BEYOND DATA: KNOWLEDGE PRODUCTION IN BUREAUCRACIES ACROSS SCIENCE, COMMERCE, AND THE STATE

Workshop at the German Historical Institute Washington (GHI) in cooperation with and co-sponsored by the Max Planck Institute for the History of Science, Berlin (MPIWG). Conveners: Sebastian Felten (MPIWG), Philipp Lehmann (MPIWG / University of California Riverside), Christine von Oertzen (MPIWG), Simone Lässig (GHI). Participants: Maria Avxentevskaya (MPIWG), Rüdiger Bergien (Zentrum für Zeithistorische Forschung, Potsdam), Dan Bouk (Colgate University), Maura Dykstra (Caltech), Anna Echterhölter (Humboldt University, Berlin), Elisabeth Engel (GHI), Devin Fitzgerald (Harvard University), Tom Ewing (Virginia Tech), Adam Fulton Johnson (University of Michigan, Ann Arbor), Kyrill Kunakhovich (University of Virginia), Anindita Nag (University of Calcutta/MPIWG), Kathryn Olesko (Georgetown University), Simon Ottersbach (University of Gießen), Chris Phillips (Leeds Trinity University), Martha Poon (Columbia University/ New School), Theodore Porter (University of California, Los Angeles), Axel Jansen (GHI), Renée Raphael (University of California Irvine), John Sabapathy (University College London), Hanna Turner (Simon Fraser University), Richard Wetzell (GHI), Chelsea Zi Wang (Claremont McKenna College).

This workshop was a collaboration between the Max Planck Institute for the History of Science in Berlin and the GHI. The GHI's new focus on the history of knowledge gelled with the MPI's longstanding interest to bring historians of different disciplines into a dialogue on how, through which means, and to what ends data is collected, stored, and transformed — this time applied to bureaucratic frameworks, a human endeavor that, just like science, produces knowledge in highly structured ways.

The workshop explored how bureaucratic practices of making and using knowledge emerged and evolved in a broad range of settings — from local administrations to mining offices, from colonial trade companies to insurance firms, and from the Middle Ages to the recent past. By bringing together scholars from diverse historical disciplines, the workshop addressed questions that are difficult to answer through individual case studies alone: Did the need to know shared by government, commerce, and science result in similar material practices of collecting and transforming large amounts of data? Or did the divergent internal logics of these domains produce

idiosyncratic approaches? In what ways did tools to classify, order, and process information migrate from one institutional context to another, and how did they change? Do we have to specify — and possibly diversify — our notion of knowledge when the aim of information processing is not so much getting the facts straight but making decisions?

The seven panels were structured thematically to stimulate thinking across domains, regions, and periods. The presenters in panel I, "Antagonizing Logics," explored how state-bureaucratic and economic tools and practices of knowledge production, as well as the values and rhetoric substantiating them converged or clashed. In his opening talk, Dan Bouk traced the rise of depreciation accounting systems for public utilities based on actuarial models for machinery alongside the development in corporate life insurance of human life valuations intended to support public health expansion. By focusing on this technical aspect of large-scale management, he troubled our narratives of state/corporate antagonism by revealing how much U.S. state and corporate bureaucracies learned from one another. Sebastian Felten explored the role of "bureaucratic imagination" to explain why the rationality of bureaucrats and business people can be at odds sometimes. Using the failed attempts by a Saxon mining official to secure capital in Amsterdam as an example, Felten explained that objectives such as transgenerational "sustainability" were fully rational only when many mines and long timespans were condensed in one imaginative grasp, which Saxon officials routinely did. Chris Philips demonstrated that the First World War was an industrial war not just in terms of the weaponry produced and used, but also regarding the relentless pursuit of industrial productivity and efficiency on the ground where diagrams, graphs, and formulae became key components of an "information infrastructure" decisive for evaluating battlefield performance and success.

In his keynote lecture, Theodore Porter reflected on how and to what ends commercial and state bureaucracies collected and processed numerical information. Drawing on rich source material from his work on the struggle over cure rates in nineteenth-century asylums, Porter demonstrated how numbers become sites of contestation when used for official assessments of institutions. Presenting facts in numerical form is not so much an export of science but emerged through tools and means developed within bureaucratic and political frameworks to reconcile central control with local autonomy. The

unruly outcomes of such struggles were often "funny numbers," which concealed the evasion of goals and corruption of measures. The mismatch between the "boring" appearances of statistics and dishonest backstage manipulations of numbers, Porter concluded, presents a danger evident in many recent efforts to decentralize the functions of governments and corporations using incentives based on quantified targets.

