Interview - H. Akin Ünver

https://www.e-ir.info/2020/05/21/interview-h-akin-unver/

E-INTERNATIONAL RELATIONS, MAY 21 2020

Akin Ünver has a PhD in Government from the University of Essex and is currently an Associate Professor of International Relations at Kadir Has University, Istanbul. He is the Resident Fellow of the Cyber Research Program at the Centre for Economic and Foreign Policy Research (EDAM), a Research Associate at the Center for Technology and Global Affairs, Oxford University, and a Research Fellow at the Alan Turing Institute, London. His areas of specialization are conflict research, computational methods and digital crisis communication. He is author of *Defining Turkey's Kurdish Question: Discourse and Politics Since 1990*, and the recently published *Internet, Social Media and Conflict Studies: Can Greater Interdisciplinarity Solve the Field's Analytical Deadlocks?* and *Computational IR: What Coding, Programming and Internet Research Can Do for the Discipline?.*

Where do you see the most exciting research/debates happening in your field?

I think the volume and quality of the scholarship in my field (broadly, as international relations, narrowly, as conflict studies) are consistently getting more impressive. Both fields are enjoying a 'golden age' of sorts, in terms of rigor and creativity, which also renders both fields very challenging to follow because such rapid progression of any field makes reading alone a full-time job.

At the sub-national level, namely, conflict and civil war literature, communicative dimensions of armed dissent, or 'rebel diplomacy' provides an exciting new opportunity. So far, the conflict studies field has focused extensively on how violence emerges, diffuses and terminates. Yet, parties to a conflict also engage in non-violent bargaining, communication and propaganda efforts as a part of their strategic interactions. Groups bargain with each other to merge, break-away, or peacefully contest territory, as well as with external state sponsors to acquire more resources and legitimacy. The dynamics of such efforts have been either neglected by the traditional literature mainly due to data unavailability issues or got marginalized because of the past focus on violent behavior as the main focus of inquiry.

However, newer forms of data (such as the Internet and social media), as well as broader opportunities to do fieldwork in conflict settings (although these opportunities are largely confined to US-based scholars) have changed this prioritization. Now, we have ample research demonstrating how rebel communications are among the vital performative acts in a civil war setting, and often generate a broader range of 'real-world' empirical implications compared to fighting. Since rebel groups don't have a functioning diplomatic service, they generally conduct diplomacy over social media and digital communication channels, providing us with both new data types and previously neglected conflict dynamics.

At the interstate level, I think the most exciting current debate is on how nations invent, share and adapt to new technologies. In the last two decades, the progress of scientific advancement has been so rapid that it has begun to impact great power dynamics and how alliances or enmities form in relation to it. The debates on the rise of China, erosion of US-based unipolarity, erosion of WW2 institutions all derive significantly from this changing technological competition, thereby rendering technology a central independent variable IR scholars employ to discuss system-level shifts.

The most promising and cutting-edge of these technology debates regards artificial intelligence (AI). AI is a comprehensive concept whose militarization is challenging to measure, and it is even harder to accurately assess what drives the diffusion of 'weapons-grade' AI technology. There are some impressive conceptualization

studies on how to define 'weaponized AI', how to think about measuring its diffusion and identify the main variables that explain how allies share or not share AI-based military technology. This debate is similar to the military technological adaptation debate in the early 20th century and during the inter-war period (1919-38), which is why it is a debate that is worth closely following.

How has the way you understand the world changed over time, and what (or who) prompted the most significant shifts in your thinking?

It is hard to pinpoint specific events, but broadly speaking, teaching in four different countries – the US, UK, Turkey and Switzerland – and getting exposed to the political worldviews of these four cultures changed how I see the world and interpret world events. I'd suggest all emerging IR scholars (master's, PhD-level or early-career) lecture in as many countries as possible (not short courses or seminars, but at least one semester-long commitments) and spend several years outside the country where they got their PhDs and country they were born in or raised. This may be risky over the short-term because relocating internationally soon after getting an MA or PhD will inevitably impact early publication quantity and thus, the CV. However, over the medium-to-long term, extended multi-country living and teaching experiences will make sure that an IR scholar's work is genuinely IR and not just a version of their own country's foreign policy analysis.

Prolonged teaching in different cultures not only enables a scholar to learn more about other countries but also pits the scholar's assumptions, worldview and axioms to a more 'international' challenge. Can a scholar's assumptions, built in their home country or country where they got their PhDs, work outside of those countries? Or do these assumptions suddenly become less relevant or less engaging to students in different countries? This is one of the best ways to calibrate a scholar's intellectual prowess and make sure they remain relevant internationally. This is true for all disciplines, but far more important for our discipline – International Relations – which, by definition, has to be 'international'.

