Ian Lowrie

PhD Candidate

Rice University

# Building an Information Economy: Work and Education in the Russian Data Sciences

**Research Abstract:**

My research focuses on elite efforts to restructure work and education in the Russian data sciences in order to build a robust Russian information economy. By investigating the novel forms of training and research emerging at hybrid industrial-academic institutions, my project provides critical and timely insight into the rapidly changing Russian information technology sector during a time of heightened geopolitical volatility. As Russia responds to its current economic crisis with radical efforts to supplement natural resource extraction with a robust information technology sector, a nuanced understanding of how work, education, and research are being restructured in the emerging data science ecosystem will be essential for regional economic and scientific policy-making. Further, the chance to observe this unique ecosystem as it develops promises substantial contributions to our scholarly understanding of the ongoing transformations of post-socialism and the global knowledge economy more broadly.

**Research Goals:**

With the ruble plummeting alongside the price of oil and gas, growing political friction between Russia and the West, and the Kremlin preparing for a major recession in the coming year, the

future of Russian computer science does not weigh heavily on the minds of most Russian citizens. For the data scientists among whom I have been conducting research, however, the current situation has provoked not only fear of obsolescence, but a renewed commitment to leveraging their expertise with machine learning and the analysis of big data to build the future Russian economy. These academic and industrial researchers, together with a constellation of business and policy elites, are aggressively forging new modes of work and training, pushing the boundaries of post-Soviet science (Graham and Dezhina 2008). Bringing together contemporary social scientific research on education, the information economy, and post-socialism, my research has been designed to provide a timely look at their attempts to build a truly modern, truly Russian information technology sector.

Situated at the intersection of statistics and computing, data science draws upon techniques such as natural language processing, image recognition, and machine learning to extract knowledge and make predictions from data. To date, this field has been dominated by Western scholars profiting from the ubiquity of easily-accessible data sets relevant to policy and business applications in their home countries, as well as a highly flexible approach to post-secondary education that allows for both rapid development of new courses and mutually-profitable collaborations between industry and academy. Russia, conversely, has been hampered by a high wall between academy and industry, lack of easy access to data specific to the local context, and highly conservative and bureaucratic educational system. The current moment of crisis, however, offers an opportunity for ongoing elite plans to build a Russian knowledge economy (Graham 2013). My research follows their projects as they blur the boundaries between marketcraft and

statecraft, investigating new forms of work and education that require both professional and academic skillsets, to ask three interlocking questions:

(1) What new pedagogical techniques are being developed in these hybrid spaces, and how do they shape the outlook and practical activities of young data scientists?

(2) What new research programs are emerging in the Russian data sciences as a result of collaboration between industry and academy? What distinguishes these projects from their counterparts in the West?

(3) How do these collaborations fit into the broader social organization of the emerging Russian knowledge economy?

**Research Activities**

To answer these questions, I have been conducting interviews with members and participating in the intellectual life of the Moscow data science community, with some supplemental interviews being done with participants from St. Petersburg and Yekaterinburg. My methodology for selecting interviewees has been a careful snowball sampling, building upon trusted contacts to extend my network of research participants through personal introductions, as well as by developing working relationships with other professionals through my participation in their community more broadly.

Most intensively, I have been working with two closely-related institutions that stand as the best current examples of ongoing efforts to develop a Russian information economy, and reform higher education in computer science, respectively. First, I have been working with developers, managers, and industrial researchers at a major web infrastructure company often called the "Russian Google," widely considered to be *the* leading data science firm in Russia. Second, I have been speaking with graduate students, lecturers, professors, and administrators at a newer National Research University (NRU), which recently partnered with this technology firm to form a school of Computer Science with the explicit aim of producing elite academic and industrial data science researchers. Additionally, I have been observing data science courses and participating in seminars and master-classes at this university.

In addition to these core field-sites, I have been participating in a range of seminars, lectures, and other events organized by and for data and computer scientists in Moscow. My perspective on my core field sites has been supplemented by extending my sample to include interviews with students and academics from other universities and institutes, as well as other information technology firms with a strong Russian presence.

