

How Data Science Is Transforming Health Care

Solving the Wanamaker Dilemma



**Tim O'Reilly, Julie Steele,
Mike Loukides & Colin Hill**



Strata+ Hadoop

WORLD

Make Data Work
strataconf.com

Presented by O'Reilly and Cloudera,
Strata + Hadoop World is where
cutting-edge data science and new
business fundamentals intersect—
and merge.

- Learn business applications of data technologies
- Develop new skills through trainings and in-depth tutorials
- Connect with an international community of thousands who work with data

How Data Science Is Transforming Health Care

*Tim O'Reilly, Mike Loukides,
Julie Steele, and Colin Hill*

O'REILLY®

Beijing • Cambridge • Farnham • Köln • Sebastopol • Tokyo

How Data Science Is Transforming Health Care

by Tim O'Reilly, Mike Loukides, Julie Steele, and Colin Hill

Copyright © 2012 O'Reilly Media. All rights reserved.

Printed in the United States of America.

Published by O'Reilly Media, Inc., 1005 Gravenstein Highway North, Sebastopol, CA 95472.

O'Reilly books may be purchased for educational, business, or sales promotional use. Online editions are also available for most titles (<http://my.safaribooksonline.com>). For more information, contact our corporate/institutional sales department: (800) 998-9938 or corporate@oreilly.com.

Cover Designer: Karen Montgomery

Interior Designer: David Futato

August 2012: First Edition

Revision History for the First Edition:

2012-08-20 First release

See <http://oreilly.com/catalog/errata.csp?isbn=9781449345006> for release details.

Nutshell Handbook, the Nutshell Handbook logo, and the O'Reilly logo are registered trademarks of O'Reilly Media, Inc.

Many of the designations used by manufacturers and sellers to distinguish their products are claimed as trademarks. Where those designations appear in this book, and O'Reilly Media, Inc. was aware of a trademark claim, the designations have been printed in caps or initial caps.

While every precaution has been taken in the preparation of this book, the publisher and authors assume no responsibility for errors or omissions, or for damages resulting from the use of the information contained herein.

ISBN: 978-1-449-34500-6

Table of Contents

1. Introduction.....	1
2. Making Health Care More Effective.....	5
3. More Data, More Sources.....	9
4. Paying for Results.....	11
5. Enabling Data.....	15
6. Building the Health Care System We Want.....	19
7. Recommended Reading.....	21

CHAPTER 1

Introduction

The best minds of my generation are thinking about how to make people click ads.

— JEFF HAMMERBACHER
early Facebook employee

Work on stuff that matters.

— TIM O'REILLY

In the early days of the 20th century, department store magnate John Wanamaker famously said, “I know that half of my advertising doesn’t work. The problem is that I don’t know which half.”

The consumer Internet revolution was fueled by a search for the answer to Wanamaker’s question. Google AdWords and the pay-per-click model began the transformation of a business in which advertisers paid for ad impressions into one in which they pay for results. “Cost per thousand impressions” (CPM) was outperformed by “cost per click” (CPC), and a new industry was born. It’s important to understand why CPC outperformed CPM, though. Superficially, it’s because Google was able to track when a user clicked on a link, and was therefore able to bill based on success. But billing based on success doesn’t fundamentally change anything unless you can also change the success rate, and that’s what Google was able to do. By using data to understand each user’s behavior, Google was able to place advertisements that an individual was likely to click. They knew “which half” of their advertising was more likely to be effective, and didn’t bother with the rest.

Since then, data and predictive analytics have driven ever deeper insight into user behavior such that companies like Google, Facebook, Twitter, and LinkedIn are fundamentally data companies. And data isn't just transforming the consumer Internet. It is transforming finance, design, and manufacturing—and perhaps most importantly, health care. How is data science transforming health care? There are many ways in which health care is changing, and needs to change. We're focusing on one particular issue: the problem Wanamaker described when talking about his advertising. How do you make sure you're spending money effectively? Is it possible to know what will work in advance?