Computational Social Science (CSS) is emerging as a research method in traditional social science disciplines, such as Sociology, Anthropology, Political Science and Economics. Could you explain CSS and what accounts for its emergence in these disciplines in the last couple of years?

Defining CSS is difficult because as a new and developing field, its researchers try to mold and stretch the definition in a way that fits into their research agenda. Another part of this definitional problem owes to the fact that both the data types and the tools of analyzing them are rapidly changing in tandem. This causes the confines of CSS to expand continually. There are also competing terms such as 'social data science', or 'new media studies' that often overlap with the objectives of CSS. A straightforward definition of CSS is that it is the use of computational tools to study social science. The problem is that Excel is also a computational tool, and running a simple crosstab calculation to study a social science phenomenon isn't precisely CSS. Same goes for conducting statistical analysis of social events using Stata. Stata *is* a computational tool and *does* social science, but this practice wouldn't constitute computational social science.

I think the best way to break through this confusion is to focus on the data type. Most foundational studies of CSS use fluid digital human data – be it social media text/image/audio/video content, mobile phone data, or real-time log-in/log-off, check-in/check-out data to learn more about why groups of people do the things that they do. So, rather than focusing on which tools we use on a computer, the focus of the CSS has to be on the type of data that we use: user-generated, digital, fluid and open-source. Using mobile phone metadata to study poverty in Nigeria? That is CSS. Harvesting geo-coded user-generated images from Flickr to predict protest onset? That is CSS too. Using fancy machine learning models on large World Bank inequality indicator datasets to understand rebellion? That is not really CSS because the data is static and not user-generated.

So, my personal interpretation is that it is the data type that renders CSS genuinely unique and different from mainstream quantitative research methods. Users themselves create computational data – their tweets, their location information, their texts, videos, images are all user-generated and can freely be found online. Some of this is intentional data (i.e. you know that your data is public) and some of them are intended to be private, but

they either leak or get scraped without user consent. That's why computational social science ethics is and should continue to be a critical pre-requisite discussion. Contrast this to the regular statistical or other quantitative work that is built on static datasets that aren't user-generated (i.e. researcher or interviewer codes them), and we have a clear understanding of how CSS is different from quantitative methods. That's why the rise of CSS owes to the mass spread of social media, smartphones, Internet connectivity and penetration. As more and more people began using these new tools that bestow them to create their own data and share it with the world in real-time, CSS simultaneously became an essential and exciting new area of research, as well as a research method.

Which approaches or tools of CSS are the most promising? How could they contribute to the advancement of International Relations as a discipline, theoretically and methodologically?

It is hard for me to pick because I teach all major approaches: social network analysis, simulations, text analysis, mapping and so on. Same goes for IR as a discipline; I can't think of any CSS approach that can't be for the benefit of an IR scholar. I have reviewed these approaches in-depth in terms of how they best related to IR in a recent publication titled 'Computational International Relations'. Also, please see my 'Turing lecture' on how to combine computational methods with the research focus of IR.

Text mining and analysis tools for example, can work very well with researchers focusing on constructivist and post-positivist IR approaches. Geospatial tools work best with IR students of borders, space and geography. Network analysis can be employed by those studying fluid complex relationships such as terror networks, advisory circles of leaders or ideological proximities between groups. The most important contribution of CSS to any field, I think, is that it uses more granular self-generated human data. While the popularization of the Internet and social media has provided people with a broader range of self-expression opportunities, they have also bestowed researchers with an unprecedented volume and detail of self-generated human data. Computational tools will definitely improve and change over the next 10 years and so will the interest in user-generated fluid data forms. The potentials of IR-CSS cooperation are so vast, that it is hard to provide even a quick snapshot in a paragraph.

What drew you to CSS? How have you applied methods of CSS to your own research?

The Syrian Civil War was my turning point. Having studied foreign policy discourses and militant group discourses previously, I got frustrated with how little we know about the realities on the ground in Syria. By 2013, I started to explore social media data, blogs and websites created by the militants and Facebook pages related to these groups. Gradually, these earlier explorations turned into exhaustive research sessions that included locating, interpreting and dissecting how those groups use the Internet and social media both as a diplomatic bargaining tool and also a governance/administrative method. I began to learn web scraping and text analysis on my own, but I was exposed to a far broader range of CSS methods during my fellowship at the Oxford Internet Institute.

