Research Statement of Vitalii Zhukov

My research interests are firmly within the realm of data science. In my doctoral dissertation, I investigated the patterns of human performance in unfettered competitions, drawing data from as diverse domains as Olympic athletes, WWII pilots, and faculty in US universities competing for NSF and NIH grants. My analysis, using advanced statistical modeling, pointed to a universal human performance law conforming to log-normality, provided the competitors are well-trained and there is sufficient time for the competition to evolve. Practically, this means that a minute percentage of super-achievers is the ultimate product of such competitions. When competitions do not produce such super-achievers, my analytical evidence suggests that either the people entering the competition are not well-trained or not enough time was given to the competition to evolve. This finding has implications across many domains of human activity, ranging from education to military training. It can be operationalized as an acid-test for the effectiveness of organizational policies focused on human development.

Furthermore, I studied in detail the competition between two famous schools of thought in psychology – affectivism and cognitivism, which are dedicated to the behavioral role of affective and cognitive phenomena, respectively. Analyzing nearly half a million relevant publications from PubMed with over 15 million citations, I found the citation impact of both affectivism and cognitivism to be log-normal, much like performance in all other domains I studied earlier. Affectivism, however, has been inching ahead of cognitivism impact-wise. This impact advancement is associated with higher level of topical diversity in the papers that cite affective publications versus those that cite cognitive publications. It appears that between two competing and maximally performing intellectual ecosystems, broader usefulness is the likely driver of gradual domination.

Many doctoral research lines are terminal – they die as soon as the student graduates. This is because such doctoral research lines look more like case studies rather than generalizable frameworks. My doctoral research is the latter! It provides the methodological framework for studying the state of competition in any domain of human activity, where the underlying mechanism of gradual domination may be different than that of affectivism vs. cognitivism. This puts me in an excellent position to develop a long-term research agenda, which is paramount to the success of a young faculty member. Importantly, the methodological skills I acquired in pursue of my doctoral research constitute a `Swiss knife' that I can use to expand my investigations in any domain characterized by big data and analytical challenges. These skills include data acquisition and wrangling, and advanced statistical modeling. I would like to emphasize the data acquisition and wrangling part, as few people have it. I did not simply analyze datasets, but I constituted the datasets myself, collecting data from very difficult sources, like the case of WWII pilots' victories scores. Such dataset construction did not only require initiative and investigative spirit, but also the ability to author data extraction scripts, complete with data quality and validity controls. Additionally, I would like to underline my expertise in large language models, which I developed to address research questions in affectivism vs. cognitivism, where I had to analyze a corpus with over one million papers.

To conclude, my in-depth and at breadth expertise across all layers of data science – i.e., acquisition, wrangling, statistics – plus the generalizable framework I inherit from my doctoral work, give me a significant advantage in a `hot' but very competitive research landscape. To that I add my solid team science credentials, as I collaborated harmoniously not only with my advisor but also with faculty from the University of California (Prof. A. Petersen) and Milano Politechnico (Prof. P. Tsiamyrtzis) during my dissertation development. Accordingly, I plan to submit proposals to the Computer and Information Science and Engineering (CISE) directorate of the National Science Foundation (NSF), forming teams with other colleagues from within and without the department. As I explained, my research framework and skills provide me with significant flexibility in choosing my targets, which will serve me well as NSF continuously unveils programs in new areas. Having studied and modeled human competitions for five years, I cannot begin to express my excitement and determination, as I am about to potentially enter a new competitive career as faculty in your department.