US-Serbia and West Balkan Data Science Workshop Final Report

Location: Belgrade, Serbia, Aug. 26-28, 2018
PI: Zoran Obradovic, Temple University (zoran.obradovic@temple.edu)

Background

Scientific collaboration between US, Serbia, and other countries from the West Balkan region has a long tradition starting in the 19th century and including contributions of world-renowned Serbian-American scientists like Nikola Tesla and Mihajlo Pupin. However, for many reasons such collaborations remained mostly at a personal level, limited to particular basic science fields (e.g. mathematics, biology, chemistry, physics). In 2010, Serbia and the United States signed the Agreement on Science and Technology Cooperation, which sets a framework for bilateral research collaboration between the scientific communities of the two countries. The aim of that agreement is to strengthen the scientific and technological capabilities of both countries, to broaden and expand relations between their scientific and technological communities, and to promote scientific and technological cooperation in areas of mutual benefit. Among the priorities of the foreseen collaboration are data science, mathematics, archaeology, anthropology, cultural heritage protection, etc. Multiple instruments of collaboration were proposed under the agreement, including joint research.

Motivated from this agreement, in December 2016, a meeting was organized by the National Science Foundation (NSF) Program Director (PD), Seta Bogosyan, NSF Office of International Science and Engineering (OISE), to explore specific collaboration opportunities with Serbian researchers. The Serbian delegation consisted mostly of leading scientists and researchers from Serbia (University of Belgrade, University of Novi Sad, Mathematical Institute of the Serbian Academy of Sciences and Arts, Archaeological Institute of the Serbian Academy of Sciences, Institute of Biological Research), Aleksandra Drecun, the President of Intersection Center for Science and Innovation in Serbia and Nikola Tanic, an Assistant Minister for Science in the Ministry of Education, Science and Technological Development of the Government of the Republic of Serbia. The host, NSF, was represented by Rebecca Keiser (Head of OISE), Erwin Gianchandani, (CISE Directorate Deputy Assistant Director), Lynne Parker (CISE IIS Division Director), Barry Johnson (ENG IIP Division Director), John Yellen (PD, Archeology), John Gillapsy (PD, MPS/Physics), Alex Simonian (PD ENG/CBET) and Tie Luo (PD, MPS/DMS), as well as Stephanie Morris (PD, NIH). The US agency representation was determined based on areas identified as research strengths of Serbia and the broader Balkan region (mathematics, archeology, biomedical informatics, engineering, big data, etc). This meeting revealed strong potential for synergy from the collaboration, and was hence followed by a visit of a group of NSF PDs to Serbia in the summer of 2017 which included meetings with Mladen Sarcevic (the Minister of Education, Science and Technological Development of Serbia), Zoran Popovic (Vice President of the Serbian Academy of Sciences) and Ivanka Popovic (Vice Rector of University of Belgrade). It was during this visit that a roadmap was established for a workshop in Belgrade and "data science" was determined to be the most appropriate platform to build a framework for larger-scale international and interdisciplinary collaboration between the US, Serbia and the West Balkan region.

Reasons for this collaboration choice include the large interest for data science in Serbia as technology companies have opened innovation centers in Serbia and have employed a number of well-educated young people in the data science field. Also, in 2015-2017 researchers from US and Serbia also organized three successful joint data science workshops in Belgrade, partially supported by the US Office of Naval Research, and these workshops included many researchers from other West Balkan countries thanks to strong existing scientific links in the West Balkan region.

Executive Summary

Following these and other ongoing outreach efforts between the two countries, this Data Science Workshop in Belgrade brought together a wide range of researchers from the US, Serbia, and West Balkan Region, addressing various challenges related to extracting value from data, including renowned experts, early stage investigators, and students from Serbia, the West Balkans, and the US.

The workshop was organized by Prof. Zoran Obradovic (Temple University) and supported jointly by NSF CISE and OISE. The objective of the workshop was to determine particular areas of demonstrated strength by Serbia and West Balkans, which also generate big data and require advanced knowledge discovery methods that would benefit from multidisciplinary international collaborations to further advance these fields with state-of-the-art methods and technologies of data science and mathematics.

The overall impression was that the synergy built by the collaboration framework of investigators in Serbian and Western Balkan countries with US colleagues would contribute to the advancement of these fields and data science. Some of the benefits of such a framework include:

- Serbia and the Western Balkan region have significant capacity to be involved in joint research projects with US partners within data science. These collaborations would benefit from strong research capacities of Serbia and the Balkan region in science, technology, engineering and mathematics (STEM) built on a very well-developed education system in these areas and track record of outstanding achievement in mathematics.

- Investment in research collaboration with Serbia and the Western Balkans will be especially cost effective as smaller initial investments will be able to fund high quality research in data science at a lower cost-to-results ratio as compared to many countries with comparable research capacities that have higher cost of living and therefore require larger research funding. In addition, bilateral research collaboration with a non-EU country like Serbia could be easier to implement by avoiding often restrictive EU regulations and associated costs.

- Serbia and the Western Balkans have access to unique data sets and environments for data science research, including a diverse and extensive range of archeological sites and populations for genetics research as well as federated education and critical infrastructure data warehouses and institutes and organizations actively collecting and analyzing such data.

The Serbian model of funding research is currently undergoing significant changes towards an open and competitive model. The Government of Serbia has submitted the Draft Law on the Science Fund of the Republic of Serbia to the Parliament, and it is expected that the Law will be adopted by the end of 2018. The new law anticipates a science foundation to encompass competitive criteria for research evaluation and foster research with high impact. This should enhance the cooperation capacity on the Serbian side but may also partially delay the implementation of bilateral projects between the US and Serbia since it is anticipated that the Fund would be fully operational by the end of 2019. The following general initiatives have been singled out as pre-requirements for the successful implementation of future collaboration between the US and Serbian research environments:

- Establishing a formal framework for collaborations and partnership between the US and Serbia, which informally exists between investigators in the US and Serbia, particularly between investigators in the cross-domains relevant for data science. This partnership can be formally agreed upon under the existing umbrella of the Agreement on Science and Technology Cooperation. The partnership should be developed taking into account the asymmetry between the partners involved.

- The national research funding and research performing agencies should explore legal and institutional options for the envisioned cooperation, as well as funding opportunities, all based on the favorable framework already established by the Agreement. Those efforts should focus both on identifying the existing concrete opportunities and programs for scientific exchanges, and increasing awareness on both sides of such schemes, as well as on exploring novel means of collaboration suggested by workshop participants – such as test beds, living labs, international research networks and hubs, and a joint research institute.

Several specific initiatives were suggested related to building living labs in Serbia and the West Balkan region to facilitate data science related interdisciplinary and international research activities in the area of smart and connected life. Towards this objective, contacts were identified at the USAID, World Bank and relevant industries to apply to for financial help in building the proposed living labs. Priority living labs topics were identified based on participants data science research aimed to facilitate decision making in healthcare, resilience in energy and ecology, and quality of education. Examples of living labs-based collaboration opportunities were discussed focused on integration of data science research and innovations activities in Serbia and West Balkans with complementary activities in the US. The overall objective of the proposed living labs is developing an experiential environment where data science investigators, policy makers and citizens from both sides would immerse themselves in designing, exploring, experiencing and refining new methodologies, policies and regulations in real-life scenarios. Following are three examples of living lab platforms that could be developed quickly with modest investments:

Balkan Regional Smart Grid Analytics Living Lab. The ongoing NSF Big Data SPOKE project on Smart Grids (Co-PIs: Mladen Kezunovic from Texas A&M, Zoran Obradovic from Temple Univ. and Santiago Carlos Grijalva from Georgia Tech) brought together over 200 investigators from more than 100 organizations at 4 continents that engaged in collaboration to identify the needs and challenges for big data research in smart power grids. This group is identifying data sources, developing the predictive data analytics for various smart grid applications, and providing infrastructure for evaluation and testing of new

solutions. Within this platform an interdisciplinary team at Texas A&M University and Temple University (Co-PIs: Mladen Kezunovic and Zoran Obradovic) is developing a real-time predictive framework aimed to improve decision-making in operations, operations planning, planning and market transaction for improved power system resilience. They are interested in federating their large-scale testbed for phasor measurement units (PMUs) with a Balkan region PMU living lab to jointly study synchrophasor-recorded disturbances as to develop advanced machine learning based situational awareness tools for a smart power grid and to characterize technical and economic benefits in context of state-of-the-art cyber security best practiced while also considering ecology and climate-related aspects. PMU sensors are currently installed in a power grid in Slovenia, Croatia, Serbia and Macedonia and high-level executives at power systems utilities from these countries recently expressed interest in collaborating with Kezunovic and Obradovic team on such a project. The federated Smart Grid Analytics living lab would allow a more comprehensive resilience research and pro-active maintenance and operation of power system infrastructure upon evolving events both in the US and in Europe, potentially making the power system more resilient and less costly to operate.

