

Sharing Data: Focusing on the Use of International Scientific Databases as an Important Tool in Science Diplomacy

The Chair, **Ernesto Fernandez Polcuch**, stated that this session differed from the proceedings of the morning insofar as its panelists would present on challenges and opportunities in multilateral, rather than bilateral, science diplomacy. To that end, he broke apart the concept of sharing data into its constituent parts.

Sharing, inherently, requires cooperation and since UNESCO's inception over 70 years ago, the organization has worked to achieve multilateral goals. He distinguished between the big science of multilateral institutions such as CERN and the multilateral agreements which emerge from meetings in the spirit of COP21 Sustainable Innovation Forum and the Sustainable Development Goals. Global agreements often cause real policy change in participant nations, and he hoped that this would be framed so that participants could learn from its effects.

Data, he continued, are mired in issues at each level from their production to movement about the world. He referenced the advent of the Internet of Things and the issues that would be faced in integrating standards between nations as well as common ground in development of this technology. He, too, spoke of data accessibility and the challenges faced in moving data, be they of natural resources, health or other personal information, or other research results between nations and the agreements that must be struck in coordinating collaboration over borders.

In closing, he stated that data sharing will grow to encompass all of the UN's 203 members and observers and that the panelists will address these opportunities.

The podium was passed to **Carson Lederer**. He discussed the history, mechanics, and epidemiology of haemoglobinopathies, which were once localized in the global south, or 'source countries'. They have changed their locality due to human migration, which poses challenges to the health systems of 'sink countries' in the global north. He then cited the high costs of treatment of these conditions, which tops 2 million USD in some countries.

Lederer continued that further stress would be placed on health systems by the current migration crisis from Syria and the Middle East. He then proposed action items, including the need for relationships with clinicians and educators on the ground, followed by a swift identification of infrastructural needs. This would then set the stage for autonomy through management, prevention and education of the population affected.

Lederer then spoke of databases, which he stated were a means to forge links, symbolize international commonalities, and widen collaboration. He explained the history of Ithant, a database dedicated to thalassemia and other haemoglobinopathies. In the same vein, the Global Globin 2020 Challenge was designed as a sister project to BRCAExchange, a database for the breast

cancer community. It is coordinated by South Africa and Malaysia and targeted to Low and Middle Income Countries (LMCs). The purpose of this project is to build capacity for genome analysis, establish best practices in genomic medicine and variant data sharing, as well as to forge a sustainable network of stakeholders.

Lederer also spoke of the Cyprus-Greek-Turkey 'Bermuda Triangle', so called due to the historical propensity for multilateral agreements between those stakeholders to 'sink'. There is great need to develop infrastructure in these areas of substantial haemoglobinopathy burden, but nuance is needed, as shared activities appear to legitimize the illegal regime in control of the north of Cyprus.

Aaron Wolf then took the stage to speak on international water policy. He began his talk with an admonition from former UN Secretary General Kofi Annan on the risks of conflict over fresh water in the future. He cited the six examples that the field of international water policy had noted as high-risk water zones before his lab began its work.

He then explained the path that his group took in amassing over 3,600 water treaties which exist between nations with transboundary watersheds, as well as information on water crises and conflicts over 4,500 years. The results showed that in 2/3 of cases, nations with transboundary watersheds have worked cooperatively and in cases of conflict 80% of exchanges were verbal. The group found no instances of all-out war over fresh water, even in times of drought.

From the database they concluded further that arid areas were more likely to cooperate than humid areas, and that the capacity to absorb change in water access yielded the least conflict. This was assessed by comparing new dams built with treaties against those built without treaties. They determined that treaties increased the institutional capacity to absorb change in water access. This finding was actionable for diplomats and policymakers; Wolf hopes that well-planned treaties will mitigate conflict in the future.

Wolf closed with an anecdote about Kura-Araks and parallel bilateralism. Azerbaijan and Armenia cannot share data; however, a neutral third party can build a database with each of them, which can be accessed by a fourth party. In this case, data can be aggregated without conflict.

Paula Dunbar closed out the session with the history of tsunami data sharing and international integration. She explained that NOAA holds historical data from 2000BC to the present and all types of data are aggregated, from cuneiform tablets to ships logs, newspapers, and seismological data. Historical data are only incorporated with an associated date.

One such historical tsunami occurred in 1700 and was dubbed an orphan event as it did not have a corresponding seismic event. In 1996, however, this was determined to be the result of a shift at the Cascadia fault line on the west coast of North America. This was determined through database use and creative dating of tree rings from trees that died from the resulting wave.

She then outlined the ways that tsunamis are measured, including death toll, sum of damages in USD, or whether or not they strike more than 1000km from the source. The last measure of tsunamis has had a great effect on policy and in three instances lead to changes in the way tsunamis are handled internationally.

In 1946, a tsunami sourced in Alaska and caused death in Hawaii, California, Tahiti, and Peru, as well as damage in New Zealand, which lead the US to create a warning system. Disaster struck again in 1960, which also caused death across the Pacific Ocean. The nations affected banded together and created a warning system along the shared coasts. Finally, in 2004, the system was upgraded in response to a tsunami in the Indian Ocean. It now covers every coastline and is composed of 500 warning stations.

Question and Answer Session

Q: Collaboration in databases is a virtuous cycle; the more partners there are, the more needs are met. How should fields start or be more effective in database construction and expansion?

A: Requiring data to be deposited in the public domain in sources connected together over the Internet relieves some of the burden from investigators. Generosity with data leads to better science overall and is a good incentive for investigators to release data.

Q: Do people in LMCs face barriers to database use?

A: While data are publicly available, a barrier to entry is sometimes the speed of internet connection. LMCs occasionally have difficulty contributing to databases, as the majority of their records are not digitized.

Q: Does the tsunami database include oral records?

A: Unless a date is included, no.

Q: Is there a place for public-private partnership in database maintenance and accession improvements?

A: In some cases, it is difficult to receive funding to maintain databases. However, private sector parties who express interest often want to charge a fee for data access, which is against the model outlined in most talks. The panelists hope that common ground can be found between the sectors.

Q: Is there a formal process to cross-pollinate between the fields so that happy accidents are more frequent?

A: It is good practice to attend meetings; however, much of the cross pollination outlined here was serendipitous.

Take-away points:

1. Interconnectivity between nations and the scope of big science has created a climate in need of multilateral agreements, which must be effective at both meeting the needs of stakeholders, as well as diplomatically nuanced to avoid conflict.
2. Parallel bilateralism has been an effective way to aggregate data between nations who do not engage in cooperative relations; however, this must be treated with great care so as to preserve the friendly status of neutral nations.
3. Scientists and database holders should be encouraged to combine modern experimental data with high-quality historical data, where practical and wise, to better inform the current state of the field.
4. Public data deposition should remain the expectation of scientists across all nations.