]: Morgan Kaufmann, 2013.

Shekhar, Jayant. "Building Recommendation Platforms with Hadoop." September 23, 2013.
    http://strataconf.com/strata2013/public/schedule/detail/27475.

Shron, Max. "Data's Missing Ingredient? Rhetoric. - Strata," October 2, 2013.
    http://strata.oreilly.com/2013/04/datas-missing-ingredient-rhetoric.html.

Shutterstock. *Rickshaw*, 2013. https://github.com/shutterstock/rickshaw.

Sid Anand. "Building a Modern Website for Scale (QCon NY 2013)." Design, June 14, 2013.
    http://www.slideshare.net/r39132/q-con-ny2013modernwebsitescalabilityfinal-22989785.
———. "LinkedIn Data Infrastructure Slides (Version 2)." June 21, 2012.
    http://www.slideshare.net/r39132/linkedin-data-infrastructure-slides-version-2-13394853.

Siegel, Eric. *Predictive Analytics: The Power to Predict Who Will Click, Buy, Lie, or Die*.
    Hoboken, N.J.: Wiley, 2013.

Sigelman, Benjamin H., Luiz Andre Barroso, Mike Burrows, Pat Stephenson, Manoj Plakal,
    Donald Beaver, Saul Jaspan, and Chandan Shanbhag. "Dapper, a Large-scale Distributed
    Systems Tracing Infrastructure." *Google Research* (2010).
    http://sites.google.com/site/atlediadema/training-documents/Google%20Technical%20Re
    port%20-%20dapper,%202010.pdf.

Sigoure, Benoit, and Viral Bajaria. *AsyncHBase*, 2013.
    https://github.com/OpenTSDB/asynchbase.

Silva, Francisco Javier Blanco. *Learning Scipy for Numerical and Scientific Computing.* [S.l.]:
    Packt Publishing Limited, 2013.

Simon, Phil. *Too Big to Ignore: The Business Case for Big Data*, 2013.

Singh, Chandeep. "Ingesting XML in Hive Using XPath." *The Data Stack*, September 26, 2013.
    https://communities.intel.com/community/datastack/blog/2013/08/15/hadoop-tutorials-in
    gesting-xml-in-hive-using-xpath.

SlamData. *Precog Platform*, 2013. https://github.com/precog/platform.

Smeets, Bart. "Creating a Business Dashboard in R," September 26, 2013.
    http://www.r-bloggers.com/creating-a-business-dashboard-in-r/.

Smiley, David, and Eric Pugh. *Apache Solr 3 Enterprise Search Server Enhance Your Search
    with Faceted Navigation, Result Highlighting, Relevancy Ranked Sorting, and More*.
    Birmingham: Packt Pub., 2011. http://site.ebrary.com/id/10517472.

Smith, Tim. "aRrgh - A Newcomer's (angry) Guide to Data Types in R." *GitHub*, October 4,
    2013. https://github.com/tdsmith/aRrgh.

Smolan, Rick, and Jennifer Erwitt. *The Human Face of Big Data*. Sausalito, Calif.: Against All
    Odds Productions, 2012.

Snowplow Analytics. *Snowplow*, 2013. https://github.com/snowplow/snowplow.

Soares, Sunil. *Big Data Governance an Emerging Imperative*. Boise: MC Press, 2012.
    http://site.ebrary.com/id/10621576.

Solid IT. "DB-Engines Ranking," October 4, 2013. http://db-engines.com/en/ranking.

Soztutar, Enis. "Apache HBase and HDFS - Understanding Filesystem Usa..." Technology, July

9, 2013. http://www.slideshare.net/cloudera/internals-session-5.

spark-project. "Deep Dive with Spark Streaming - Tathagata   Das - Spark Meetup 2013..." June 23, 2013.
http://www.slideshare.net/spark-project/deep-divewithsparkstreaming-tathagatadasspark meetup20130617.

Spotify. *Luigi*, 2013. https://github.com/spotify/luigi.

———. *Snakebite*, 2013. https://github.com/spotify/snakebite.