Serbia Learning Analytics Living Lab. Improving graduation rate while also improving quality of college education is a high priority challenge in Serbia as well as in the United States. Data scientists from Temple University and the University of Belgrade have shown interest in jointly addressing this high-impact problem by extending collaboration of research teams at Zoran Obradovic's and Boris Delibasic' laboratory, established eight years ago. Their objective is to build a learning analytics living lab at the University of Belgrade to facilitate large-scale integration of proactive decision-supporting tools into the academic advising process. The University of Belgrade (UB), the oldest and the largest university in the West Balkans, is home to almost 100,000 students. At UB, data is continuously collected and analyzed about course and exams quality and corrective measures are implemented if large discrepancies are noticed. In addition, at UB projects are being undertaken to track student activities in the classrooms (by building portals for student-professor interactions), and also to track students' movements through the university (presence at labs, lectures, etc.). This platform can nicely complement an evaluation environment built at Temple University where risk-based advising interventions are characterized on a dataset of about 24,000 undergraduate students in collaboration of data scientists with a team of academic advisors and information technology experts (PI: Z. Obradovic). The purpose of the Serbia learning analytics living lab would be to study how to effectively transfer discovered risk factors learned in one environment for cost effective risk monitoring in another environment (e.g. another university) where many factors of student dropout risk and declining academic performance may be different, but some important aspects are shared.

Serbia BioScience Living Lab. Serbia was a major crossroads of migration between Asia, Europe and Africa, and provides an unprecedented snapshot of human diversity in ancient times. An interdisciplinary team based primarily at Temple University (Co-PIs: Allison Hayes-Conroy, Rob. Kulathinal, Sudhir Kumar, Ralph Horwitz, Michael Henderson and Zoran Obradovic) has interest in partnering with a team in Serbia towards developing a *BioScience living lab* based on human diversity of ancient DNA available in Serbia to jointly study integrated social, environmental, and biological factors of embodied life, which has been hindered by the intricacy of the human relationship to surrounding social and physical environments – a relationship with emotional, political, economic, and ecological

dimensions. In this project, diverse features of life will be collected at the Serbia BioScience living lab using wearable devices, tracking software, social media, omics, electronic health records and other advanced instruments. Integrated data will be analyzed with the objective of making scientific discoveries potentially leading to new targeted treatments for a wide variety of complex diseases known to have environmental and genetic factors, as well as analytical improvements in how data is used in life sciences and related fields, innovations in techniques for collecting sensitive biographical data, and overall extension of the scientific evidence base for development and decision making in the area of human health.

A few hundred participants were physically present at the workshop, but there was a further global presence, as the workshop was live streamed to the entire world, with the total number of accessed pages showed at Figure 1.
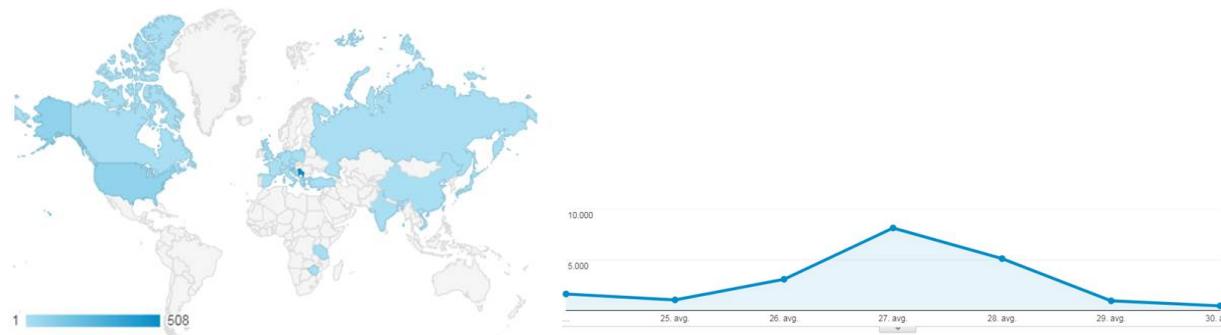

Figure 1: Spatial and temporal coverage of the Data Science workshop live streaming (728 participants in North America, Europe, Asia and Africa accessed 35,383 pages)

The workshop general chair was Z. Obradovic (Temple University, USA) while co-organizer and engagement chair was A. Drecun (Intersection Center for Science and Innovation, Serbia) and local co-chairs were Z. Ognjanovic (Mathematical Institute of the Serbian Academy of Sciences and Arts, Serbia) and S. Ghilezan (University of Novi Sad, Serbia). The field trip host was M. Korac (Archeological Institute, Serbia). The workshop had 2 keynotes on the following topics:

- "Data Science: Challenges and Opportunities when Playing in Everyone Else's Backyard" (Z. Obradovic, Temple University, USA)

- "On the Importance of Causal Models in Making Sense of Big Data" (V. Honavar, Pennsylvania State University, USA)

The program (see https://nsfserbia.rs/), included 39 invited lectures, 101 poster presentations and 3 panels grouped in the following focus areas: Foundations, Biomedical Informatics, Critical Infrastructures, and Digital Archeology.

Each of the four tracks has provided a report on the outcomes and identified the potentials and opportunities for collaborative research between the US and West Balkan researchers (included as 4 sections of this report).

Track A Report: Data Science Foundations

In this section we summarize the outcomes of the Data Science Foundations Track and identify the potentials and opportunities for collaborative research between the US and West Balkan researchers.

This track was co-chaired by A. Scedrov (Univ. Pennsylvania, USA), Lj. Kocarev (Macedonian Academy of Sciences and Ss. Cyril and Methodious Univ., Macedonia) and B. Delibasic (Univ. of Belgrade, Serbia. Eight invited talks were delivered to present current research possibilities and research strengths from the US & Serbia and Western Balkans in Data Science Foundations.

Three invited talks were delivered from US researchers:
1. "Towards Explainable and Efficient Learning Systems" (S. Parhasarathy, Ohio State University)
2. "Just Machine Learning" (T. Eliassi-Rad, Northeastern University)
3. "Structured Prediction: Data Analytics Meets Applications" (H. Rangwala, George Mason University)

Two talks were delivered from researchers from Serbia and Western Balkan countries:
4. "Meta-Model Framework for Surrogate-Based Optimization with Machine Learning" (Lj. Todorovski, J. Stefan Institute and Univ. of Ljubljana, Slovenia)
5. "Continuous optimization methods in distributed environment" (N. Krejic, Univ. of Novi Sad, Serbia)

Three talks were delivered from PhD candidates from Serbia and the Western Balkan countries that are doing their PhD studies in the United states
6. "Generalization-Aware Structured Regression towards Balancing Bias and Variance" (M. Pavlovski, Temple Univ. USA and Macedonian Academy of Sciences, Macedonia)
7. "Gaussian conditional random fields extended for directed graphs" (T. Vujicic, Mediterranean Univ., Montenegro)
8. "Fast Learning of Scale-Free Networks Based on Cholesky Factorization" (V. Jelisavnic, Mathematical Institute of the Serbian Academy of Sciences and Arts, Serbia)

The US researchers pointed out the importance of building fair, interpretable, and efficient machine learning algorithms. The value of building structured machine learning algorithms was also emphasized.