The presenters of the second panel, "Laws of the Local," focused on the knowledge practices that empires deployed to exercise control in far-away localities. Maura Dykstra reconsidered bureaucratic reporting and reviewing practices in the Qing legal system. To reconcile local autonomy with imperial law, the Qing established a procedure in which most local cases were kept outside of the law. However, the "local case" is a focal point for examining how central, provincial, and local institutions combined in the everyday search for justice. The distinction between legal and local entailed a distinct set of reporting procedures that created fundamental differences between those cases which were summarized for legal review and those which were not slated for scrutiny by supervisory officials. Anna Echterhölter explored the tension between universal measurement units and local realities in the statistical reports that the colonial government in German New Guinea sent year after year to Berlin. Local systems of counting, indigenous measures of value, and calculated imprecisions proved valuable instruments of colonial administration, as officers on the scattered islands often showed a profound ignorance about the conditions on the ground. Adam Fulton Johnson investigated how ethnological knowledge production in the U.S. was transformed during the nineteenth century from a bureaucratic archival technique into a contextual and often idiosyncratic practice of documentation: scribbled field notes and miscellaneous anecdotes. The conventions of knowledge inscription established in these ethnographic encounters structured both the development of disciplinary anthropology as well as the continuing relationship that southwestern Indian communities had with the U.S. bureaucratic state.

In Panel III, "Promise of the Pattern," presenters addressed how bureaucracies engage in big data collection, infrastructure, and how they derive knowledge from these data. Rüdiger Bergien explored if and to what extent computerization by the West German Police and the East German Ministry of State Security led to a change in their particular information management and knowledge production. He

found that computerization did not turn these organizations into almighty surveillance instruments. Rather than overcome the issue of uncertainty, as intended, computerization reproduced existing and even introduced new uncertainty, e.g. about who should be given access to the database. Tom Ewing reflected on the ways in which the 1889 New York "Russian influenza" epidemic posed a unique challenge to knowledge production in bureaucracies involved in public health, medicine, and social policies. Not the agencies alone, but also daily newspapers, medical journals, and doctors' meetings debated what was viable information and how it should be analyzed and disseminated. Martha Poon focused on Microsoft's current concerns to pursue a new data-driven business model based on cloud computing. Poon argued that this may be a result of shifting dynamics in global financial markets, and not just of technical innovation, with fundamental impact on the architecture of the commercial system of data-driven administration — a business model that some are calling "surveillance capitalism."

Panel IV, "Noise and Signal," highlighted the media of bureaucratic data collection and dissemination. Anindita Nag discussed how in 1902, the Colonial Office established its Visual Instruction Committee to produce photographic evidence of Britain's improvement of its colonial territories, mostly in the form of lantern slide lectures for British classrooms, lecture halls, and libraries. In identifying visualization as a significant way to think about bureaucratic knowledge production, Nag shed light on the technological, representational, and aesthetic enterprise of the British state in India. Simon Ottersbach demonstrated how Radio Free Europe (RFE) accumulated a wealth of data, which was used for short — or long-term analyses of societal, cultural, political, or economic processes in the "East." Unparalleled in the Cold War media landscape, research emerging from this archive was shared not just in-house but also with governmental task forces, university lecturers, and global media entities, thus shaping what the West knew about the East. Devin Fitzgerald discussed the nature of translation in the early Qing (1644-1720) archive, after the Manchu invasion in 1636. Manchu materials constituted a different "stream" of documents, and only entered the "open archive" when the state disseminated them in histories, proclamations, and institutional compilations. While the Chinese side of the state archive functioned much as it did prior to 1644, the "open archive" for the integrated Qing Empire only existed with direct court participation.

Panel IV, entitled "Keeping Track of the Field," focused on the workflow and frictions that arose between state institutions and those they employed to collect but also interpret data and specimens. Philipp Lehman examined how data from the notes and recordings of officials in German Southwest Africa traveled to Berlin, were the Foreign Office assumed a gatekeeper function, forwarding or withholding particular data (e.g. weather observations to the naval observatory). Lehmann used these data journeys to discuss the often tense relationship between different interests groups among the institutions of the German empire. Hannah Turner examined how at the Smithsonian's National Museum of Natural History, material technologies such as the circular, the ledger book, and the catalog card played a role in establishing what information was "valid." Knowledge originating in the North American indigenous communities was often excluded. While practices have changed at the museum, Turner argued that the collected objects and the information associated with them came to form fundamental concepts in anthropology and material culture research upon which we still rely.