Spring Data: Modern Data Access for Enterprise Java. Sebastopol, CA: O'Reilly, 2013.

SpringSource. *Spring for Apache Hadoop*, 2013.
https://github.com/spring-projects/spring-hadoop.

Square. *Crossfilter*, 2013. https://github.com/square/crossfilter.

Sripati, Praveen. "Hadoop Tips: Debugging a Hadoop MapReduce Program in Eclipse," n.d.
http://www.thecloudavenue.com/2012/10/debugging-hadoop-mapreduce-program-in.html.

Stacey, Mark, Joe Salvatore, and Adam Jorgensen. *Visual Intelligence: Microsoft Tools and Techniques for Visualizing Data*. Indianapolis, IN: John Wiley & Sons, Inc., 2013.

Stado Global Development Group. *Stado*, 2013. https://launchpad.net/stado.

Stanford Visualization Group. *Data Wrangler*, n.d. http://vis.stanford.edu/wrangler/app/.

Story, Rob. *Bearcart*, 2013. https://github.com/wrobstory/bearcart.

———. *Vincent*, 2013. https://github.com/wrobstory/vincent.

Strata Conference. *Strata Conference Santa Clara 2013: Complete Video Compilation*. [Calif?].: O'Reilly Media, 2013. http://my.safaribooksonline.com/video/databases/9781449365394.

Strathausen, Johann Philipp. *Graph Dracula*, 2013. https://github.com/strathausen/dracula.

Stripe. *MoSQL*, n.d. https://github.com/stripe/mosql.

Sturm, Brian W. "The Enchanted Imagination: Storytelling's Power To Entrance Listeners." *School Library Media Research* 2 (1999).
http://www.eric.ed.gov/ERICWebPortal/recordDetail?accno=EJ593526.

Sumbaly, Roshan, Jay Kreps, Lei Gao, Alex Feinberg, Chinmay Soman, and Sam Shah. "Serving Large-scale Batch Computed Data with Project Voldemort." In *Proceedings of the 10th USENIX Conference on File and Storage Technologies*, 18–18, 2012.
http://dl.acm.org/citation.cfm?id=2208479.

Sumbaly, Roshan, Jay Kreps, and Sam Shah. "The Big Data Ecosystem at LinkedIn."
1125–1134, 2013. http://dl.acm.org/citation.cfm?id=2463707.

Sweet, Nathan. *Kryo*, 2013. http://code.google.com/p/kryo/.

SymPy Development Team. *SymPy*, 2013. http://sympy.org.

SYSTAP. *Bigdata*, n.d.
http://sourceforge.net/apps/mediawiki/bigdata/index.php?title=Main_Page.

Tableau Software. *Tableau*, 2013. http://www.tableausoftware.com.

Talend. *Talend Big Data*, 2013. http://www.talend.com/products/big-data.

Tavory, Ami. *DAGPype*, 2013. https://pypi.python.org/pypi/DAGPype.

Taylor, James. *Decision Management Systems a Practical Guide to Using Business Rules and Predictive Analytics*. [Boston, Mass.]: IBM Press/Pearson Education, 2012.
http://www.books24x7.com/marc.asp?bookid=45443.

Teetor, Paul. *R Cookbook*. 1st ed. Beijing ; Sebastopol, CA: O'Reilly, 2011.

The Gephi Consortium. *Gephi*, n.d. https://gephi.org/.

The Legacy of Fischer Black. New York: Oxford University Press, 2005.

The Miso Project. *D3.chart*, 2013. https://github.com/misoproject/d3.chart.

———. *Dataset.js*, 2013. https://github.com/misoproject/dataset.

The OpenTSDB Authors, Benoit Sigoure, Geoffrey Anderson, Ion Savin, and Will Moss. *OpenTSDB*, 2013. http://opentsdb.net/.

The PostgreSQL Global Development Group. *PostgreSQL*, 2013. http://www.postgresql.org/.

The PyData Development Team. *Pandas*, 2013. https://github.com/pydata/pandas.