*National Research University*

The NRU with which I have been working has been exceptionally welcoming and supportive of my research. My focus at the NRU has been on the computer science faculty, with particular attention paid to the chairs working on data science and artificial intelligence. However, I have also conducted interviews with students and faculty from affiliated departments, such as applied

mathematics. All told, I have interviewed twelve graduate students, ten faculty, and five post-doctoral researchers at the NRU. These interviews have been essential for constructing a nuanced understanding of the backgrounds, motivations, capabilities, and goals of my subject population. Further, they have provided critical material for the construction of both a historical narrative about the decision-making process behind the foundation of this new faculty as well as a sophisticated institutional description of its current operation. Most anthropologically-speaking, however, they have provided a wealth of insight into the daily lives of students and academic researchers in the data scientists, painting a critical picture of their intellectual work and how it intersects with their life projects more generally.

In addition to interviews, I have been participating more broadly in the intellectual life of the department, both by observing seminars and courses as well as through more informal socialization with my interviewees. Recording the microscalar interactions between participants in these events has been crucial for my understanding of the epistemological and political dynamics of intradepartmental relations, the character of relationships between advisors and advisees, and most generally how data scientists work together to solve the problems that they have collectively set themselves.

*Web Infrastructure Firm*

My access to the corporate side of data science has been mixed, given the particular sensitivity of many topics with which data scientists work as well as understandable concerns about the release of proprietary information. Robust participant observation, that is to say, access to business

meetings, working sessions, and corporate decision-making more generally has proved unfeasible. However, the firm that I have been working with has provided very generous access to employees for interviews. I have conducted around 20 interviews at this firm, with industrial researchers, managers, and both front- and back-end developers. These interviews have provided a wealth of information about the internal culture of this company, the organization of its work-flow, the variegated projects on which it is currently working, and the demographics of its employees.

Beyond these institutional-ethnographic questions, the interviews have also explicitly focused on eliciting participants' educational backgrounds and attitudes towards the ongoing collaboration of their company with academia. In addition to the interviews conducted with personnel of the company itself, I have also recorded a number of conversations with lecturers at the company's own, private data analysis school, which is designed to rapidly develop the professional competencies necessary to work at the firm.

*Other Activities*

In order to contextualize the research conducted at my core field sites, I have participated in the intellectual community of computer and data scientists in Moscow, and conducted interviews with a range of professionals and academics from other leading institutions. Participation in open events such as Moscow Big Data Week and public meetups for data scientists has been critical for understanding the particularities of the approaches taken at the NRU and web services firm I have been studying, as well as for observing how the participants in my study interact with their

colleagues. Approximately twenty interviews with elite data scientists from both leading international and Russian information technology firms, as well as more traditional applied mathematics and computer science departments, have helped to paint a broader picture of the data science community in Russia, as well as contextualize the particular career trajectories of my primary informants.

**Important Research Findings:**

My time in Moscow has given me excellent overview of the institutional and demographic arrangements of the Russian data science community. My participant observation has provided a nuanced view of the range of potential career trajectories, forms of sociality, and modes of work and study typical of that community.

The world of Russian data science is populated by a surprisingly diverse array of actors. More senior data scientists generally hold advanced degrees in applied mathematics or computer science, although a great number received their formal training in other fields such as physics, the social sciences, or various types of engineering. In part, this is facilitated by the highly universal approach characteristic of the data sciences; they are designed to be a flexible network of related techniques that can be applied to the resolution of any number of concrete problems. However, what all of these researchers share is a high level of fluency in fundamental mathematics. Indeed, exceptional competence in this area, demonstrated through avenues such as successful competition in mathematics Olympiads, can often make up for lack of formal post-baccalaureate education, especially in the private sector. Mathematical acumen is viewed as a

necessary tool for the process of rapidly learning the new programming languages and applications typical of the ever-changing field of computer science, as well as an essential trading language for discussing projects with colleagues whose specific areas of expertise might differ from one's own. While many of those whom I interviewed participate in both private-sector and academic research, the salient thematics that emerged in each case are somewhat different.

The picture that is emerging from the academic side of my research is one of rapid, but not altogether smooth, change in the bureaucratic management of data science education. The NRU faculty, founded with deep support from industry, was designed to correct what were understood to be two key deficiencies in Russian computer science education. First, undergraduate and graduate students alike were held to be overly focused on foundational theoretical knowledge and lacking in concrete, applied skills. Second, the work of both students and professors in Russia is believed to be suffering from an overly strong barrier between the academy and industry. The new faculty attempts to correct this "imbalance" in several ways, such as by incorporating members of industry in the coursework planning and oversight committees, facilitating internships and work experience by students, using corporate data to train students, and hiring what, in America, might be called "professors in the practice" to teach programming and data analysis.

