Google's headquarters in Mountain View, Calif. As one employee put it, Google is a "candy store" for O.R. practitioners (inset).

Source: Google

Operations research at Google

Advanced analytics permeates work at Google, making the multinational technology giant a 'candy store for O.R. practitioners.'

By Brian Thomas Eck and Amber Richter

The mission of Google, Inc. is "to organize the world's information and make it universally accessible and useful." This has spawned efforts as diverse as optical fiber to the home (gFiber), longevity research (Calico), smart home automation (Nest), YouTube, glucose-detecting contact lenses (Verily), self-driving cars and many others. Such broad-reaching innovations are possible due to robust search and ads businesses.

It is well known that web search was the foundation of Google. Recognition that a web page is important if it is pointed to by other important pages translates into math: PageRank's importance scores are the stationary values of an enormous Markov chain [1]. With this start, it is not surprising that Google's culture goes hand in hand with analytical literacy.

Working as an operations research (O.R.) practitioner surrounded by highly analytical colleagues brings, by contrast, new meaning to the term "isolated practitioner." Wandering around Google, one sees whiteboards everywhere, filled with equations, graphs, pseudo code and probability distributions. Widespread respect for data to inform decisions is accompanied by healthy skepticism; numbers can also mislead.

For example, peer reviews and presentations at all levels have the primary intent of improving the analysis. A presentation where the audience politely listens and applauds at the end represents a failure to engage. A successful presentation features frequent interruptions, challenges to assumptions and analysis, and lively debate with the audience. This holds true even in executive presentations, where the O.R. is vetted with deep experts in computer science and statistics, technical minds imbued with a broad grasp of the business. In contrast to organizations where advanced analytics methods remain shielded, Googlers pry open the Black Box and engage. For experienced hires from other companies, this can be initially disconcerting, but over time, they discover that they gain trust and impact by embracing this method of collaboration.

Organizing information at scale often relies on software. Each day, Google's systems crawl 20 billion web pages, stream hundreds of millions of hours of YouTube videos and activate 1.5 million Android devices. This scale requires a massive physical infrastructure: Google's unparalleled worldwide cluster computing system. This infrastructure includes 13 data center campuses of staggering size. The Council Bluffs, Iowa, campus is the largest in the world, with multistoried data center buildings that have building pads over a third of a mile long. In addition, Google has presence in dozens of cities across more than 33 countries, with a global network of fiber optic cables connecting it all, to bring information and services quickly and reliably to its end users.

Building and growing this infrastructure requires insights provided by hundreds of advanced analytics projects. Ongoing operations require efficient allocation of compute, storage and network resources across internal product areas and external cloud customers. Advanced analytics also make up the core function for many Google products, from improving users' search results to finding an optimal driving route for Google Maps directions.

Advanced analytics techniques go beyond what we traditionally define as O.R., and include methods from fields such as statistics, robotics, control systems, game theory, econometrics and risk analysis. For example, machine learning (ML) is used to improve search results, automate language translation, protect Gmail and Chrome users from spam and malware, and even improve data center energy efficiency (Google is the

All About the Roundtable

The Roundtable consists of the institutional members of INFORMS with member company representatives typically the overall leader of O.R. activity. The Roundtable is composed of about 50 organizations that have demonstrated leadership in the application of O.R. and advanced analytics. The Roundtable culture is peer-to-peer, encouraging networking and sharing lessons learned among members.

The Roundtable meets three times a year. Roundtable goals are to improve member organizations' OR/MS practice, help Roundtable representatives grow professionally and help the OR/MS profession to thrive. Further information is available at http://roundtable.informs.org.

The Roundtable also has an advisory responsibility to INFORMS. According to its bylaws, "The Roundtable shall regularly share with INFORMS leadership and advise the INFORMS Board on its views, its suggested initiatives and its implementation plans on the important problems and opportunities facing operations research and the management sciences as a profession and on the ways in which INFORMS can deal proactively with those problems and opportunities." The Roundtable meets with the INFORMS president-elect each spring to discuss practice-related topics of interest to him or her, and with the entire INFORMS Board each fall to discuss topics of mutual concern.

This series of articles aims to share with the INFORMS membership at large some information and insights into how O.R. is carried on in practice today.

largest corporate purchaser of renewable energy on the planet; Google's data centers are among the most energy efficient in the world). Google has published hundreds of papers related to ML (see http://research.google.com) and has open-sourced many ML tools through TensorFlow (see www.tensorflow.org).