The R Foundation. *The R Project*, 2013. http://www.r-project.org/.

The Saddle Development Team. *SADDLE*, 2013. https://github.com/saddle/saddle.

theguardian.com. "All Our Datasets: The Complete Index," September 26, 2013. http://www.theguardian.com/news/datablog/interactive/2013/jan/14/all-our-datasets-index.

Thorp, Jer. "Visualization as Process, Not Output," n.d. http://blogs.hbr.org/cs/2013/04/visualization_as_process.html.

Ting, Kathleen, and Jarek Jarcec Cecho. *Apache Sqoop Cookbook*. Sebastopol, CA: O'Reilly Media, 2013.

TinkerPop. *Gremlin*, 2013. https://github.com/tinkerpop/gremlin.

Tosi, Sandro. *Matplotlib for Python Developers: Build Remarkable Publication Quality Plots the Easy Way*. Birmingham [u.a.]: Packt Publ., 2009.

Tran, Nga, Sreenath Bodagala, and Jaimin Dave. "Designing Query Optimizers for Big Data Problems of The Future." *Proceedings of the VLDB Endowment* 6, no. 11 (2013). http://db.disi.unitn.eu/pages/VLDBProgram/pdf/flash/Vertica.pdf.

Trifacta. *Triflow*, 2013. https://github.com/trifacta/triflow.

———. *Vega*, 2013. https://github.com/trifacta/vega.

Trudeau, Richard J, and Richard J Trudeau. *Introduction to Graph Theory*. New York: Dover Pub., 1993.

Tsvetovat, Maksim. *SnowWhite*, 2013. https://github.com/maksim2042/snowwhite.

Tsvetovat, Maksim, and Alexander Kouznetsov. *Social Network Analysis for Startups: [finding Connections on the Social Web]*. Sebastopol, CA: O'Reilly Media, 2011.

Tufte, Edward R. *Visual Explanations: Images and Quantities, Evidence and Narrative*. Cheshire, Conn: Graphics Press, 1997.

Tupper, Charles. *Data Architecture: From Zen to Reality*. Burlington, MA: Morgan Kaufmann/Elsevier, 2011.

Turkington, Garry. *Hadoop Beginner's Guide*. Birmingham: Packt Pub., 2013. http://public.eblib.com/EBLPublic/PublicView.do?ptiID=1140151.

Twitter. *Ambrose*, 2013. https://github.com/twitter/ambrose.

———. *Chill*, 2013. https://github.com/twitter/chill.

———. *Elephant Twin*, 2013. https://github.com/twitter/elephant-twin.

———. *Finagle*, 2013. https://github.com/twitter/finagle.

———. *Hadoop-LZO*, 2013. https://github.com/twitter/hadoop-lzo.

———. *HDFS-DU*, 2013. https://github.com/twitter/hdfs-du.

———. *hRaven*, 2013. https://github.com/twitter/hraven.

———. *Scalding*, 2013. https://github.com/twitter/scalding.

———. *Snowflake*, 2013. https://github.com/twitter/snowflake.

———. *Summingbird*, 2013. https://github.com/twitter/summingbird.

———. *Zipkin*, 2013. https://github.com/twitter/zipkin.

Twitter, and Cloudera. *Parquet*, 2013. http://parquet.io/.

