Idea 1 -- Quantifying the reach of Russian disinformation on COVID into post-Soviet countries:

[PEOPLE HERE LOVE TO TALK ABOUT RUSSIAN DISINFORMATION, BUT MUCH OF THE PROBLEM IS LOCAL. SO TYING DISINFORMATION TO RUSSIAN LANGUAGE BOTS MAKES ME NERVOUS; LOOKING AT THE ROLE OF BOTS MORE GENERALLY IN DISCOURSE IS VERY INTERESTING. BUT OVERALL, I AM MUCH MORE INCLINED TO YOUR IDEAS BELOW STARTING AT NUMBER 3] – **DT: so much disinfo is in Russian domestically, and so much native-lang disinfo in from Russia. Doesn’t mean don’t do it… but go over the points.**

* Some preliminary questions to answer before doing an investigation:
  + What is the reach of Twitter/ other social media in these countries? Who is on these platforms? DIFFERS BY COUNTRY. TWITTER NOT HUGE **not as big as FB and VKontakte, FB is the one universal**. TELEGRAM QUICKLY BECOMING MOST IMPORTANT FOR SHARING INFORMATION IN CENTRAL ASIA, **used by government too, teachers using it to teach. Look at some surveys that are out there, IRI has done some—plus their rival NDI.**
  + What other platforms, if not social media platforms, should we consider when assessing how Central Asians consume information?
  + What is the history and aims of Russian disinformation on these platforms?
* Proposed project outline, assuming an analysis of a platform like Twitter:
  + First, identify Russian-language Twitter tags associated with Covid misinformation. (In English, this would be hashtags like #fakevirus, #covidfake, etc.). This would be a qualitative exercise. Then, we could pull large samples of tweets from the Twitter API that use these terms.
  + Next, identify content that is thematically similar to these tweets from a random sample of Russian-language tweets, for a more complete dataset.
  + After that, build a network of the twitter users who engage in/ follow/ reply to this content. (I know of algorithms that can do this) I LIKE THE IDEA OF DOING THIS BUT STARTING IN A SMALL, RUSSIAN-SPEAKING POPULATION, LIKE BISHKEK
  + See how many of them are bots, based on habits of tweeting (again, bots have telltale habits). Doesn’t mean they are from Russia (and tying twitter users to a place is hard)… but from where else would they be? EVERYWHERE
  + In which countries do these bots predominate as a share of the discourse? Does it have anything to do with the country’s “alignment?” E.g. is there a difference between their prevalence in Georgia and Armenia, one of which is pro-Russia and the other more anti, but both with somewhat-free information ecosystems? IN BOTH THE PROBLEM SEEMS TO BE VERY HOMEGROWN

Idea 2 -- What is really going on with Russia’s COVID death toll? How is it possible that it’s so low? Can we tell where the under-reporting bias is coming from? Potentially more of a deep-reporting story than a data story. A NUMBER OF JOURNALISTS IN MOSCOW ARE TRYING TO DO THIS AND I DON’T FEEL LIKE WE CAN COMPETE. WE’RE NOT A PLACE PEOPLE COME TO FOR BREAKING NEWS ON RUSSIA. CAN YOU DO THIS FOR KAZAKHSTAN? THEY SEEM TO BE DOING THE SAME THING AS IN RUSSIA, I.E. NOT COUNTING COVID DEATHS WHERE THE PATIENT HAD UNDERLYING ISSUES. **Look at excess deaths data—spend some time. If it gets far, then would introduce to guy in Almaty.**

* As others have said: the numbers are likely to be downplayed for all sorts of reasons (e.g. here: <https://abcnews.go.com/International/data-suggests-russias-coroanvirus-deaths-higher-reported/story?id=70683286>)
* Do we have contacts in Russia who can give us more specific excess mortality numbers than what is publicly available? Any qualitative insight on why the mortality numbers may not be accurate?