Researchers from Serbia and the Western Balkan countries presented research in enhancing the possibilities of machine learning algorithms with innovative approaches. Learning of algorithms in environments that are distributed and where data is protected was also emphasized.

Young researchers presented their ideas on building structured machine algorithms that improve the current state-of the art in science.

The overall impression of this track is that the ongoing research activities in the area of data science foundations are compatible, and that there will be potentially large benefit of developing additional collaboration mechanisms to support joint research among investigators in Serbia and Western Balkan countries with US colleagues working in this area.

Track A unique research opportunities in Serbia

An invited talk by Z. Markovic, a former director of the Mathematical Institute of the Serbian Academy of Sciences and Arts, provided an overview of the mathematical research eco-system in Serbia. In the last decades, researchers from Serbia have achieved prominent results in mathematics, especially in the areas that represent the foundations of data science (mathematical logic, optimization, graph theory, among others). These achievements are the result of the unique system of organization of research and education in mathematics.

The Mathematical Institute of the Serbian Academy of Sciences and Arts is a unique institute in the field of mathematically oriented research, not only in Serbia, but also in the former Yugoslavia. Within the Institute, a large number of national projects in mathematics and computer science are coordinated, as well as significant regional and international collaborations within Horizon2020, FP7, FP6, COST, Erasmus+, Tempus and bilateral frameworks. These projects bring together numerous interdisciplinary teams with over 400 researchers (experts, postdocs, and doctoral students) from all main Serbian universities. The positions Serbia holds on world-ranking lists in the above-mentioned area (e.g., https://www.scimagojr.com/countryrank.php) are generally between the 30th and 40th place in the world, which is totally disproportionate with respect to the size of the country, the number of researchers, the national GDP, and funding of science. Since the 1960s a considerable number of notable Serbian mathematicians have obtained their PhD degree in the USA and/or have had their postdoc positions in the USA, and many of them have/had tenure positions at prestigious US universities.

Currently, the strongest groups in Serbia in mathematical logic are focused on non-classical logics and formalization of uncertain reasoning, formal methods, and computer-assisted formal reasoning. By developing various logics, the first group addresses some of the important issues in AI, for example: modeling the behavior of intelligent agents in the presence of incomplete or uncertain information. In formal methods, research has been spread from the design of formal models for sequential, concurrent and distributed systems to their application and implementation. Most recent research interests focus on privacy in linked data, formal methods for data privacy in large scale systems, and formal methods in AI.

The main goal of the group for optimization is to consider real-world large-scale systems that need to be improved using mathematics and optimization methods, particularly using the method of Variable Neighborhood Search, proposed by a Serbian researcher, and further developed worldwide. The group for graph theory is mainly focused on the theory of graph spectra and applications related to chemistry and study of topologies of social networks. The research group

for high performance computing based on a dataflow platform has a system installed at the Mathematical Institute of the Serbian Academy of Sciences and Arts that is also used in courses by several US universities.

At the University of Belgrade – Faculty of Organizational Science at the Center for Business Decision Making, researchers are doing unique research for developing data analysis algorithms and methods for emerging datasets that are not largely open to the scientific community. In 2011, the group analyzed datasets from ski lift transportation records from the largest Serbian ski resort, Kopaonik, for a period of six years (2006-2011). Research using this data has resulted in several research papers published in highly ranked journals, and several scholar visits to the United States. Over time, the group has obtained access to several major Serbian datasets, like human trajectory data from shoppers in Serbian shopping malls (in 2018), data from Serbian road traffic accidents (from 2011), data from cell phone traffic (in 2018), and traffic data from Serbian highway toll ramps (in 2018), which all open opportunities for joint collaboration between the US and Serbia & Western Balkan countries, as several of those datasets are very hard to obtain globally.

Data privacy is one of the most urgent issues of modern society. Due to the tremendous developments of information technologies, new disciplines emerge rapidly with application in all segments of life and society. The fast changes on the applicational side permanently urge for the development of foundational aspects based on solid theoretical concepts. Technology raises new privacy concerns. Logical methods have been ubiquitous for foundations of mathematics, physics and computer science. Mathematical models and formal methods have become the base tools in computer science for developing reliable software and hardware. New paradigms of information technologies, such as internet of things, cloud computing, and blockchain, also require reliability that can only be provided by mathematical models. The complexity of this problem requires multidisciplinary teams of mathematicians, computer scientists, information scientists, lawyers, sociologists and psychologists. It is necessary to encourage mathematical and multidisciplinary research to address issues of data privacy. These topics have been studied independently by different groups of US and Serbian scientists involved in Track A and Track B. The interaction between the groups started in 2017. In informal discussions during the meeting, we have identified this topic as a good possibility for collaborative research. We have publicly announced this topic of joint interest during the brain-storming panel on the last day of the Workshop.

Track A existing collaborations

The Workshop chair of Track A on foundations of data science, together with the two local co-chairs of the Workshop, are co-directors of a series of very successful workshops and schools of Logic and Applications - LAP (http://imft.ftn.uns.ac.rs/math/cms/LAP2018). The meeting, which has been organized annually in Dubrovnik (Croatia) since 2012, gathers prominent researchers from around the world along with early stage researchers from the region. The LAP meeting is partly tailored to be a summer school for PhD students from the region. This well-established meeting is a good opportunity for US doctoral students and early stage researchers to gain knowledge, exchange experience, and create collaboration networks in this part of the world.

There is a long-standing fruitful individual collaboration between US and Serbian researchers in foundations of data science, which has resulted in a series of joint publications, communications

at conferences, and visits. These existing collaborations between established scientists are a good opportunity to exchange doctoral students and postdocs on a regular basis between the research groups in this foundational area and to promote collaborative research. This good practice provides a basis for collaboration on a wider scale.

One interesting example of a fruitful collaboration is the collaboration Temple University (PI: Z. Obradovic) has established with the University of Belgrade – Faculty of Organizational Sciences in 2009. This collaboration enabled obtaining a Fulbright Visiting Scholar Grant in 2011/2012 from Serbia to the USA for a visiting scholar in Z. Obradovic's lab at Temple University. After this visit, in 2012, a dual doctoral degree agreement was established between Temple University and the University of Belgrade which allows obtaining dual doctoral degrees from both Universities, but also fosters scientific collaboration and joint research possibilities between academic institutions. Since 2012 Zoran Obradovic has provided several visiting scholar opportunities for scholars from the University of Belgrade, and several joint doctoral degrees have been realized. The collaboration also resulted in multiple joint research publications and in organizing multiple joint scientific workshops, financed by the Office of Naval Research, in 2016 (South East European Data Science Forum), 2017 and 2018 (Computational Decision Making and Data Science Workshop).

Track A strengths in the local research community

Strong research capacities in science, technology, engineering and mathematics (STEM) in Serbia are built on a very highly developed education system in these areas.

Serbia has several centers for excellence in educating students in Mathematics. Mathematical faculties are present at all major Universities across Serbia. Mathematics is a unique science in Serbia where four universities (Belgrade, Niš, Novi Sad, Kragujevac), out of five public universities in Serbia, are among the top 500 in the Shanghai Mathematics List for the year 2017 (Shanghai Ranking's Global Ranking of Academic Subjects 2017 - Mathematics). Serbia has an educational system that is distributed across the whole country in the elementary and secondary school system. The Mathematical Gymnasium in Belgrade is an outranking school in Mathematical education that allows talented students from all over Serbia to get the best mathematical secondary school education in the world. This secondary school has students that regularly achieve the biggest successes on international mathematical competitions. The best students from the mathematical gymnasium are often awarded with free scholarships at the best higher education institutions worldwide, e.g. Oxford, Cambridge, etc. Similar mathematical oriented special classes exist in gymnasiums in Novi Sad, Niš, Subotica and Senta.