Too often, when doctors order a treatment, whether it's surgery or an over-the-counter medication, they are applying a "standard of care" treatment or some variation that is based on their own intuition, effectively hoping for the best. The sad truth of medicine is that we don't always understand the relationship between treatments and outcomes. We have studies to show that various treatments will work more often than placebos; but, like Wanamaker, we know that much of our medicine doesn't work for half of our patients, we just don't know which half. At least, not in advance. One of data science's many promises is that, if we can collect enough data about medical treatments and use that data effectively, we'll be able to predict more accurately which treatments will be effective for which patient, and which treatments won't.

A better understanding of the relationship between treatments, outcomes, and patients will have a huge impact on the practice of medicine in the United States. Health care is expensive. The U.S. spends over \$2.6 trillion on health care every year, an amount that constitutes a serious fiscal burden for government, businesses, and our society as a whole. These costs include over \$600 billion of unexplained variations in treatments: treatments that cause no differences in outcomes, or even make the patient's condition worse. We have reached a point at which our need to understand treatment effectiveness has become vital—to the health care system and to the health and sustainability of the economy overall.

Why do we believe that data science has the potential to revolutionize health care? After all, the medical industry has had data for generations: clinical studies, insurance data, hospital records. But the health care industry is now awash in data in a way that it has never been before: from biological data such as gene expression, next-

generation DNA sequence data, **proteomics**, and **metabolomics**, to clinical data and health outcomes data contained in ever more prevalent electronic health records (EHRs) and longitudinal drug and medical claims. We have entered a new era in which we can work on massive datasets effectively, combining data from clinical trials and direct observation by practicing physicians (the records generated by our \$2.6 trillion of medical expense). When we combine data with the resources needed to work on the data, we can start asking the important questions, the Wanamaker questions, about what treatments work and for whom.

The opportunities are huge: for entrepreneurs and data scientists looking to put their skills to work disrupting a large market, for researchers trying to make sense out of the flood of data they are now generating, and for existing companies (including health insurance companies, biotech, pharmaceutical, and medical device companies, hospitals and other care providers) that are looking to remake their businesses for the coming world of outcome-based payment models.

CHAPTER 2

Making Health Care More Effective

What, specifically, does data allow us to do that we couldn't do before? For the past 60 or so years of medical history, we've treated patients as some sort of an average. A doctor would diagnose a condition and recommend a treatment based on what worked for most people, as reflected in large clinical studies. Over the years, we've become more sophisticated about what that average patient means, but that same statistical approach didn't allow for differences between patients. A treatment was deemed effective or ineffective, safe or unsafe, based on double-blind studies that rarely took into account the differences between patients. With the data that's now available, we can go much further. The exceptions to this are relatively recent and have been dominated by cancer treatments, the first being Herceptin for breast cancer in women who over-express the Her2 receptor. With the data that's now available, we can go much further for a broad range of diseases and interventions that are not just drugs but include surgery, disease management programs, medical devices, patient adherence, and care delivery.

For a long time, we thought that Tamoxifen was roughly 80% effective for breast cancer patients. But now we know much more: we know that it's **100% effective in 70% to 80% of the patients**, and ineffective in the rest. That's not word games, because we can now use genetic markers to tell whether it's likely to be effective or ineffective for any given patient, and we can tell in advance whether to treat with Tamoxifen or to try something else.

Two factors lie behind this new approach to medicine: a different way of using data, and the availability of new kinds of data. It's not just

stating that the drug is effective on most patients, based on trials (indeed, 80% is an enviable success rate); it's using artificial intelligence techniques to divide the patients into groups and then determine the difference between those groups. We're not asking whether the drug is effective; we're asking a fundamentally different question: "for which patients is this drug effective?" We're asking about the patients, not just the treatments. A drug that's only effective on 1% of patients might be very valuable if we can tell who that 1% is, though it would certainly be rejected by any traditional clinical trial.