My first CSS project was 'Militant Selfies' project where I scraped selfies containing text that corresponds to a pre-arranged corpus of words most frequently associated with ISIS violence in Syria. Using these geotagged selfies, I was able to build one of the most detailed conflict maps of Syria. That project won an award from Oxford University's Cybersecurity Program in 2017 and was later featured on the BBC. Then, I partnered with Dr Mirco Musolesi from the University College London to start a project titled 'Computational Agent-based Models of Civil Wars'. In this project, we scraped all social media posts of militant groups in Syria and still build a 'tit-for-tat' behavioral model that explains which violent acts force the other side to retaliate and under what conditions do certain violent acts don't get retaliated. This project also won an award and funding from the Alan Turing Institute in London.

What are some of the current limitations of contemporary Cybersecurity research? Which CSS tools could we use to overcome them?

There are two chronic problems with cybersecurity research: data availability and attribution. What constitutes a 'cyber-attack'? DDoS (Distributed Denial of Service), MitM (man-in-the-middle), phishing and spear-phishing

types, SQL (structured query language) injections, cross-site scripting (XSS) attacks are some of the best-known examples. The problem is, if a cyber-attack is well-strategized, it becomes very difficult to spot, because most 'sophisticated' cyber-attacks are very hard to notice, or it takes weeks or months before they are spotted. Most successful hacking operations are never discovered; the attacker gains access into the target system and stays in without being spotted for months! Even when spotted, the second chronic problem – attribution – comes into play. Cyber-attacks are easy to mask and mislead, rendering it very difficult to identify who the attacker is.

This forces the researcher to either focus on a smaller pool of empirical evidence of cyber-attacks that incurred measurable damage to physical infrastructure or find out ways to partner with cybersecurity firms to gather corporate research data. The first option gives us very limited data to work with because cyber-attacks that have physically observable implications (like explosions or service disruptions) constitute a very small portion of total cyber-attacks that happen every second, worldwide. The second option requires significant privilege – either rich research funding to buy cyber 'event' data from firms or special connections that will get the researcher this data for free. Even in the latter case, the data firms release is the data that they want to be released, not necessarily the sample required by the researchers. In either case, the second option gives us a limited-access data trove that is not replicable by other researchers and hence making it very hard to externally verify the validity of the dataset.

There are only a handful of studies that have managed to solve these problems but only for narrowly-defined empirical case studies. CSS doesn't solve the problems of cybersecurity research but provides a working pathway for cybersecurity scholarship to emerge out of its data availability problems. CSS works with publicly-available data, and thus, CSS studies are easier to replicate and easier to verify in terms of data validity. Second, CSS has provided models for industry-academia partnership to supply researchers ample data to work with, whereas such a model has thus been unforthcoming in the cybersecurity field. Getting granular cyber 'event' data is still a matter of privilege, so its results are hard to generalize and replicate, rendering it overall problematic as a scientific field. Further research is needed in cybersecurity to clarify and refine how to measure and identify cybersecurity-related events, how to account for 'unknown' or 'unknowable' (concealed) cyber incidents and how to think about the attribution problem in cybersecurity research. Should we try to solve it, or can we generate empirical knowledge without attribution?

How can we address the tensions between the benefits of computational methods and technology and the vulnerabilities that data science and digital media platforms brings to democracy?

This is an evolutionary and iterative process. Some of the vulnerabilities of platforms, such as human and bot-oriented disinformation, algorithmic bias and privacy were identified by researchers in the first place. Although some researchers choose to exploit these vulnerabilities for research purposes (to acquire larger datasets, more invasive personal data), some others work with these platforms to address these vulnerabilities and to better protect users against exploits. Some researchers also establish public awareness platforms to bring these issues to the mainstream social debate and raise greater awareness around them. As greater public pressure forces platforms to improve their privacy practices, these platforms then turn back to the researchers to find ways to address those concerns.

Some of these problems will be solved along the way as societies become more aware of them. But if I had to point to two mechanisms, I'd point to a greater public debate on these issues, as well as more robust platform-researcher partnerships that can continue to solve these problems as they arise. The main thrust of our efforts has to be building greater data literacy skills among society, especially young people. Most people don't really know how their tweets, Instagram photos or Facebook albums are harvested by the government and platforms for surveillance or monetization purposes, and thus, don't understand why they need to assume control over what they share online. Once they have a good understanding of how their data are used, they become strong advocates of privacy and raise awareness among their friends that end up successfully pressuring the platforms into improving their problematic data retention and harvesting methods. As far as the agency-structure problem is concerned, I believe online threats against democracy has to be solved at the agency-level. This is also valid for CSS researchers – data ethics and privacy must be taught at the beginning of CSS programs, which is what we

do at the Summer Institute in Computational Social Science (SICSS).