Idea 3 – 30 years after the end of USSR, what is the status of the Russian language in the near abroad? Potentially, an audio series:

* Preliminary question to answer: how to reduce the sprawl/ specify the scope of this question we’re answering? Is there a specific domain we should focus on, like education or **government communication**? Or maybe **music—maybe Youtube (e.g. Kyrgyz rap with Russian subtitles)** or cultural output? VERY COOL. IT IS SO VAST, THOUGH, I WOULD CONSIDER LOOKING AT ONE COUNTRY TO START
* Potential data sources:
  + Scraped government documents. Simply add up number of documents by language over time. KAZAKHSTAN PROBABLY HAS THE BEST ONLINE RECORDS IN CENTRAL ASIA. **Likely to go back over time. Maybe first pub in Russian and then first in Kazakh.**
  + Ditto for press releases, official announcements, parliamentary debates
  + Ditto for number of schools offering instruction by language (where might this come from?)
  + Social media speech. Difficult to assess over time, though.
* Are there specific types/ domains of communication that are more likely to be undertaken in Russian? ANYTHING TECHNICAL. MOST URBAN AREAS. NGOS
  + To answer this, for the documents we want to use, we can use cross-linguistic topic modeling (using a universal sentence encoder—I can explain in person) to cluster documents across languages by theme or subject matter, over time.
* Are there specific types of PEOPLE who speak one or the other language, and in what context? URBAN TYPES ARE MORE LIKELY TO SPEAK RUSSIAN, AS ARE ETHNIC MINORITIES EVEN FROM REMOTE AREAS
  + Would need to do a lot of legwork on biographies, potentially feasible for politicians in some countries.
* Domain knowledge could be shaped by interviews with country-level experts, language activists, educators, etc.—on both sides. Would want to hear from linguistic Russophobes as well as Russophiles in various countries. I THINK YOU WOULD FIND A LOT OF PEOPLE WILLING TO TALK ABOUT THIS IN KAZ, KYRGYZSTAN, UZBEKISTAN, AZERBAIJAN (WHICH HAS SEEN RUSSIAN DROPPED AS LANGUAGE OF URBAN ELITE VERY QUICKLY)
* Depending on the scope/ findings, package this as a **radio story**. Ripe for sound bites in different languages, and for personal stories about linguistic heritage from activists and researchers. SURE
  + One episode per country
  + Accompanied by a digital guide that includes visualizations—e.g. line charts showing the quantifiable decline of Russian in different spheres.
* A coda: what are the effects of Covid on this? Have governments preferred one language or another to communicate with their citizens? GOOD QUESTION. I’VE SEEN A GOOD EFFORT TO TRY TO DO THINGS IN MULTIPLE LANGUAGES IN CENTRAL ASIA, LESS SO IN THE CAUCASUS, WHERE ENGLISH IS BECOMING THE GO-TO SECOND LANGUAGE

Idea 4 – Similar to the above, but with a focus on Kazakhstan (which you singled out) and with a deep dive on the links between quantified language use and political developments. Could be print or audio.

* Same analytical techniques as above, but peg to important political developments
  + Nazarbayev’s push to move government services to Kazakh in the 90s
  + The more recent push in 2018, which apparently had to be walked back. <https://today.ku.edu/2018/01/25/kazakhstan-seeks-use-language-tool-establishing-independence-scholar-says>
  + Etc.
* As an additional dimension of analyzing social media data: you can train an algorithm to recognize an ethnic Slavic vs. ethnic Turkic name, I’m sure. To the extent that people use their real names on Twitter (probably a terrible assumption), what share of ethnic Kazakhs are using Russian online? THEY’RE PROBABLY MORE LIKELY TO USE THEIR REAL NAMES ON FACEBOOK, WHICH IS POPULAR IN KAZAKHSTAN
* A historical reference I’ve relied on here: <http://mason.gmu.edu/~cenasia/Fierman.pdf>
* Is digitization in government, as accelerated by COVID, tipping the scales of official use in one direction or another? <https://astanatimes.com/2020/05/lockdown-measures-speed-up-nur-sultans-smart-city-digitization-plans/>
* **If you narrow the focus to parliaments, just cookie-cutter the analysis from Ukraine!**
* **Kyrgyzstan—very sensitive for a time, there was some bloodletting. But has calmed down a lot. Definitely been a trend towards Kyrgyz in public life.**
* **Scrape individual deputies.**

Idea 5 -- Food security in the region and its links to climate change + COVID: I LOVE THIS AND THINK THERE IS A TON TO EXPLORE HERE AND IT IS ONLY GROWING WORSE WITH COVID