Track A risks, limitations and weaknesses

Serbian universities generate each year very capable young minds in STEM areas. Due to their capacities they enroll in graduate schools at prestigious universities worldwide. However, the vast majority of them do not find suitable jobs and research positions in Serbia after graduation. This kind of brain-drain is one of the major issues the Serbian research community in STEM is facing.

Track A opportunities for future collaboration

The following activities were proposed for future collaboration:

To support education and cross-training (e.g., in the form of summer schools) of students and more experienced specialists in interdisciplinary fields, relevant for mathematics and data science. As one potential source for this activity, we have identified the NSF program IRES – International Research Experiences for students. A higher-level NSF program is to be discussed.

To establish a regional research center in mathematics and data science, which will support and boost excellence in research between the US and Serbian scientists as well as regional and international collaboration.

The pre-requirement for the successful implementation of future collaboration between the US and Serbian research environments, as stated in the Executive Summary, is to establish a formal framework for collaborations and partnership between the US and Serbia under the umbrella of the Agreement on Science and Technology Cooperation, which has been signed by Serbia and the US in 2010. Collaboration exists on the individual level between investigators in the US and Serbia, particularly between investigators from the US and internationally recognized fundamental science research groups in Serbia and the region that address problems relevant for data science.

Track B Report: Data Science in Critical Infrastructures

This track was co-chaired by N. Chawla (Univ. Notre-Dame, USA), Y. Kompatsiaris (Center for Research and Technology, Greece) and M. Mihaljevic (Mathematical Institute of the Serbian Academy of Sciences and Arts). As a background for consideration of possible joint research activities related to data science in critical infrastructures, eight talks have been delivered: Four presentations were by experts from the US, two by experts from the region (Macedonia) and two from Serbia. The invited talks addressed the following topics:

1. "Intelligent Climate Adaptation and Resilient Engineering (I-CARE)" (A. Ganguly, Northeastern Univ., USA)
2. "Graph Signal Processing" (J. Moura, Carnegie Mellon Univ., USA)
3. "A Sub-linear Time Algorithm for Substring Matching in Big Data" (K. Narayanan, Texas A&M Univ., USA)
4. "Big Data for Urban Traffic Management Control (UTMC) systems in Smart Cities"(I. Mishkovski, S. Cyril and Methodious Univ., Macedonia)
5. "Critical Infrastructure Challenges: Paradigm Shift in Electrical Power Systems – A Prospective from Belgrade" (S. Vukosavic, Univ. Belgrade, Serbia)
6. "The role of industry, occupation, and location specific knowledge in the survival of new firms" (C. Jara-Figueroa, Massachusetts Institute of Technology, USA)
7. "An ICT framework for addressing sustainable development" (V. Zdraveski, S. Cyril and Methodious Univ., Macedonia)
8. "Critical Infrastructures Research System in Serbia" (M. Mihaljevic, Mathematical Institute of the Serbian Academy of Sciences and Arts, Serbia)

The delivered talks illustrate certain related and complementary research interests of US and Serbia/Region experts, which could be considered as a fruitful initial position for collaboration and joint projects. The talks have addressed the following important topics of data science:

- Applications of Big Data and Data Science in engineering for electrical power systems, environmental issues, transportation, telecommunication;
- Big Data & Data Science Management and Frameworks including distributed data bases, knowledge management and engineering, novel data models, novel data management methods;
- Security & Privacy in the Era of Data Science & Big Data: cryptography; threat detection using big data analytics; privacy threats of big data; socio-economical aspect of big data in the context of privacy and security, blockchain technology.

As illustrative examples we point out the following talks.

Talk [2] has been delivered by the President of IEEE, pointing out top level encouragement from the US for joint research activities within data science relevant for critical infrastructures.

Talk [1] addressed climate change adaptation for infrastructure, pointing out that data science techniques for climate change adaptation are critical for civil, agricultural and environmental engineering, as well as for communities in the 21st century, including solutions for efforts that maximize the performance and life-cycle cost-effectiveness of climate-ready infrastructure investments.

Talk [4] has addressed data science issues in transportation as another illustrative critical infrastructure and points to certain interesting outcomes of the performed activities.

Talk [5] has pointed out interest and potentials of Serbia regarding a worldwide hot topic: Applications of Data Science for developing advanced electrical power systems. In particular, the smart-grid/micro-grid approach is addressed and importance of data management and applications in engineering are pointed out, including data protection as well as data integrity and privacy issues.

As a very important issue, talk [8] has pointed out the significant capacity of Serbia to be involved in joint research projects with US partners within data science for critical infrastructures. This talk shows that there are internationally recognized research groups at the universities, research institutes and certain business entities with activities and respectful achievements within data science for critical infrastructures. These achievements include over 1,000 internationally recognized results published as journal papers, book chapters and conference papers, with over 10,000 citations in the international publications - the areas of these results include: electric power systems, information-communication systems, data management systems, health care and medical systems, and cyber security.

As the concluding notes, we point out the following comparative advantages of Serbia and References for collaboration with US. It is worth noting that, based on publications and their citations, Serbia appears as a regional leader in basic research for operational research over big data, cyber security, and big data processing for electric power and information-communication systems. Also, within the mentioned research domains, Serbia has established a certain level of international collaboration with EU countries as well as some overseas countries like the US and Japan. Accordingly, certain groups in Serbia have proven skills for establishing additional collaborations and extending the existing joint research activities with US partners to the benefit of both sides. A particular interest of the US could be collaboration focused towards priority research directions for the US where Serbia has high-skill researchers and the US needs additional work power for fulfilling challenges within certain critical infrastructures.

Accordingly, it is concluded that Serbia could be a desirable partner of the US in certain basic research activities of high interest for the US where Serbia has certain advantages in comparison with a number of other countries. These comparative advantages of Serbia include: (i) significant number of high-skill internationally recognized experts; (ii) successful participation in many international realized projects; (iii) lower requirements of theoretical research funding. Significantly lower requirements for theoretical research funding and very high rate of outcomes/funding in comparison with the most developed countries could be also considered as a unique opportunity for performing research activities in Serbia and the West Balkans.

Track B unique research opportunities in Serbia and the West Balkans

It could be claimed that the domain of data and critical infrastructures is universal, and it is not easy to claim that any part of the World provides unique research opportunities in this domain. On the other hand, the following feature could be considered non-universal. Let research activity within data and critical infrastructures be decomposed as follows: (i) basic research; (ii) simulations and (technological) development; (iii) testing in a real environment. In the previous decomposition and taking into account the strong technical and mathematical capabilities among Serbian researchers, partners from Serbia and the West Balkans could greatly contribute to collaborative work in activities (i) & (iii) in their local environments and collaborating on issue (ii) in the US.

In addition, Serbia is almost unique as a relatively small country with four (80%) national universities ranked among first 500 in the Shanghai ranking list of world universities in the area of mathematics, which is a key area in data science – This fact implies a strong potential of Serbia for performing basic research in the data and critical infrastructures domain.

Track B potential NSF concern

The main challenge of a collaboration would be a short record of already-performed joint activities within the NSF framework.

Track B specific further steps

Currently, the main problem in Serbia, which is also at the root of the region's previously described brain drain, is insufficient stimulative funding of research projects as a consequence of the overall economic environment. Consequently, a number of discussions at the workshop were focused towards establishing mechanisms which could provide additional external funding of high-skill researchers, who could work on projects addressing topics of high priority for the US where experts from Serbia could perform a number of research activities at a lower research cost in comparison to the US. Establishing such a framework will lead to win-win collaboration between the US and Serbia.

Accordingly, it appears that future steps should be built over the following framework: (i) research bilateral projects selected to be highly beneficial for all parties involved; (ii) partnership-based selection of the projects; (iii) asymmetric funding of the projects, which reflects asymmetric economic potential of the project partners.