Twitter, Kevin Weil, Dmitriy Ryaboy, Chuang Liu, and Florian Leibert. *Elephant Bird*, 2013.
        https://github.com/kevinweil/elephant-bird.
Typesafe. *Akka*, 2013. http://akka.io/.
Üngür, Osman. *Simmetrica*, 2013. https://github.com/o/simmetrica.
University of Athens. *madIS*, n.d. http://code.google.com/p/madis/.
University of Ljubljana. *Orange*, 2013. http://orange.biolab.si/.
University of Waikato. *Weka 3*, 2013. http://www.cs.waikato.ac.nz/%7Eml/weka/.
Uruchurtu, Linda. "38 Tools For Beautiful Data Visualisations," September 26, 2013.
        http://www.elisa-dbi.co.uk/blog/38-tools-for-beautiful-data-visualisations/.
Valiant, Leslie G. "A Bridging Model for Parallel Computation." *Commun. ACM* 33, no. 8
        (August 1990): 103–111. doi:10.1145/79173.79181.
Vantomme, Jan. *Processing 2 Creative Programming Cookbook: Over 90 Highly-effective
        Recipes to Unleash Your Creativity with Interactive Art, Graphics, Computer Vision, 3D,
        and More*. Birmingham: Packt Pub., 2012.
        http://search.ebscohost.com/login.aspx?direct=true&scope=site&db=nlebk&db=nlabk&A
        N=492090.
Venner, Jason. *Pro Hadoop*. Berkeley, CA: Apress, 2009.
        http://public.eblib.com/EBLPublic/PublicView.do?ptiID=478209.
Vernat, Emeric. *JavaMelody*, 2013. https://code.google.com/p/javamelody/.
Vertica. *Vertica Avro Parser*, 2013. https://github.com/vertica/Vertica-Extension-Packages.
———. *Vertica-Hadoop-Connector*, 2013.
        https://github.com/vertica/Vertica-Hadoop-Connector.
Verwoerd, Jos. "Data, Data, Data: Thousands of Public Data Sources," n.d.
        http://blog.bigml.com/2013/02/28/data-data-data-thousands-of-public-data-sources/.
Villellas, Oscar. "Faster, More Memory Efficient SQL Queries via IOPro," n.d.
        http://continuum.io/blog/iopro-pyodbc-performance.
Voldemort Development Team. *Voldemort*, 2013. https://github.com/voldemort/voldemort.
Waber, Ben. *People Analytics How Social Sensing Technology Will Transform Business and
        What It Tells Us About the Future of Work*. Upper Saddle River, N.J.: FT Press, 2013.
        http://proquest.safaribooksonline.com/?fpi=9780133158342.
Wampler, Dean. "Hadoop Data Warehousing with Hive." September 23, 2013.
        http://strataconf.com/strata2013/public/schedule/detail/26899.
Wang, R. "Ray", and 2012. "What a Big-Data Business Model Looks Like." *Harvard Business
        Review*, October 5, 2013. http://blogs.hbr.org/2012/12/what-a-big-data-business-model/.
Warden, Pete. *Big Data Glossary*. Sebastopol, CA: O'Reilly, 2011.
        http://proquest.safaribooksonline.com/9781449315085.
———. *Data Source Handbook*. 1st ed. Sebastopol, CA ; Beijing [China]: O'Reilly, 2011.
Warren, Norman P, Mariano Teixeira Neto, Stacia Misner, Ivan Sanders, and Scott A Helmers.
        Business Intelligence in Microsoft SharePoint 2013. Redmond, Wash.: Microsoft, 2013.
White, Brandyn. *Hadoopy*, 2013. https://github.com/bwhite/hadoopy.
White, Tom. *Hadoop: The Definitive Guide*. Farnham; Sebastopol, Calif.: O'Reilly, 2012.
Wickham, Hadley. "Advanced R Development," October 1, 2013. http://adv-r.had.co.nz/.
———. "Bin-summarise-smooth: A Framework for Visualising Large Data" (September 26,
        2013).
        http://www.polisci.ucla.edu/workshops/ap-workshop-papers/Hadley%20Wickham%20Pa
        per-%2004-08-13.pdf.

———. *Ggplot2*, 2013. http://ggplot2.org/.

———. *Lubridate*, 2013. https://github.com/hadley/lubridate.

———. *Plyr*, 2013. https://github.com/hadley/plyr.

———. "The Split-apply-combine Strategy for Data Analysis." *Journal of Statistical Software* 40, no. 1 (2011): 1–29.

William, Sam. "Building Scalable Big Data Infrastructure Using Open Source Software." September 23, 2013. http://strataconf.com/strata2013/public/schedule/detail/27419.