More than that, asking questions about patients is only possible because we're using data that wasn't available until recently: DNA sequencing was only invented in the mid-1970s, and is only now coming into its own as a medical tool. What we've seen with Tamoxifen is as clear a solution to the Wanamaker problem as you could ask for: we now know when that treatment will be effective. If you can do the same thing with millions of cancer patients, you will both improve outcomes and save money.

Dr. Lukas Wartman, a cancer researcher who was himself diagnosed with terminal leukemia, **was successfully treated** with sunitinib, a drug that was only approved for kidney cancer. Sequencing the genes of both the patient's healthy cells and cancerous cells led to the discovery of a protein that was out of control and encouraging the spread of the cancer. The gene responsible for manufacturing this protein could potentially be inhibited by the kidney drug, although it had never been tested for this application. This unorthodox treatment was surprisingly effective: Wartman is now in remission.

While this treatment was exotic and expensive, what's important isn't the expense but the potential for new kinds of diagnosis. The price of gene sequencing has been plummeting; it will be a common **doctor's office procedure** in a few years. And through Amazon and Google, you can now "rent" a cloud-based **supercomputing cluster** that can solve huge analytic problems for a few hundred dollars per hour. What is now exotic inevitably becomes routine.

But even more important: we're looking at a completely different approach to treatment. Rather than a treatment that works 80% of the time, or even 100% of the time for 80% of the patients, a treatment might be effective for a small group. It might be entirely specific to the individual; the next cancer patient may have a different protein that's out of control, an entirely different genetic cause for the disease.

Treatments that are specific to one patient don't exist in medicine as it's currently practiced; how could you ever do an FDA trial for a medication that's only going to be used once to treat a certain kind of cancer?

Foundation Medicine is at the forefront of this new era in cancer treatment. They use next-generation DNA sequencing to discover DNA sequence mutations and deletions that are currently used in standard of care treatments, as well as many other actionable mutations that are tied to drugs for other types of cancer. They are creating a patient-outcomes repository that will be the fuel for discovering the relation between mutations and drugs. Foundation has identified DNA mutations in 50% of cancer cases for which drugs exist (information via a private communication), but are not currently used in the standard of care for the patient's particular cancer.

The ability to do large-scale computing on genetic data gives us the ability to understand the origins of disease. If we can understand why an anti-cancer drug is effective (what specific proteins it affects), and if we can understand what genetic factors are causing the cancer to spread, then we're able to use the tools at our disposal much more effectively. Rather than using imprecise treatments organized around symptoms, we'll be able to target the actual causes of disease, and design treatments tuned to the biology of the specific patient. Eventually, we'll be able to treat 100% of the patients 100% of the time, precisely because we realize that each patient presents a unique problem.

Personalized treatment is just one area in which we can solve the Wanamaker problem with data. Hospital admissions are extremely expensive. Data can make hospital systems more efficient, and avoid preventable complications such as blood clots and hospital readmissions. It can also help address the challenge of **health care hot-spotting** (a term coined by Atul Gawande): finding people who use an inordinate amount of health care resources. By looking at data from hospital visits, **Dr. Jeffrey Brenner** of Camden, NJ, was able to determine that "**just one per cent of the hundred thousand people who made use of Camden's medical facilities accounted for thirty per cent of its costs.**" Furthermore, many of these people came from only two apartment buildings. Designing more effective medical care for these patients was difficult; it doesn't fit our health insurance system, the patients are often dealing with many serious medical issues (addiction and obesity are frequent complications), and have trouble trusting

doctors and social workers. It's counter-intuitive, but spending more on some patients now results in spending less on them when they become really sick. While it's a work in progress, it looks like building appropriate systems to target these high-risk patients and treat them before they're hospitalized will bring significant savings.

Many poor health outcomes are attributable to patients who don't take their medications. [Eliza](#), a Boston-based company started by [Alexandra Drane](#), has pioneered approaches to improve compliance through interactive communication with patients. Eliza improves patient drug compliance by tracking which types of reminders work on which types of people; it's similar to the way companies like Google target advertisements to individual consumers. By using data to analyze each patient's behavior, Eliza can generate reminders that are more likely to be effective. The results aren't surprising: if patients take their medicine as prescribed, they are more likely to get better. And if they get better, they are less likely to require further, more expensive treatment. Again, we're using data to solve Wanamaker's problem in medicine: we're spending our resources on what's effective, on appropriate reminders that are mostly to get patients to take their medications.