However, it remains unclear how successful these changes will be, as they are taking place within a highly inertial institutional and bureaucratic context. Many professors, particularly those working in administrative roles, have reported frustration with their inability to rearrange

education sufficiently quickly, although they readily acknowledge that it has become substantially easier, especially in comparison with their colleagues at places such as Moscow State University. Perhaps ultimately more compromising, however, is the fear that the baby of a historically rich tradition of sophisticated theoretical knowledge might be thrown out along with the bathwater of academic disconnection from the rapidly-changing landscape of industrial computer science. Compounding this fear is a worry that working exclusively with one large industrial firm, while they are universally acknowledged as an excellent partner, might ultimately result in overly-specific departmental foci in both research and training. As my research continues, I am eager to follow the interdepartmental conversation about the "rationalization" and "updating" of the curriculum and education, as well as these ambivalent feelings about its "industrialization."

The industrial research community faces related, but somewhat different and occasionally orthogonal problems. Like academic researchers, my interlocutors in industry are intensely interested in the direction that their community will take in the near future, but their concerns are somewhat different. Given that my research was primarily conducted at a well-established and highly successful firm, it is perhaps unsurprising that they are optimistic about the future of the Russian technology sector. They believe that they will continue to make money, and provide an essential service to the Russian economy more generally. However, they are less certain that the institutional and political support for the ascendance of knowledge work more generally can be counted upon. They are all quite cognizant that the economy in Russia to date remains firmly tied to natural resource extraction, and worry that short-sighted market and state actors will not

take the necessary steps, both legal and organizational, to provide a solid basis for the kind of flourishing information economy many of them view as the best possible future for Russia.

Given this interest in their future, many industrial researchers with whom I spoke are intensely invested in the educational projects outlined above, and are keen to work closely with academic researchers in the development of both new pedagogical programs and, indeed, concrete business applications and ventures. This high degree of collaboration, particularly in light of the failure of industry and academy to work seamlessly together in other sectors, is perhaps the most striking finding of my research.

**Policy Implications and Recommendations:**

Rapid global circulation and international collaboration have long been recognized as hallmarks of contemporary scientific knowledge production in general, and of the computational and mathematical sciences in particular. Despite the current geopolitical and economic tensions between Russia and the West, scientific and educational policy should continue to encourage the types of collaboration that forged strong bonds between Russian and international academics in the years immediately following the collapse of the Soviet Union. This collaboration is a powerful form of cultural diplomacy and a long-standing source of soft power for Western actors. More specifically, the particularities of the Russian data sciences make it an ideal candidate for collaboration. Supporting ongoing modernization efforts in the Russian data sciences promises tangible benefits for the international scientific community and knowledge economy more generally.

United States actors from the public and educational sectors, and the international scientific community more broadly, should return to a policy of active involvement with ongoing reforms in the Russian scientific and educational systems. Entering into these discussions as allies of reform-minded academics, providing logistical support and empirically-based, outside consultation would be a critical boost for those parties committed to returning the Russian system to a more meritocratic, agile, and contemporary form of the organization of academic labor. The United States in particular has a long history of more-or-less successfully managing industrial-academic relations in ways that provides for both basic research and actionable innovations. Supporting such management in Russia would benefit not only local actors but stakeholders in the knowledge economy the world over.

My research has shown that Russian academics are putting their unique traditions of fundamental mathematics to work in innovative and sophisticated ways, but are struggling to build the necessary institutional groundwork for their development into well-developed programs of research or concrete applications. Indeed, they are often doing so outside of existing state structures. International support from scientific organizations could ensure that these researchers' unique contributions begin to circulate more freely among their international peers, rather than ending up lost in the halls of bureaucracy or the state scientific establishment. Further, overcoming the current isolation of the nascent Russian computing sector would provide unique opportunities for economic collaboration in a heretofore sparse marketplace. As Russian elites attempt to supplement the petrostate economy with a more robust informational technology sector, the potential for Western companies to participate are enormous.