Wilson, Chris. "Binify + D3 = Gorgeous Honeycomb Maps," October 4, 2013. http://mechanicalscribe.com/notes/binify-d3-topojson-tutorial/.

Wilson, Greg. "Software Carpentry: A Rational Computing Process: How and Why to Fake It," October 5, 2013. http://software-carpentry.org/blog/2013/05/rational-computing-process.html.

Wise, Lyndsay. *Using Open Source Platforms for Business Intelligence: Avoid Pitfalls and Maximize ROI*. Waltham, MA: Morgan Kaufmann, 2012.

Wittenauer, Allen. "Hadoop Operations at LinkedIn." March 21, 2013. http://www.slideshare.net/allenwittenauer/2013-hadoopsummitemea?ref=http://eventifier.co/event/hadoopsummit2013/slides.

———. "Hadoop Performance at LinkedIn." January 19, 2013. http://www.slideshare.net/allenwittenauer/2012-lihadoopperf.

Wong, Dona M. *The Wall Street Journal Guide to Information Graphics: The Dos and Don'ts of Presenting Data, Facts, and Figures*. 1st ed. New York: W.W. Norton & Co, 2010.

Wood, David. *Linked Data*. [S.l.]: O'Reilly Media, 2013. http://www.manning.com/dwood/.

Xavier, Amatriain, and Justin Basilico. "The Netflix Tech Blog: System Architectures for Personalization and Recommendation," October 5, 2013. http://techblog.netflix.com/2013/03/system-architectures-for.html.

Xin, Reynold. "An Introduction to the Berkeley Data Analytics Stack (BDAS) Featuring Spark, Spark Streaming, and Shark." September 22, 2013. http://strataconf.com/strata2013/public/schedule/detail/27438.

Xin, Reynold, Josh Rosen, Matei Zaharia, Michael J. Franklin, Scott Shenker, and Ion Stoica. "Shark: SQL and Rich Analytics at Scale." *arXiv Preprint arXiv:1211.6176* (2012). http://arxiv.org/abs/1211.6176.

Yahoo. *Storm-yarn*, 2013. https://github.com/yahoo/storm-yarn.

Yang, Fangjin. "Fast, Cheap, and 98% Right: Cardinality Estimation for Big Data." *Metamarkets*, October 4, 2013. http://metamarkets.com/2012/fast-cheap-and-98-right-cardinality-estimation-for-big-data/.

Yau, Nathan. *Data Points Visualization That Means Something*. Hoboken, N.J.: Wiley, 2013. http://lib.myilibrary.com?id=471991.

———. *Visualize This: The FlowingData Guide to Design, Visualization, and Statistics*. Indianapolis, Ind: Wiley, 2011. http://www.myilibrary.com?id=317758.

Yelp. *Mrjob*, 2013. https://github.com/Yelp/mrjob.

Yhat. *Pandasql*, 2013. https://github.com/yhat/pandasql.

Zadrozny, Peter. *Big Data Analytics Using Splunk*. [New York, N.Y.]: Apress, 2013.

Zaitlen, Benjamin. "Wakari and Big Finance," October 5, 2013. http://continuum.io/blog/wakari-and-big-finance.

Zeyliger, Philip. "Debug Servlets, or 'HTTP Won; Use It' - Philz," n.d.

http://omel.ette.org/blog/2013/02/06/debug-servlets/.

———. "Tricks for Distributed System Debugging and Diagnosis." September 23, 2013. http://strataconf.com/strata2013/public/schedule/detail/27194.

Zhai, Ke, Jordan Boyd-Graber, Nima Asadi, and Mohamad L. Alkhouja. "Mr. LDA: A Flexible Large Scale Topic Modeling Package Using Variational Inference in MapReduce." 879–888, 2012. http://dl.acm.org/citation.cfm?id=2187955.

Zikopoulos, Paul, Dirk deRoos, Krishnan Parasuraman, Thomas Deutsch, James Giles, and David Corrigan. *Harness the Power of Big Data The IBM Big Data Platform*. McGraw Hill Professional, 2012.