CHAPTER 3

More Data, More Sources

The examples we've looked at so far have been limited to traditional sources of medical data: hospitals, research centers, doctor's offices, insurers. The Internet has enabled the formation of patient networks aimed at sharing data. Health social networks now are some of the largest patient communities. As of November 2011, **PatientsLikeMe** has over 120,000 patients in 500 different condition groups; **ACOR** has over 100,000 patients in 127 cancer support groups; **23andMe** has over 100,000 members in their genomic database; and diabetes health social network **SugarStats** has over 10,000 members. These are just the larger communities, thousands of small communities are created around rare diseases, or even uncommon experiences with common diseases. All of these communities are generating data that they voluntarily share with each other and the world.

Increasingly, what they share is not just anecdotal, but includes an array of clinical data. For this reason, these groups are being recruited for large-scale crowdsourced clinical outcomes research.

Thanks to ubiquitous data networking through the mobile network, we can take several steps further. In the past two or three years, there's been a flood of personal fitness devices (such as the **Fitbit**) for monitoring your personal activity. There are mobile apps for taking your pulse, and an **iPhone attachment for measuring your glucose**. There has been talk of mobile applications that would constantly listen to a patient's speech and detect changes that might be the precursor for a stroke, or would use the accelerometer to report falls. **Tanzeem Choudhury** has developed an app called **Be Well** that is intended primarily for victims of depression, though it can be used by anyone.

Be Well monitors the user's sleep cycles, the amount of time they spend talking, and the amount of time they spend walking. The data is scored, and the app makes appropriate recommendations, based both on the individual patient and data collected across all the app's users.

Continuous monitoring of critical patients in hospitals has been normal for years; but we now have the tools to monitor patients constantly, in their home, at work, wherever they happen to be. And if this sounds like big brother, at this point most of the patients are willing. We don't want to transform our lives into hospital experiences; far from it! But we can collect and use the data we constantly emit, our "data exhaust," to maintain our health, to become conscious of our behavior, and to detect oncoming conditions before they become serious. The most effective medical care is the medical care you avoid because you don't need it.

CHAPTER 4

Paying for Results

Once we're on the road toward more effective health care, we can look at other ways in which Wanamaker's problem shows up in the medical industry. It's clear that we don't want to pay for treatments that are ineffective. Wanamaker wanted to know which part of his advertising was effective, not just to make better ads, but also so that he wouldn't have to buy the advertisements that wouldn't work. He wanted to pay for results, not for ad placements. Now that we're starting to understand how to make treatment effective, now that we understand that it's more than rolling the dice and hoping that a treatment that works for a typical patient will be effective for you, we can take the next step: Can we change the underlying incentives in the medical system? Can we make the system better by paying for results, rather than paying for procedures?

It's shocking just how badly the incentives in our current medical system are aligned with outcomes. If you see an orthopedist, you're likely to get an MRI, most likely at a facility owned by the orthopedist's practice. On one hand, it's good medicine to know what you're doing before you operate. But how often does that MRI result in a different treatment? How often is the MRI required just because it's part of the protocol, when it's perfectly obvious what the doctor needs to do? Many men have had **PSA tests for prostate cancer**; but in most cases, aggressive treatment of prostate cancer is a bigger risk than the disease itself. Yet the test itself is a significant profit center. Think again about Tamoxifen, and about the pharmaceutical company that makes it. In our current system, what does "100% effective in 80% of the patients" mean, except for a 20% loss in sales? That's because the drug

company is paid for the treatment, not for the result; it has no financial interest in whether any individual patient gets better. (Whether a statistically significant number of patients has side-effects is a different issue.) And at the same time, bringing a new drug to market is very expensive, and might not be worthwhile if it will only be used on the remaining 20% of the patients. And that's assuming that one drug, not two, or 20, or 200 will be required to treat the unlucky 20% effectively.