You have been studying crisis communication and decision-making processes**. How have social media and the digital space been used to protect the reputation of authoritarian and semi-authoritarian leaders during periods of crisis and high uncertainty?**

This is a very rich and rapidly developing field. Following the Arab Spring and Occupy movements, governments worldwide have begun to see the Internet as another battlefront of global influence. Authoritarian countries have gone a step further and viewed the Internet as a potential battleground with their own citizens. In fact, authoritarian control over public communications is an old research field, but the Internet and social media introduce new dynamics to existing theories. A major control area is at the infrastructure layer; authoritarian governments generally first sever physical connection (Internet shut-downs, bandwidth throttling), or disable the entire infrastructure of connectivity. These measures happen during perceived or actual attempts towards regime change, or crises that impact government stability.

The second control area is at the connection-level; namely censorship of accounts, keywords, or groups of people, or state-sponsored DDoS attacks that take down opposition websites and blogs. These measures are taken during crises that are serious, but not government-destabilizing. Third, authoritarian governments co-opt digital connection practices and try to influence public opinion through the Internet and social media. These measures include 'troll armies', bots, or state-sponsored disinformation efforts that seek to mislead, distract and de-mobilize masses. At the fourth-level authoritarian governments use digital public diplomacy tools (official accounts) to convey policy messages and manage audience costs. These levels represent the government's perception of the intensity of a crisis, so if the perceived crisis is greater, the government uses more radical measures to limit communication. The problem is, in recent years, countries labeled as 'democracies' have also begun to rely on second and third-degree Internet meddling policies. This blurs out any clear difference between authoritarian and democratic governments in terms of how they respond to online dissent during emergencies.

You are currently building an Artificial Intelligence Weaponization Dataset (AIWD). What is the AIWD and what can we expect it to cover?

This is a new project funded by the Turkish Academy of Sciences, so I can't get too much into the detail at this point. Broadly speaking, there are two definitional problems: a) what constitutes 'weaponization' of technological advancement in the field of artificial intelligence, b) how do we measure these 'weaponization' processes when their progress is extremely fast? The first part of the project seeks to build a dictionary and event dataset that will allow us to speak the same language when it comes to the study of how countries use AI-based technologies for military purposes. This is important because AI research has to escape the traps that plagued the cybersecurity scholarship for so many years; there has to be a common lexicon, vocabulary and event pool from which we can build new knowledge and form a sub-field of science.

The second part of the project uses machine-learning and live-scraping techniques to auto-detect and log every new AI-based military technology progress made by the countries into the Artificial Intelligence Weaponization Dataset (AIWD). This will be a live dataset that surveys the entire Internet like a radar, searching for terms, phrases and word combinations that we have identified in the first part of this project. The main purpose of the project is to provide an extremely granular and 'live' event dataset for the use of scholars that explore the uses of A.I. within the military domain, as well as researchers that study how technology transfers and military technology adaptations take place around the world.

What is the most important advice you could give to young scholars of International Relations?

First, if you can, spend as much time abroad as possible, and in several different countries. This will be uncomfortable financially, friendships and adaptability-wise. Yet the study of IR necessitates getting out of your own identity comfort zone; both in terms of your ethno-nationalist and religious/sectarian background. The entire purpose of our field is to be 'international'. It is hard to be international and think in true IR terms if you do a PhD,

MA, BA all in the same country you were born and raised in. I realize that this will be hard for scholars that come from less privileged socio-economic backgrounds, that's why I began this paragraph with 'if you can'. But even if you can barely make it financially, do it. Living in multiple countries, in a tolerable degree of hardship is the best IR training you can ever get.

Second, spend around 40 minutes each day watching news from at least 3 different countries in different continents. The US, South Africa and Russia, or Brazil, China, Egypt or any other combination. Most countries have English-language news channels, so daily exposure to the news agendas and priority issues of different countries will get you outside of your own country's news bubble. Things that are important for your own country are generally not that important for other countries. What your country thinks as 'right' is often not so justified for other countries. This provides a much-needed perspective and vision for any young IR scholar.

Third, deepen your interest and enthusiasm in comparative sociology and history. IR can be a lonely discipline if it doesn't import concepts and perspectives from other disciplines. Sociology and history cover a very large span of fundamental IR-related topics, enabling a scholar to derive from two major fields of social science at once. This will improve your assessment of world events, system-level problems and country-specific political processes better than scholars that don't have an interest in both fields.