As one of the topics for a joint project of high potential interest for both partners, the following is pointed out: "Advanced Techniques for Deployment and Business Operations of Smart/Micro-Grid Employing Data Science Paradigms". The project addresses the following challenges, where solutions provide an advanced framework for Smart/Micro-Grid electrical power systems:

(a) Advanced processing of big data obtained from the sensors for efficient control and management of Smart/Micro Grid;

(b) Advanced information-communication infrastructure for data transmission and storage (including dedicated error correction encoding/decoding);

(c) Advanced solutions for privacy and security issues;

(d) Advanced approaches for business management including employment of smart contracts and blockchain technology.

(e) Development of a living labs framework for achieving the identified critical infrastructures related goals as a user-centered, open-innovation ecosystem.

Please note that the talks [1]-[8] provide supporting elements for the above proposal. In addition, following the NSF Data Science workshop, on November 26, 2018, a Serbian national meeting on "Energy and Climatic Changes" was organized in Belgrade by the Serbian Academy of Sciences and Arts, with participants from academia, industry, and government. The meeting has recognized the importance of the so called "Triple D": De-carbonization, Distribution, and Digitalization. Digitalization should provide synergy of energy technologies and information-communication technologies. In particular, digitalization assumes the presence of a framework for: (i) big data collection and sophisticated data processing, and (ii) the so-called Living Labs, where data are collected, and smart energy solutions are tested and developed in a transparent ecosystem. This initiative can provide complementary support for follow up activities related to data science in critical infrastructures.

Track C Report: Biomedical Informatics

This section summarizes the outcomes of our interactions in the area of Biomedical Informatics, identifying several unique collaborative opportunities between the US and West Balkan scientists.

This track was co-chaired by P. Radivojac (Northeaster Univ., USA), G. Stiglic (Univ. of Maribor, Slovenia) and N. Veljkovic (Institute for Nuclear Sciences "Vinca", Serbia). As a background for consideration of possible joint research activities related to data science in biomedical informatics eight talks were delivered to present current research possibilities and research strengths from the US & Serbia and Western Balkan.

Three invited talks were delivered by US researchers:

1. ''From Genomics to Therapeutics: Uncovering and Manipulating the Genomic Circuitry of Human Disease'' (R. Kulathinal, Temple University, USA)
2. ''Human Aware Artificial Intelligence'' (S. Natarajan, Univ. Texas at Dallas, USA)
3. ''Deeper understanding of microbiomes as a benefit of forgetting microbial names'' (Y. Bromberg, Rutgers Univ., USA)

Three invited talks were delivered by researchers from Serbia and Western Balkan countries:

4. ''Learning to annotate from multiple data sources and representations'' (T. Smuc, R. Boskovic Institute, Croatia)
5. ''Automatic annotation of the human proteome: Prediction of protein-protein interactions and disease phenotype associations'' (V. Perovic, Institute Vinca, Serbia)
6. ''Infant Neuromotor Development and Childhood Problem Behavior – sensitivity to non-ignorable missingness'' (E. Tahirovic, International University of Sarajevo, Bosnia and Hercegovina)

One invited talk was delivered by a PhD student from Serbia that is doing her PhD in the United States:

''A pilot cognitive computing system to understand immunization programs'' (M. Stanojevic, Temple University, USA)

One invited talk was delivered by a PhD student from Serbia that is doing his PhD in Serbia:

''Integration of Domain Knowledge into Machine Learning Algorithms – Applications in Healthcare'' (S. Radovanovic, Univ. of Belgrade, Serbia)

The US researchers pointed out the importance of developing systematic technologies for unraveling genetic characteristics that underlie complex diseases and cancer. It was emphasized that insights gained through systematic approaches have significant potential to be translated into therapeutic treatments. One US speaker presented a reasonable model of human interaction with machine learning algorithms, where human advice is combined with data. New approaches to microbiome research aiming to map functionalities of condition-specific microbiomes were also presented. Researchers from Serbia and the Western Balkan countries presented new machine learning based approaches to annotations of biological data. Young researchers presented their work on applications of machine learning algorithms in healthcare.

Track C has shown that researchers from both sides tackle similar present-day problems with significant complementarities in their approaches. Our overall impression is that stronger cooperation between them would contribute to increased knowledge and better applications in the area of biomedical informatics.

Track C unique research opportunities in Serbia and the West Balkans

Serbia and the West Balkans offer exciting opportunities for collaborations intersecting the biomedical and data sciences, which have the potential to benefit both the region and the global community. In the areas of genetics and genomics, the emergence of relatively inexpensive whole-genome sequencing during the past decade has greatly improved our understanding of extant human variation, our genetic histories, and clinically actionable genetic information. However, smaller local populations remain underrepresented in large-scale studies, thus, presenting opportunities to discover new variants and markers of demographic, evolutionary, and biomedical significance. According to the 1,000 Genomes Project, 10% of variants observed in a

population are private to that population. Identifying population-specific variants in any group will improve the identification of recent gene flow between populations, the estimation of ancestral demographic patterns, and the characterization of germline and somatic disease-associated variants in a wider context. The Balkan region contains a disproportionately high number of ethnic populations with ancestral connections to populations including northern Europe and the near and far East. Ideally, the collaboration between regional and leading US scientists would lay the groundwork to ultimately result in ethnically relevant reference genomes, as has been performed, for example, in the Korean population, which would be subsequently integrated into precision medicine efforts including diagnostics, treatment, and prognostics related to individuals of Serbian and West Balkan origin around the world. Preliminary research in this direction is already ongoing, however, an important milestone that will enable this collaboration is advancing the recently formed private-public partnership to deep-sequence 1000 Balkan genomes. While these projects target extant individuals from contemporary populations, this Data Science Workshop also revealed exclusive opportunities to sequence ancient DNA from several recently uncovered Roman settlements, such as those in Viminacium and other sites in Serbia that could shed new light on the lives of regional populations in the past 2000 years, including their demographic and biomedical histories. Similar opportunities have been identified in the area of microbiome studies that can contribute to new discoveries of global importance. In terms of individual and population health, the mix of rural and urban populations, its cultural diversity, and the wide-spread adoption of smart phones in Serbia present an excellent opportunity for mHealth and population health studies, again with both regional and global significance.

Track C strengths in the local research community

An invited talk by N. Tanic, PI at the Institute for Biological Research "S. Stankovic," in Serbia, provided an overview of the biomedical research eco-system in Serbia. Biomedical research capacities in Serbia are spread among four state universities (University of Belgrade, University of Novi Sad, University of Niš and University of Kragujevac) and several public research institutes (e.g., Vinča Institute of Nuclear Sciences, Institute of Biological Research "Siniša Stanković", Institute of Molecular Genetics and Genetic Engineering). Serbia has a long tradition in molecular biology and biomedical research with a respectable track record in receiving grants from European funding agencies including Horizon 2020, COST and Erasmus+. The estimated number of researchers in the field is about 1,300, and among them are research groups skillful and experienced in the molecular genetics of disease, population genetics, and various 'omics sciences.

Researchers working in the field of biomedical informatics in Serbia are strongly interconnected both informally and formally through two recently formed scientific societies: The Serbian Society for Molecular Biology (MolBioS) and Serbian Society for Bioinformatics and Computational Biology (BIRBI). The biennially held Belgrade Bioinformatics Conference (BELBI) has evolved over the years to become a leading event in the region that provides a platform to present research results and a forum for networking and collaboration. The conference was initially organized in 2012 and since then has gathered numerous leading researchers from Europe, the United States, and elsewhere. Several participants of this Data Science Workshop have already taken part in BELBI, reflecting the interconnectedness and scientific aspirations of the local research community. Balkan science strives to preserve and

improve expertise in a deeply engrained tradition of analytics that includes strong mathematical foundations, algorithm/software development, and modelling in biophysical and biomedical domains, while also demonstrating its readiness to adapt to the challenges of big data analytics.