It doesn't have to be this way.

In the U.K., **Johnson & Johnson**, faced with the possibility of losing reimbursements for their multiple myeloma drug Velcade, agreed to refund the money for patients who did not respond to the drug. Several other pay-for-performance drug deals have followed since, paving the way for the ultimate transition in pharmaceutical company business models in which their product is health outcomes instead of pills. Such a transition would rely more heavily on real-world outcome data (are patients actually getting better?), rather than controlled clinical trials, and would use molecular diagnostics to create personalized "treatment algorithms." Pharmaceutical companies would also focus more on drug compliance to ensure health outcomes were being achieved. This would ultimately align the interests of drug makers with patients, their providers, and payors.

Similarly, rather than paying for treatments and procedures, can we pay hospitals and doctors for results? That's what **Accountable Care Organizations** (ACOs) are about. ACOs are a leap forward in business model design, where the provider shoulders any financial risk. ACOs represent a new framing of the much maligned HMO approaches from the '90s, which did not work. HMOs tried to use statistics to predict and prevent unneeded care. The ACO model, rather than controlling doctors with what the data says they "should" do, uses data to measure how each doctor performs. Doctors are paid for successes, not for the procedures they administer. The main advantage that the ACO model has over the HMO model is how good the data is, and how that data is leveraged. The ACO model aligns incentives with outcomes: a practice that owns an MRI facility isn't incentivized to order MRIs when they're not necessary. It is incentivized to use all the data at its disposal to determine the most effective treatment for the patient, and to follow through on that treatment with a minimum of unnecessary testing.

When we know which procedures are likely to be successful, we'll be in a position where we can pay only for the health care that works. When we can do that, we've solved Wanamaker's problem for health care.

CHAPTER 5

Enabling Data

Data science is not optional in health care reform; it is the linchpin of the whole process. All of the examples we've seen, ranging from cancer treatment to detecting hot spots where additional intervention will make hospital admission unnecessary, depend on using data effectively: taking advantage of new data sources and new analytics techniques, in addition to the data the medical profession has had all along.

But it's too simple just to say "we need data." We've had data all along: handwritten records in manila folders on acres and acres of shelving. Insurance company records. But it's all been locked up in silos: insurance silos, hospital silos, and many, many doctor's office silos. Data doesn't help if it can't be moved, if data sources can't be combined.

There are two big issues here. First, a surprising number of medical records are still either hand-written, or in digital formats that are scarcely better than hand-written (for example, scanned images of hand-written records). Getting medical records into a format that's computable is a prerequisite for almost any kind of progress. Second, we need to break down those silos.

Anyone who has worked with data knows that, in any problem, 90% of the work is getting the data in a form in which it can be used; the analysis itself is often simple. We need electronic health records: patient data in a more-or-less standard form that can be shared efficiently, data that can be moved from one location to another at the speed of the Internet. Not all data formats are created equal, and some are certainly better than others: but at this point, any machine-readable format, even simple text files, is better than nothing. While there are

currently hundreds of different formats for electronic health records, the fact that they're electronic means that they can be converted from one form into another. Standardizing on a single format would make things much easier, but just getting the data into some electronic form, any, is the first step.

Once we have electronic health records, we can link doctor's offices, labs, hospitals, and insurers into a data network, so that all patient data is immediately stored in a data center: every prescription, every procedure, and whether that treatment was effective or not. This isn't some futuristic dream; it's technology we have now. Building this network would be substantially simpler and cheaper than building the networks and data centers now operated by Google, Facebook, Amazon, Apple, and many other large technology companies. It's not even close to pushing the limits.