Panel VI, titled "Answers and Answerability," examined the material forms in which bureaucracies elicited answers to their questions, both from the people they governed and from their own officials. Kyrill Kunakhovich traced the rise and fall of public opinion polling in Soviet bloc Poland and East Germany. Public polling data became increasingly important in the 1960s and 1970s, as communist regimes sought to project responsiveness to the people's will whilst struggling to reconcile popular desires with their political agendas. By the 1980s, Soviet bloc leaders began to ban public opinion research altogether and even to disseminate made-up statistics known as the "propaganda of success." Renée Raphael showed how many of Philip II's subjects sought to participate in the Spanish "paper state" by penning their own discourses. Using one 39-page booklet detailing the discovery of the famous mining region in Potosí (Bolivia) as an example, Raphael argued that such interventions highlight a particular feature of the Habsburg bureaucracy: that individuals with good ideas were thought to have the duty to present them directly to the king. John Sabapathy analyzed a range of inquiries that southern French lawyer (and later pope Clement IV) Gui Foucois (d. 1268) was involved in. Sabapathy showed how distinctive such practices were between spheres (ecclesistical and secular) and countries (France and England), shedding new light on questions of state formation and the role played there by bureaucratic knowledge and investigative questionnaires.

The presenters of the last panel, "Finding a Common Language," investigated how imperial bureaucracies incorporated and channeled knowledge so that it could be put to use. Maria Avxentevskaya examined how technical translation, mostly from Western languages such as Dutch, influenced the Russian Enlightenment by transforming the realms of both practical knowledge and state management. Focusing on the letters and papers of Peter the Great, Avxentevskaya showed how the "localization" of knowledge through translation influenced institutional and educational practices, while newly acquired values of translation affected the conceptualization of practical knowledge, as well as the development of Russian imperial humanism. Kathryn Olesko explained how Prussian officials managed the empire's extensive eastern "frontier" resulting from the Polish Partitions in 1772, 1793, and 1795. Self-trained technicians swept across this new landscape gathering data, creating an "information order" of the Prussian frontier and contributing to the illusion that this space could be controlled. Olesko argued that this marked a turning point in the evolution of Prussia as a scientific and technical state. The final speaker, Chelsea Zi Wang, presented a peculiar solution to the problem of data synchronization in imperial China. "Layered quotations" found as early as the eighth century and well into the twentieth century eased document storage and retrieval within the confines of a single text. Long dismissed as excessively repetitive, this particular use of administrative language, as Wang argued, allowed the Chinese state to keep track of previous communications as a case moved up and down through different levels of government.

Taken together, the range of case studies was dazzlingly diverse, a fact that proved highly productive from the very start of the meeting. During the lively and inspiring discussions, participants focused on similarities and contrasts among the cases presented and identified recurring themes. In the final discussion these were condensed into four general observations, which seem particularly relevant for the history of bureaucratic knowledge. First, in most cases, the bureaucracies at hand spent considerable resources not only on knowing the territory, populations, markets etc., that they considered their environment, but also on knowing themselves — sometimes explicitly by creating reports and organizing reviews, sometimes implicitly through the structure of their media (e.g. in the Chinese case of "layered quotations"). Second, language emerged repeatedly as a problem for actors at the boundary between the bureaucratic "self" and its environment. Incomprehension and the need to translate were

prompted not only by the difference between "natural" languages (e.g. Dutch and Russian) but — even more importantly — between the technical language employed by the bureaucrats in official communication, and the language used by everybody else. Third, in all of the cases considered bureaucracies engaged in practices of abstraction. Some of the authorities were so far removed from the Weberian ideal type of bureaucracy that concepts inherited from sociology and political science made it harder rather than easier to understand their workings. However, what appeared as a common theme was that their mode of domination was based on abstraction and technical language. At the same time, this is where the historicity of the different cases appeared most clearly. While the medieval Church, the Qing Empire and the German colonial governments all abstracted from local realities, they did so in bewilderingly different ways. Fourth, the history of bureaucratic knowledge — showcased here as if through a kaleidoscope — might provide a meaningful framework to reconsider Big Data as a recent social, political and economic phenomenon. When current practices of gathering, storing, and analyzing mounds of data are taken as the latest expression of a long quest to rule by abstraction and technical language, they may seem less new, less puzzling, and perhaps also less difficult to control.

Sebastian Felten and Christine von Oertzen (MPIWG)