The University of Belgrade, among the best ranked universities in this part of Europe, provides training in many traditional and several emerging disciplines. Instruction in biomedical data science is performed at both the masters and doctoral level at the Faculty of Biology, the Faculty of Mathematics, and the School of Electrical Engineering, which have jointly developed courses in Genome Informatics with one of the global leading biomedical data companies, Seven Bridges Genomics, also located in Belgrade. Undergraduate and graduate students represent a valuable source of human capital in Serbia, as many of them are involved in research activities and even in international collaborations. During the Data Science meeting, we learned about their work in direct communication and from the presentations during poster sessions.

Education in science, technology, engineering and mathematics or STEM subjects is of primary importance across the region. In Serbia, mathematics holds a special place in STEM, as centers of excellence in teaching mathematics and computer science have been developing at the forefront of these fields for decades. They deliver students with high STEM abilities at the end of their compulsory education and this human capital that has been growing cumulatively over a long period of time is dispersed across various disciplines. We noted that biomedical research attracts many of these brilliant minds that are eager to tackle big challenges.

Track C risks, limitations, and weaknesses

The United States provides a successful model for funding the most impactful research in an open and competitive manner. Serbia and other countries in the region have yet to establish similar funding mechanisms. A formation of an organization similar to the NSF in Serbia (The Draft Law on the Science Fund of the Republic of Serbia is submitted and is expected to be adopted by the end of 2018), with equivalent effective procedures and funding mechanisms, presents an important step forward. However, even with such an organization, the reality is that the funding of joint projects may be imbalanced due to the sheer differences in the scale of the economies of these two regions.

Track C opportunities for future collaborations

This three-day Data Science Workshop provided numerous opportunities to discuss and conceive possible collaborations between US and Serbian/Balkan scientists. Proposed collaborations in the area of genomics and genetics will offer a unique opportunity for US scientists to leverage local expertise and genomic sequence data from the past and contemporaneous Balkans. Another topic that gained keen interest among participants was the investigation of the microbiome. The major challenge in this field will be to define a healthy, physiological microbiome whose key determinants are associated with dietary habits that are specific in different parts of the world. Given that, data from regional populations should be acquired and mapped against the rest of the world to continue to build a global picture. Intersectional areas such as genomic anthropology would provide another exciting direction, where the expertise from biomedical informatics and digital anthropology could lead to increased knowledge about this critical migration route that linked many parts of the ancient world.

Digitalization and standardization of medical records in Serbia was also identified as an important next step towards deploying artificial intelligence (AI) towards computer-aided clinical decision-making. Such activities would open up the possibilities for projects related to natural language processing of the Serbian language and integration of these technologies with the global technologies and knowledge in precision medicine, including drug efficacy, drug-drug interactions, understanding and implementing best treatment practices and management of chronic disease, and ultimately pharmacogenomics and pharmacokinetics with the addition of genotyping.

Serbian and regional scholars have developed expertise across a variety of biomedical informatics fields and, in particular, have demonstrated noteworthy capabilities in building efficient algorithms. Several topics related to AI and its application in biomedicine have emerged as potential collaborations, including the utilization of compressed data without decompression and integration of information from different biomedical resources, causal modeling, etc. Automated annotation of function of biological macromolecules (individually and in communities) and high-resolution functional effects of genomic variation are research themes that were distinguished as common interests on both sides. Similarly, we observed discussions about mutual interests in biocuration, which is inherently a multidisciplinary field in which the AI community has to contribute much in the way of formal logic, knowledge representation, and probabilistic modeling. On the other hand, life science researchers underpin the annotation of human diseases, phenotypes and gene roles. During the Data Science Workshop, we identified specific projects related to the interpretation of genetic variants (including inference of detailed ancestry), including the relationship between genomic diversity and function of biological macromolecules. We could use these projects as a starting point that could be undertaken collaboratively and lead to presentations and discussions during future workshops and summer schools.

Student exchange programs and short-term visiting faculty appointments have been identified as an essential avenue of collaboration in the area of biomedical data science. During the WS, participants from Temple University and Vinca institute of nuclear sciences has agreed on joint small-scale research projects that will involve students from both sides supervised concomitantly by the US and Serbian scientist. Sessions will be organized remotely, topics will encompass biological data analyses and the initiative will start at the beginning of 2019. Also, these exchanges may involve summer schools and workshops to be held in Serbia and the West Balkans and to be attended both by US and local students collaborating in the frame of summer projects which will leverage local data sources. Lectures by regional and US scientists would contribute to the training of an already analytically-strong student population, strengthen local expertise, and eventually widen the talent pool in the field. A few specific topics have already been identified among WS participants as pilot studies to tackle population-specific genetic architecture and gene flow throughout the Balkan region. Not only are joint publications and conference presentations expected outcomes, but also long-term collaborations and exchanges. We have identified the NSF program International Research Experiences for Students (IRES) as one potential source of support for this summer school.

Track C summary

With the generation of "omics" and biometric data exceeding Moore's Law and computational resources allowing us to manage and analyze more data than ever before, a number of prominent

researchers have proclaimed that this is the "Century of Biology" [4]. This Data Science Workshop demonstrated an extraordinary potential to develop the emerging field of Biomedical Informatics by leveraging unique resources in Serbia and the West Balkans, which include the availability of extant genomic and microbiome data from several ethnic populations, extinct ancient genomes found on a historical global migratory route, current expertise and capabilities in building efficient algorithms, and a long tradition of world-class training of students in STEM. A few years ago, a member of the US team (Predrag Radivojac) has proposed an initiative to define an ethnic reference genetic panel for Serbia and the Balkan region. The genetic architecture of the region is perceived as extremely complex and the initial discussions have involved experts from Indiana University and Northeastern University in the US, whereas as Serbian representatives, scientists from the Vinca Institute of Nuclear Sciences, Institute of Molecular Genetics and Genetic Engineering and Faculty of Mathematics, University of Belgrade have been included. The activities to reach this objective are already underway, while further aims like reference genome, Balkans' gene flow and genomic comparison with other populations will be further discussed in the envisaged Summer school where this informal consortium will likely be further expanded. The small-scale student projects co-mentored by experts from both countries to tackle specific questions related to this topic are among planned activities in the Summer school. Establishment of a long-term collaborative mechanism between US and Serbian/Balkan data sciences offers the potential to develop unique and innovative projects that would greatly benefit the advancement of science in biology and medicine.

Track D Report: Digital Archaeology

In this section we present a summary of our assessment of the potential for collaborative research between US and Serbian archaeologists. This assessment is based on our experiences during the formal workshop program (https://nsfserbia.rs/), a site visit to Viminacium, a visit to the Institute of Archaeology Belgrade and to the newly refurbished Serbian National Museum, and informal conversations between the US and Balkan participants.

This track was co-chaired by F.P. McManamon (Arizona State Univ.), S. Blazevska (National Institute for Management of the Archeological Site of Stobi, Macedonia) and S. Golubovic (Institute of Archeology, Serbia). As a background for consideration of possible joint research activities related to data science in digital archaeology, eight talks have been delivered: Four presentations were by experts from the US, two by experts from the region (Macedonia) and two from Serbia. The invited talks have addressed the following topics:
"Archaeological Science, Archaeology of Science, and Tools for Closing the Gap between Practice and Ideals" (B. Marwick, Univ. Washington, USA)
"More Data and More Computation but not Necessarily Less Theory: Assessing the Status and Near-Future Directions of Archaeology" (T. Kohler, Washington State Univ., USA)
"Evidential Reasoning in Archaeological Science and the Need for Humanistic Approaches to Big Data" (C. Heitman, Univ. Nebraska-Lincoln, USA)
"Grand Challenges, Big Data, Fuzzy Data, and Digital Archaeology: integrating information about the past into the Planet Texas 2050 DataX platform" (A. Rabinowitz, Univ, Texas at Austin, USA)
"Digital archaeology at Stobi in Macedonia" (G. Pavlovski, Archeological Site Stobi, Macedonia)

"From Sketches to Digital Archaeology – Viminacium Research Project" (N. Mrdjic, Institute of Archeology, Serbia)

"Photogrammetry as a necessity: the Temple of Isis in Stobi" (D. Nikolovski, Archeological Site Stobi, Macedonia)

"Epigraphy databases – should Viminacium have one?" (I. Kosanovic, Institute of Archeology, Serbia)

The delivered talks illustrate certain related and complementary research interests of US and Serbia/Region experts which could be considered as a fruitful initial position for collaboration and joint projects.