Electronic health records enable us to go far beyond the current mechanism of clinical trials. In the past, once a drug has been approved in trials, that's effectively the end of the story: running more tests to determine whether it's effective in practice would be a huge expense. A physician might get a sense for whether any treatment worked, but that evidence is essentially anecdotal: it's easy to believe that something is effective because that's what you want to see. And if it's shared with other doctors, it's shared while chatting at a medical convention. But with electronic health records, it's possible (and not even terribly expensive) to collect documentation from thousands of physicians treating millions of patients. We can find out when and where a drug was prescribed, why, and whether there was a good outcome. We can ask questions that are never part of clinical trials: is the medication used in combination with anything else? What other conditions is the patient being treated for? We can use machine learning techniques to discover unexpected combinations of drugs that work well together, or to predict adverse reactions. We're no longer limited by clinical trials; every patient can be part of an ongoing evaluation of whether his treatment is effective, and under what conditions. Technically, this isn't hard. The only difficult part is getting the data to move, getting data in a form where it's easily transferred from the doctor's office to analytics centers.

To solve problems of hot-spotting (individual patients or groups of patients consuming inordinate medical resources) requires a different combination of information. You can't locate hot spots if you don't have physical addresses. Physical addresses can be geocoded (con-

verted from addresses to longitude and latitude, which is more useful for mapping problems) easily enough, once you have them, but you need access to patient records from all the hospitals operating in the area under study. And you need access to insurance records to determine how much health care patients are requiring, and to evaluate whether special interventions for these patients are effective. Not only does this require electronic records, it requires cooperation across different organizations (breaking down silos), and assurance that the data won't be misused (patient privacy). Again, the enabling factor is our ability to combine data from different sources; once you have the data, the solutions come easily.

Breaking down silos has a lot to do with aligning incentives. Currently, hospitals are trying to optimize their income from medical treatments, while insurance companies are trying to optimize their income by minimizing payments, and doctors are just trying to keep their heads above water. There's little incentive to cooperate. But as financial pressures rise, it will become critically important for everyone in the health care system, from the patient to the insurance executive, to assume that they are getting the most for their money. While there's intense cultural resistance to be overcome (through our experience in data science, we've learned that it's often difficult to break down silos within an organization, let alone between organizations), the pressure of delivering more effective health care for less money will eventually break the silos down. The old zero-sum game of winners and losers must end if we're going to have a medical system that's effective over the coming decades.

Data becomes infinitely more powerful when you can mix data from different sources: many doctor's offices, hospital admission records, address databases, and even the rapidly increasing stream of data coming from personal fitness devices. The challenge isn't employing our statistics more carefully, precisely, or guardedly. It's about letting go of an old paradigm that starts by assuming only certain variables are key and ends by correlating only these variables. This paradigm worked well when data was scarce, but if you think about it, these assumptions arise precisely because data is scarce. We didn't study the relationship between leukemia and kidney cancers because that would require asking a huge set of questions that would require collecting a lot of data; and a connection between leukemia and kidney cancer is no more likely than a connection between leukemia and flu. But the existence of data is no longer a problem: we're collecting the data all

the time. Electronic health records let us move the data around so that we can assemble a collection of cases that goes far beyond a particular practice, a particular hospital, a particular study. So now, we can use machine learning techniques to identify and test all possible hypotheses, rather than just the small set that intuition might suggest. And finally, with enough data, we can get beyond correlation to causation: rather than saying “A and B are correlated,” we’ll be able to say “A causes B,” and know what to do about it.

CHAPTER 6

Building the Health Care System We Want

The U.S. ranks 37th out of developed economies in life expectancy and other measures of health, while by far outspending other countries on per-capita health care costs. We spend 18% of GDP on health care, while other countries on average spend on the order of 10% of GDP. We spend a lot of money on treatments that don't work, because we have a poor understanding at best of what will and won't work.