A Vibrant Community of Quants

O.R. practitioners are often interested in how companies organize their O.R. employees: in central teams, embedded within functional domains or some hybrid. Google uses a hybrid approach and augments it by providing clear direction on how individuals' careers can advance, building and sustaining a community of quantitative analysts, and retaining that community's identity.

At Google, operations researchers need to be generalist problem solvers, and they typically work in roles such as data scientist (quantitative analyst), software engineer or research scientist. As such, they are held to the standards of their associated job ladder. These ladders describe expectations at each level throughout a contributor's career. Through committee-based decision-making, the ladders provide consistency across interviewing, hiring, calibrating performance ratings and evaluating promotions. This discipline sets a uniformly high bar for hiring and promotion across the company, and the consistent expectations facilitate rotation among teams. Furthermore, because all professionals participate, they become deeply familiar with each other's work.

The data scientist, or quantitative analyst, ladder includes several hundred analysts, most with a statistics background, a significant minority with an O.R. background, and small groups from fields such as biostatistics, economics and computational engineering. The number of analysts supporting a domain can vary from a few to a few dozen. The ladders enable that degree of domain specialization

while preserving consistently high standards for technical hiring and work and embedding the analyst in a broader technical community.

In addition to job ladders, Google uses forums for professionals to share their work freely within the company, such as informal lunch series, tech talks, a data science blog and more formal global summits. By providing these community-building activities and job ladders, Google sustains community identity and career direction for its O.R. analysts while positioning O.R. practitioners in both centralized and embedded teams.

The following section highlights two centralized teams: one with a functional focus on the technical infrastructure domain and the other with a focus on methods and tools used across multiple application domains.

Core O.R. Teams

Operations Decision Support (ODS): This Mountain View, Calif.-based team is comprised of operations research Ph.D.s who focus on Google's technical infrastructure: optimizing the hardware supply chain, planning data center and wide-area network capacity, optimizing server deployments and lifecycles, and improving the utilization of compute and storage resources. Many of the projects are variants of well-known trade-offs to optimize cost: the Newsvendor problem, timing for technology refresh and determining build frequency economic order quantities. ODS's focus on cost optimization led to its strong reputation for *total cost of ownership* management.

For example, Google positions network gear in multiple cities around the world in order to connect with peers (Internet Service Providers) closer to their end users. How many and which facilities should be used, and which gear should be placed where, require trade-offs between facility and fiber costs to connect gear across sites. The team uses simulation to cost-optimize strategic roadmaps for evolution of peering support within Google's network.

Another example is deciding when to replace an older server with one from the newest generation, which requires optimizing various costs. Analyses such as this inform many thousands of decisions, some in the form of a policy, some as a simple calculator, and some as a complex decision support tool, run either periodically or on demand.

Technician inside a Google data center.

ODS also does forecasting and capacity planning. ODS produces a range forecast of the fleet, which consists of compute, storage and power capacity needs in the data centers. This is used in making many downstream decisions including acquisition of land and utilities, new construction and network capacity augmentation. Moving from point forecasts to quantification of the variation implied by forecast error, and using this variation to set inventory buffers, necessarily involves substantial organizational transformation as well as analysis. As in most companies, this integration of hard and soft skills is an essential ingredient in the toolkit of an O.R. practitioner at Google.

Beyond these examples, ODS applies advanced analytics to optimize Google's fleet. It uses mixed integer programming (MIP) models to plan server deployments across the fleet and to optimally add and reshape compute and storage capacity within each cluster of machines. ODS also uses simulation and machine learning models to overcommit and schedule compute and storage capacity to improve utilization.

Operations Research Team (O.R.): While the ODS team is organized around application domains, the O.R. team is organized around methods. This Paris-based group develops and supports combinatorial optimization software and applies it to large-scale, real-world problems across the company. This software engineering and research team originated out of a challenge posed by Google StreetView.

Obtaining Street View imagery requires efficiently routing cars down streets around the world to capture all the needed images. Solving this classic Chinese Postman Problem led to savings on labor and car maintenance, reduced emissions and more up-to-date imagery through shorter and thus more frequented routes. This problem motivated the founding of the O.R. team as Google's in-house vehicle routing team, and the team's expertise quickly expanded from there.