* Background is that food cultivation in Soviet Union (and particularly Central Asia) was often guided by production goals and done in unsustainable ways. What harvest have these countries been reaping over the past 30 years? Have the foodways of people had to evolve? Are people eating better or worse now than they were 10, 20, 30 years ago?
* Data visualizations—maps
  + Areas under cultivation (by province)
  + Shaded by productivity by crop
  + Animate (or make user-configurable) by time
  + Extend to future per expectations of climate change (consult climate scientists about this—what is going to grow in Uzbekistan in 30 years?)
* Are there data on hunger, food insecurity, poverty we can use? I know nothing about this domain. NOT GREAT. FAO HAS SOME, ESPECIALLY FOR TAJIKISTAN AND KYRGYZSTAN, MAYBE ARMENIA.
* Inspired by this: <http://www.ponarseurasia.org/memo/food-security-central-asia-public-policy-challenge>
* Also this: <http://www.fao.org/news/story/en/item/1255716/icode/>
* **USDA has a lot of data apparently—market year, July 1-June 30. Have a press officer who replied quickly.**
  + **Georgia imports a lot of food from Russia**
  + **Ukraine, Russia, Kazakhstan are major grain exporters**
  + **Some people have linked Arab spring to bad harvest in Russia**
  + **Weekly report on food prices from UN on Tajikistan**
  + **Tajikistan, Kyrgyzstan, Armenia, Georgia—all recipients of lots of data**

Idea 6 – A more culture-oriented piece on food and eating in Central Asia. Could also be an **audio series** by country, if we want to adopt a broad scope. THIS WOULD CERTAINLY BE INTERESTING, THOUGH I FEAR DATA WILL BE HARD TO COME BY. **Could just be a reported piece.**

* Would explore: as some Eurasian countries (like Kazakhstan and Azerbaijan) have gotten richer, what are the bourgeoisie eating? Has national cuisine become more “international,” turning towards Muscovite favorites like pizza and sushi? Have culinary influences from the Islamic world strengthened since independence? How has the restaurant scene evolved? What are the trends in alcohol consumption?
* Data, if any, could consist of restaurant guides, menus/ prices, import/ export figures, if we could dig them up.
* Could link to agricultural data on food prices: how linked are trends in food to availability of raw inputs?
* “Experts” could be agricultural or regional economists, local food writers, as well as experts on lifestyle/ foodways in the countries of interest.
* Looking ahead: effects of the virus? Likely to be pure speculation.

Other fantasy ideas that seem difficult for lack of data:

* An investigation of the ingredients and methods used in Soviet and post-Soviet cookery. Would need a fairly comprehensive corpus of Russian-language cookbooks, food guides, and possibly menus, etc.
* An investigation of the “entitities” (people, places, things) or “cast of characters” that appears in Soviet and post-Soviet historiography over time. Would need a corpus of history textbooks.
  + A way to quantitatively track the rise and fall, and rise again, of characters like Stalin, or the Czars, or what have you. YES, WE’VE LONG TALKED ABOUT THIS AND I THINK A SPECIFIC LOOK AT HOW STALIN IS REPRESENTED IN REGIONAL HISTORY CURRICULA WOULD BE EXCELLENT. GETTING TEXTBOOKS, HOWEVER, IS A LITTLE TRICKY
  + With that corpus, a little more advanced NLP could tell you about the *context* that character is discussed in. Is Stalin being associated with terms like “repression,” or “industrial production,” for example? VERY COOL
* **Uzbekistan—what is level of adoption of Cyrillic?**
  + **Kyrgyzstan—hasn’t gone there**

Notes 6/29:

* Personal news
  + Accepted voter access fellowship—starts late this week or early next
  + Would still like to do stories for you, and have been approved to do so
  + Have not yet asked Justin about payment, and am open to how we str
* Present ideas
  + Kazakh language activists
  + Kazakh Twitter bot presence in COVID discussion (vs. general)
  + Food security—Tajikistan, Uzbekistan, Kyrgyzstan in particular.
* Plenty of Kazakhs in Japan—friend of mine now at Harvard did Grad work in Japan.
  + Kazakh language and domains
  + Keep digging there
  + Impression: intelligentsia and all Russian-speakers.
  + And sort of more nationalists, often not from big cities (originally)
  + Groups would butt heads a bit
  + Gets a feeling that more high-profile activists that get arrested / get into international press are Russian-speakers from Almaty. Would need to.
  + Could look at last year and few months of protests
  + Might be super interesting if on FB—what % of chatter is in different languages? But might be worth narrowing down there.
* But could also be good home in Ukraine work already done!
* How much of a particular discourse is bots? Super interesting!
  + REALLY interesting in Azerbaijan—they attack us (Eurasianet, the press in general) every time we publish about Armenia
  + Would bots on FB be trackable? That’s where the meat might be.
  + Lots of anecdotal evidence that there are **troll factories** in Azerbaijan.
* Climate + food
  + Glacial melt – long-term data and food security stuff
    - Any bumper harvests in the short term?
    - Something always wanted to do: size of glaciers in Central Asia
    - Flows in rivers
  + But tying with COVID—trickiness.