Part of the problem is cultural. In a country where even **pets can have hip replacement surgery**, it's hard to imagine not spending every penny you have to prolong Grandma's life—or your own. The U.S. is a wealthy nation, and health care is something we choose to spend our money on. But wealthy or not, nobody wants ineffective treatments. Nobody wants to roll the dice and hope that their biology is similar enough to a hypothetical "average" patient. No one wants a "winner take all" payment system in which the patient is always the loser, paying for procedures whether or not they are helpful or necessary. Like Wanamaker with his advertisements, we want to know what works, and we want to pay for what works. We want a smarter system where treatments are designed to be effective on our individual biologies; where treatments are administered effectively; where our hospitals are used effectively; and where we pay for outcomes, not for procedures.

We're on the verge of that new system now. We don't have it yet, but we can see it around the corner. Ultra-cheap DNA sequencing in the

doctor's office, massive inexpensive computing power, the availability of EHRs to study whether treatments are effective even after the FDA trials are over, and improved techniques for analyzing data are the tools that will bring this new system about. The tools are here now; it's up to us to put them into use.

CHAPTER 7

Recommended Reading

We recommend the following articles and books regarding technology, data, and health care reform:

- Ahier, Brian. “**Big data is the next big thing in health IT**,” O’Reilly Radar. February 27, 2012.
- Bigelow, Bruce. “**Big Data, Big Biology, and the ‘Tipping Point’ in Quantified Health**,” Xconomy. April 26, 2012.
- Brawley, Otis Webb. *How We Do Harm: A Doctor Breaks Ranks About Being Sick in America*. St. Marten’s Press, 2012.
- Christensen, Clayton M. et al. *The Innovator’s Prescription: A Disruptive Solution for Health Care*. McGraw Hill, 2008.
- Howard, Alex. “**Data for the Public Good**,” O’Reilly Radar. February 22, 2012.
- Manyika, James et al. “**Big data: The next frontier for innovation, competition, and productivity**,” McKinsey Global Institute. May, 2011.
- Oram, Andy. “**Five tough lessons I had to learn about health care**,” O’Reilly Radar. March 26, 2012.
- Shah, Nigam H and Jessica D Tenenbaum. “**The coming age of data-driven medicine: translational bioinformatics’ next frontier**,” *Journal of the American Medical Informatics Association* (JAMIA). March 26, 2012.
- Trotter, Fred and David Uhlman. *Meaningful Use and Beyond*. O’Reilly Media, 2011.

- Wilbanks, John. “[Valuing Health Care: Improving Productivity and Quality](#)” [PDF], Ewing Marion Kauffman Foundation. April, 2012

About the Authors

Tim O'Reilly is the founder and CEO of O'Reilly Media Inc., thought by many to be the best computer book publisher in the world. O'Reilly Media also hosts conferences on technology topics, including the O'Reilly Open Source Convention, Strata: The Business of Data, and many others. O'Reilly's Make: magazine and Maker Faire has been compared to the West Coast Computer Faire, which launched the personal computer revolution. Tim's blog, the O'Reilly Radar "watches the alpha geeks" to determine emerging technology trends, and serves as a platform for advocacy about issues of importance to the technical community. Tim is also a partner at O'Reilly AlphaTech Ventures, O'Reilly's early stage venture firm, and is on the board of Safari Books Online.

Mike Loukides is Vice President of Content Strategy for O'Reilly Media, Inc. He's edited many highly regarded books on technical subjects. He's particularly interested in programming languages, Unix and what passes for Unix these days, and system and network administration. Mike is the author of *System Performance Tuning*, and a coauthor of *Unix Power Tools*. Most recently, he's been fooling around with data and data analysis, languages like R, Mathematica, and Octave, and thinking about how to make books social.

Julie Steele is the Content Editor for Strata at O'Reilly Media. She is coauthor of *Beautiful Visualization* and *Designing Data Visualizations*. She finds beauty in exploring complex systems, and thinks in metaphors. The best part of her day is finding patterns across verticals and traditional silos, and connecting people who are working on similar problems in seemingly unrelated areas. She is particularly drawn to the visual medium as a way to understand and transmit information.

Colin Hill is the president, chairman, and cofounder of GNS Healthcare. He brings years of hands-on scientific experience to his role, with expertise in the areas of computational physics and systems biology.