The team develops its optimization software libraries to handle the speed, scalability and security that Google-scale projects demand. More than 150 teams at Google use these libraries, and most of them have been open-sourced as the or-tools suite, available on GitHub. These libraries include a gold-medal winning constraint solver, vehicle routing library, linear optimization solver, Boolean optimization solver, knapsack solver and libraries for solving flow and assignment problems (see https://developers.google.com/optimization/).

Although it grew out of StreetView, the O.R. team works on projects all across Google. The team has developed optimization algorithms to stabilize YouTube videos, direct navigation for the Loon (Internet balloon) fleet, and even assign people across Google to serve on promotion committees. The O.R. team has also worked with Terra Bella, Google's subsidiary formerly known as Skybox Imaging.

Terra Bella has satellites that orbit the earth in short cycles and capture high-resolution satellite imagery of places all around the world. Fixed orbit paths limit when locations are in view of each satellite, and data downlink opportunities are available only when the satellites are near fixed ground stations. The OR team developed a MIP approach to schedule the timing and location of satellite captures of target images and downlinks of satellite image data.

Isolated Practitioners (Not)

In addition to the large O.R.-focused teams, there are many individual and small groups of O.R. contributors all across Google linked together by the community mechanisms described above.

Several O.R. practitioners work across Google Express, Google's online delivery service providing fast delivery of products from popular retailers. They solve problems such as demand forecasting, capacity planning, scheduling and routing to help deliver products from retailers to customers. For example, some practitioners work on forecasting the number of orders by time of day and location to be able to schedule drivers and store operators via optimization algorithms that account for constraints such as the very short lead time of orders, staff preferences and consistency in individual staff schedules over time.

An O.R. problem that arises often in Google infrastructure is dynamic, multi-dimensional bin packing and load balancing. One example is job scheduling in Google's massively parallel computing environments. Here, the multidimensional items are jobs that need to be placed on machines (bins) subject to multiple hard and soft constraints, such as available CPU and RAM, job preferences, priorities and specialized hardware needs. The infrastructure-related Algorithms and Analytics teams work with the relevant engineering teams to improve both online dynamic algorithms and offline MIP-based solutions for scheduling jobs, adding resources to data centers and answering related capacity planning questions.

The Large-Scale Optimization research team, based in New York, works with the relevant engineering teams to improve the efficiency and robustness of Google's computational infrastructure, such as the backend systems that serve search and Google's external cloud offering. For example, the team applied balanced graph partitioning algorithms to cluster search terms according to how often they co-occur in search queries, then used this clustering to govern how queries are distributed among machines in the search backend. This change greatly increased the rate at which queries can be served via improved caching.

A software engineering team in Network Architecture does capacity planning and risk analysis for Google's wide area network of fiber optic cables. Their models seek to minimize cost while ensuring availability, speed and scalability, three key components of Google's network. They use MIP models to determine the cheapest network that can route flows during a given set of fiber failure scenarios. A Monte Carlo simulation tests the resulting network against availability and latency service level requirements to determine additional failure scenarios to include in the MIP in the next iteration.

O.R. is Everywhere

In summary, advanced analytics permeates work at Google. It might seem easier to describe where it hasn't been applied. But this impression is quickly contradicted by nontraditional cases such as human resources identifying an optimal number of candidate interviews or a job posting for a food service analytics and insights manager. There are always new problems to solve and new impacts to deliver. The relevance of O.R. and advanced analytics is stronger than ever in this burgeoning high-tech industry.

Working here is perhaps best summed up by a recent quote from a Google analyst: "Google is like a candy store for O.R. practitioners." ORMS

Brian Thomas Eck, Ph.D., and *Amber Richter, Ph.D.,* are quantitative analysts on the Operations Decision Support team within Technical Infrastructure. Eck is also the Google representative to the INFORMS Roundtable.

Disclaimer: The opinions expressed in this article are those of the authors and do not necessarily represent the views of Google.

There are **always new problems** to solve and **new impacts to deliver.** The relevance **of O.R. and advanced analytics** is **stronger than ever** in this burgeoning **high-tech industry.**

REFERENCE

1. Langville, A.N. and Meyer, C.D., 2006, "Google's PageRank and Beyond: The Science of Search Engine Rankings" (page 31), Princeton University Press, Princeton, N.J.