**Co-Curricular Activity:**

In addition to a great deal of research materials, my time as a Title VIII fellow has produced a number of fruitful collaborations with Russian colleagues. Perhaps most notably, the sociology department at the Higher School of Economics has been an excellent source of support, both intellectual and moral, for the current project. They are one of the few Russian departments with solid representation in the sociology of science and technology, and participation in their colloquia and informal conversations about my ongoing work has been invaluable. The department is acting as my host for the next stage of my dissertation research, from June-August 2015. The Center for Science and Technology Studies of the European University at St. Petersburg has also been an important locus of collaboration. Home to an ongoing research project on "Russian Computer Scientists at Home and Abroad," the EUSP group and I share a great deal of thematic interests. They have been crucial in facilitating interviews in St. Peterburg, and invited me to lead a short seminar on the Geography of Science for their graduate students in the fall.

**Conclusions:**

The experimental collaboration between academy and industry at the heart of my research has produced a number of concrete pedagogical innovations. These have been facilitated by the inclusion of members from industry in the academic council of the new department at the NRU. The council has a high degree of control over the content of individual courses taught in the

department, and these courses are being actively restructured to include more practice with applied tasks, greater training in a wider range of programming languages, and less focus on the fundamental mathematics underlying these tasks and languages. However, students and faculty alike remain committed to a rigorous, algorithmic understanding of their work, drawing on a long history of excellence in fundamental mathematics at all levels of Russian education. As a consequence, the current generation of young Russian data scientists are uniquely poised between the academic and the industrial worlds, seeing in industrial work the potential for highly intellectual, creative research, and viewing academic work in turn primarily as a prolegomenon, however necessary, to the real business of developing innovative applications.

The new research programs and types of researchers emerging in the Russian data sciences distinguish themselves from those in the West in a number of ways. First, most prosaically, different areas of mathematics are being leveraged to approach long-standing problems in areas such as machine learning in new ways. Second, despite the push for greater focus on applied tasks, the role of mathematics as a fundamental language for describing the work of information technology professionals is far more prominent. Everyone from academic researchers to low-level front-end developers have a great deal of respect for and competency in theoretical mathematics. However, they view it not as an end in itself or as a description of essential truths, but rather as a sophisticated form of universal language, as a *tool* for communication with peers, the assimilation of new forms of knowledge, and the development of new competencies.

The collaboration between academy and industry in the data sciences participates in, and is occasionally explicitly connected to, the broader commitment by policy elites to developing a

Russian knowledge economy. This commitment has manifested itself in the creation of technology start-up incubators, public-private collaboration on the development of innovative products, and the encouragement of applied research in a range of academic departments heretofore more focused on theoretical research. In the data sciences, however, we seem to see the most substantive effects of such policies on the actual content of education, and the most enthusiastic participation of the private sector in such schemes. Private-sector data science relies heavily upon the theoretical expertise produced by the university system, and the goal appears to be honing and supplementing this expertise for use in the private sector rather than radically restructuring computer science education into a form of trade education.

**Plans for Future Research Agenda / Presentations and Publications:**

Moving forward, I plan to wrap up this current project with return trips to Moscow from June-August of 2015, and from June-July of 2016. These visits will provide the opportunity to revisit earlier conversations with participants in light of changes in the geopolitical and economic situation, and to pursue questions which have emerged during the initial analysis of extant research materials. Funding for the 2015 trip has been secured through the Wenner-Gren foundation and Rice University in part on the strength of my current support through American Councils.

The first presentation of my research findings will be at the American Anthropological Association meetings in November 2015. Together with colleagues who are also researching university reforms, I have organized a panel investigating potential avenues of collaboration

between science studies and the anthropology of education. As I begin work on the dissertation, I plan to continue to present on this material at conferences such as the meetings of the Society for the Social Study of Science, the Society for Cultural Anthropology, and the Society for the History of Technology.

Tentatively, I am planning publications for both the specialized journals *Anthropology and Education Quarterly* and the *Journal of Eurasian Studies*, as well as the more theoretically oriented *Cultural Anthropology.* As I move forward with writing the dissertation and eventually my first monograph, I hope to continue to collaborate with my Russian colleagues and publish some of my materials in Russian-language sociology journals as well.