Track D unique research opportunities in Serbia

An invited talk by S. Golubovic, Research Assoc. Prof. at the Institute of Archeology in Serbia, provided an overview of the archeology research eco-system in Serbia. Several archaeological sites of high international research significance are located in Serbia; we highlight just a few here. Serbia's prominent geographical position at the crossroads between southwest Asia and Western Europe makes it a critical area for the study of large-scale human adaptations and population movements from the Pleistocene onwards. This makes it attractive to US-based researchers who want to work on key questions of international significance relating to human evolution and major transitions and transformations in prehistory and the more recent past. The path of the Danube river through the country offers additional opportunities for the study of long-term environmental history, human-environment interactions, and economic and cultural connectivity.

Serbia is a karst-rich region with around 460 known caves, suggesting immense potential for discovering stratified hominid habitation sites. Pešturina Cave is a Palaeolithic site with a relatively well-dated sequence that is typical of many Serbian sites in having one of the latest transitions of Middle-to-Upper Palaeolithic in Europe, at around 39 ka. This indicates a persistence of Neanderthals in refugia, where they persisted later than in many other places in Europe. Given this pattern of late persistence, there is a high potential for sites in Serbia to provide unique insights into the demise of the Neanderthals, interactions between Neanderthals and modern human populations, and the process of replacement of Neanderthals by modern humans. To date, only brief excavation reports are available in English on Serbian Middle and Upper Palaeolithic sites. There is much scope for more detailed publication and specialist reporting to contribute to our understanding of major transformations in human evolution. Current work is being undertaken by experts from Belgrade University and the Institute of Archaeology with some long-term US and European collaborators.

Mesolithic and Early Neolithic settlements in Serbia are especially notable (or at least especially well investigated) in the Iron Gates region, a gorge on the Danube River containing two hydroelectric dams. The distinctive archaeological record of this area spans the period from the beginning of the Bølling–Allerød Interstadial to the Middle Holocene, ca. 12,700–5600 cal BC. Over 50 caves/rockshelters and open-air sites are known from this period and location, which together have produced large numbers of burials and architectural remains, and rich inventories of faunal material and portable artifacts, including personal ornaments, early stone sculptures, and other works of art. Over 450 burials, containing the remains of more than 700 individuals,

have been recorded from 15 sites, with five sites (Lepenski Vir, Hajdučka Vodenica, Padina, Schela Cladovei and Vlasac) accounting for over 95% of all burials known from the Iron Gates. These remains have been the subject of isotopic analysis to investigate mobility, marriage patterns, and diet. Many Serbian archaeologists are actively researching the Iron Gates Mesolithic and it has also been the subject of intense interest to English-speaking archaeologists (e.g., Hodder 1990). Serbian archaeologists have ongoing international research collaborations in this region, for example with the University of Cambridge, which houses several Serbian-trained research staff and fellows in the Department of Archaeology and has carried out a recent project focused on the microscopic analysis of dental calculus from individuals at the site of Vlasac (Cristiani et al. 2016).

One of the most significant sites of the Iron Gates Mesolithic is Lepenski Vir, which is important because it contains a rich burial and artefactual record of the Meso–Neolithic transition, and distinctive art traditions. This makes it an ideal location to investigate the timing and nature of contacts between Mesolithic foragers and Early Neolithic farmers. Although the site has been intensively studied, chronological questions remain. Notable current work on the Iron Gates is being led by the experts from Belgrade University on reconstructing and interpreting changes in demography and fertility across the Mesolithic–Neolithic transition. Given the rich burial record, there is great potential for new genomic and isotopic analyses to improve our understanding of the population structure, diversity, and mobility during this important transition period.

The first traces of the Neolithic anywhere in Europe are in the southern Balkans, and farming appears shortly after 6000 cal BC in the western Balkans (which includes Serbia). Sites of exceptional importance in the Neolithic include Starčevo and Vinča-Belo Brdo. These are notable because they are among the earliest distinctive cultural groups of some of the first farmers in Europe. Although their chronologies are still somewhat contested, they represent uniquely rich assemblages suitable for investigating the shift from hunting and gathering to full-time farming. There is scope here for investigating patterns of migration and subsistence change using genomic and isotopic methods. These sites, and later Eneolithic sites, contain important records of the emergence of social and wealth inequality, and some of them are beginning to be analyzed in ways that contribute to current dialogs on these same issues by US archaeologists. Future collaborations on this period could also examine the local Neolithic Demographic Transition in more detail and could include analyses of the relationship between population growth and inequality measured in various ways.

The earliest metallurgy in Eurasia has also been documented at the large Neolithic settlement site of Belovode, which is currently being investigated by a joint Serbian-British-German project. Early metallurgy has also been claimed for the site of Rudna Glava. This site has evidence of copper mining and was extensively investigated from 1960-1980 but remains enigmatic since lead isotopic analysis of a substantial sample of Eneolithic copper artifacts from the Balkans showed that none of them derived from Rudna Glava ores. This suggests some scope for future research to revise the chronology of metallurgical activities at Rudna Glava. These sites, and others like them in Serbia (e.g. Pločnik), represent an exceptional opportunity to investigate the technological and social contexts of the origins of metallurgy and a potentially fruitful area of collaboration between US and Serbia/West Balkan archaeologists.

Iron-Age sites in Serbia also provide significant possibilities for the examination of population mobility, culture contact, and the development of long-distance trading networks. The earlier part of this period has been the subject of relatively little study, but the emergence of historically-documented tribal groups and their interactions from the 5th to the 1st centuries BC have enjoyed more attention. Processes of collaboration, conflict, and state-formation can be observed in the relationships between Thracian, Illyrian, and Celtic cultural groups, which culminate in the establishment of the powerful kingdom of the Scordisci by the 3rd century BC. In the context of a vigorous exchange network focused on metals, amber, and wine, among other products, these local groups engaged in trade with both Greek communities to the south and Italian communities to the west. A Greek or mixed settlement was even established at Kale-Krsevica in the 5th or 4th century BC, in the area of an earlier Bronze and Iron Age stronghold. The archaeological remains from this period thus offer a rich tapestry of imported goods, local imitations, and elaborate metal luxury items of the sort that appear in the elite graves at the site of Trebenište in Macedonia (some of which are currently on display in the National Museum in Belgrade).

These Iron-Age groups formed part of a broader cultural zone that in antiquity stretched into the areas now occupied by Romania and Bulgaria, as well as Macedonia, and their interconnection can be seen in the incorporation of the Scordisci state into the Dacian kingdom expanded by Burebista in the 1st c. BC. It is even more apparent, however, in the results of a gradual process of Roman conquest, which brought Serbia, Romania, Bulgaria, and parts of Hungary under the control of the Roman Empire between the 1st c. BC and the 2nd c. AD. The Serbian and Hungarian territories were incorporated into the provinces of Dalmatia, Pannonia, and Upper Moesia, while the Bulgarian and Romanian territories spread across the provinces of Thrace, Dacia, and Lower Moesia. Perhaps the most dramatic archaeological reflections of this conquest in Serbia are the bridge across the Danube constructed by Trajan's engineers, some of the piers of which are still visible, and the Roman military road cut into the cliffs of the Iron Gates, now drowned by the hydroelectric dams.

The Danube formed a natural common frontier across these areas, especially as the Roman Empire abandoned the trans-Danubian province of Dacia in the 3rd century AD. As a result, these regions formed a koine of Roman provincial culture at the height of the Empire, when a series of Roman fortresses along the Danube became flourishing towns and provincial capitals. Histories of Serbia are quick to note that no fewer than 17 Roman emperors were born within its bounds between the 2nd and the 4th centuries AD, including Constantine the Great. The size and elaboration of its Roman sites reflect this importance. One of these sites, Felix Romuliana, is already on the UNESCO World Heritage list, and the series of fortresses spread out along the Danube are currently included in a multi-site, multi-nation World Heritage nomination that celebrates the organization of the Roman limes (frontier or border) in Europe.

This Roman legacy, and the long tradition of Roman archaeology in Serbia, provide another unique opportunity for US-Serbia research collaboration. Roman provincial archaeology is a vibrant subfield of Classical Archaeology in the US, and as Classical archaeologists become increasingly aware of the potential for environmental archaeology to illuminate history, the combination of frontier culture and long-term environmental research is particularly attractive. For historical reasons, US-based archaeologists are only now beginning to establish long-term collaborations in the former Republic of Yugoslavia and in Romania, so the current environment

is especially favorable for research partnerships. This potential is exemplified by the site of Viminacium, which attendees visited on the first day of the workshop. This site encompasses both a well-preserved Roman military camp and a thriving town of perhaps 50,000 people that served as the provincial capital of Upper Moesia. Its excavation was spurred by the construction of a major electrical power-plant and an adjacent open-air coal mine; archaeologists have been working furiously to excavate and document both the site and its associated cemetery, one of the largest provincial cemeteries of this period. To date, more than 14,000 graves have been excavated and documented, and the remains of approximately 1000 individuals are currently stored for further study.

A new campaign of excavation in the cemetery will begin this fall, in advance of new construction at the power plant. The excavation team is multidisciplinary and well-trained in both traditional and digital methods of documentation. Terrestrial and drone-based photogrammetry are routinely carried out, together with various methods of geophysical prospection, and the project has the benefit of a new research and outreach center that includes lab space for specialized analysis, a library, meeting rooms, and lodgings for students and staff. A team of specialists work on the ceramic record, and the staff includes a physical anthropologist and an archaeozoologist, both highly qualified. In addition to the extensive ceramic material and metal and bone artifacts, past excavations have recovered very large quantities of animal bones, with a particularly interesting deposit from the Roman amphitheater that includes the bones of bears, lions, and leopards. Parasitological studies are underway on soil from the drains of a public latrine in a large Roman bath complex, and composition analysis has been carried out on architectural mortars. The project is thus involved in a number of scientific investigations, both quantitative and qualitative, that present opportunities for collaboration or data-sharing with US partners. A collaboration focused on the human population from demographic, epidemiological/pathological, genetic, and isotopic perspectives is particularly attractive, as is a collaboration that might apply new analytic techniques to questions of human-environment interactions and animal exploitation (especially fish, which have so far not been studied in any detail).

The presence of these unique sites in Serbia means that collaborations between US and Serbian archaeologists will give US researchers excellent access to key resources for investigating archaeological problems of international significance, from the Middle Palaeolithic through the Roman period and beyond. We emphasize here the importance of the archaeological record in Serbia for understanding past human mobility, especially with respect to large-scale population movements and their short- and long-term effects.

Track D strengths in the local research community

Archaeologists in Serbia have a deep knowledge and strong appreciation of the unique archaeological resources of their country. Up to now there have been several generations of local archaeologists, and there is strong institutional knowledge about the work of prior generations of researchers. They are well connected to archaeologists elsewhere in the Balkans and throughout Europe. They have several locally-produced English-language archaeology journals that are peer-reviewed, open access and available online. There are strong collegial relationships between archaeologists at the Institute of Archaeology (Ministry of Education, Science and Technology),

those based in the Faculty of Philosophy at the University of Belgrade (Ministry of Education, Science and Technology), those at the Museum (Ministry of Education, Science and Technology), and those employed by the Institute for Protection of Cultural Heritage (Ministry of Culture). Undergraduate and graduate students appear to have many opportunities to gain field and laboratory experience.

The local archaeological community is highly skilled in complex technologies for field data collection, such as remote sensing and photogrammetry for 3D modelling. Amongst the archaeologists working at Viminacium in particular there is a strong interest to further develop digital archaeology capabilities, such as databases to organize their massive quantities of rescue excavation data. Throughout the spectrum of archaeologists we met in Serbia there is an awareness of the need to develop more problem-oriented research programs, to engage with anthropological frameworks, to employ qualitative and quantitative models to understand archaeological data, to expand professional publication opportunities (critical to career advancement), and to expand training opportunities for students.

Track D risks, limitations and weaknesses

Retaining highly-capable students in careers in archaeology is a challenge for the Serbian archaeological community. Although graduate students can find employment at the Institute or on projects such as rescue excavations at Viminacium, there is no funding scheme for post-doctoral research to support recently graduated students who do not yet have a university position. New academic positions at the University of Belgrade are very rare. Because cultural resource management activities are only permitted to be carried out by the Museums, the Institutes or the University, there is no possibility for private companies to be formed and generate employment for recent graduates and early career researchers. These factors limit the viability of a career in archaeology for Serbian students, and may make it hard for US researchers to find junior collaborators who have the time to take on new projects and make meaningful contributions to collaborative research programs. However, we are aware of plans to substantially expand the institute by hiring more archaeologists and work towards official affiliation with the Serbian Academy of Arts and Sciences. There also appear to be limitations on monograph publication opportunities for archaeologists within the institute, which creates a double bind: career advancement requires publication but annual support for producing publications is limited.

A limitation of the way activities are organized at Viminacium is that because of the pressure of the nearby mining activity, excavations are extensive and conducted quickly. This leaves archaeologists with little time to contextualize the finds and prepare publications for an international audience. This means that collaborative research on that site would need to invest substantial time in digitizing, cleaning and organizing data before more research-oriented activities can be carried out.

Track D options for future collaboration

The extensive range of unique archaeological resources in Serbia offer many possibilities for collaboration between US and Serbian archaeologists. At the end of our workshop, individual

members of the US delegation could easily conceive of small-scale collaborations with Serbian archaeologists we have met during our short visit. However, to further assess what topics are most ready for long-term or large-scale collaborative research that might broaden the community of collaborators from both countries, we propose that the next step should be a research-focused summer school where a small number of US faculty and students combine with Serbian faculty and students in a multi-site summer school in Serbia.

One option for the structure of this summer school is for the Serbian archaeologists to teach photogrammetry, 3D methods, and cultural transmission analysis methods, and the US faculty to teach anthropological theory, quantitative modelling, and digital humanities approaches to data analysis. This plan would ensure an equal contribution from US and Serbian archaeologists, according to the skills the Serbian archaeologists have, and would like to acquire from the US archaeologists.

The summer school could be organized around two of Serbia's unique archaeological resources, for example Viminacium, using data collected by previous work at these sites. The goal would be for students and faculty to work in groups to initiate and begin small research projects that might be suitable for publication, or as pilot work to support a funding application.

We have identified the NSF program International Research Experiences for Students (IRES) as a potential source of support for this summer school.

This summer school could also include a substantial focus on the digitization of legacy data and the creation of interoperable databases, which is a major desideratum on the part of the Viminacium team. It might also help to catalyze and support more specific international research projects: both the physical anthropologist and the zooarchaeologist associated with the project are interested in international collaboration in order to deal with the study of large backlogs of material and to expand the range of scientific techniques for analysis employed at the site.

One member of the US group (Rabinowitz) has recently initiated a project focused on the genetic analysis of both human and oral microbiome (dental calculus) using DNA from a Roman cemetery of approximately the same date at Histria in Romania; this project also envisions a program of isotopic analysis to understand both population mobility (Sr/O) and dietary variation (C/N) in the context of social status and equity within this community. A significant zooarchaeological component is also anticipated for this project, which could form the basis for collaborative work on comparative collections in this area as well (especially with respect to the riverine ecosystem of the Danube at various points along its course). Given the geographic and chronological proximity of the Serbian and Romanian stretches of the Roman limes, this project might offer a starting-point for a broader international effort to understand demographics, mobility, diet, and resource distribution along the Roman border.