

BIO-FI: LEVERAGING THE POWER OF BIOLOGICAL BIG DATA FOR ADVANCED ANALYTICS AND MODELING OF CHEMICAL AND BIOLOGICAL THREATS

Towards Risk Assessment: A Quantitative Analysis Of Global Biological Research

Jonathan Pinney Gryphon Scientific LLC **Rocco Casagrande** Gryphon Scientific LLC **Jim Nolan** Gryphon Scientific LLC
Laurel MacMillan Gryphon Scientific LLC

Academic research drives scientific innovation and discovery and provides the training system for today and tomorrow's most talented scientists. The size and geographic diversity of the biological research community complicates the development of a clear understanding of the quantity and quality of pioneering research in various countries. This knowledge gap prevents an estimate of the ability of researchers to create novel dual-use information or techniques or to contribute to the development of novel biological agents. To accurately assess the risk posed by the creation and use of new bioweapons, we must gain a better understanding of the prevalence and quality of biological research across different regions of the world. We have employed a generalizable methodology that combines data mining and machine learning techniques to collect, extract, aggregate, and resolve disparate information about global biological research expertise to support risk estimates of global biological proliferation. Our combinatorial strategy enabled us to quantify research quality, normalize research metrics across countries, and apply machine learning techniques to compare nations and regions based on characteristics of their scientists. The study provides valuable insight on the international biological research acumen and capabilities of regions across the world relevant to the creation of new biological agents. These findings can be used by the broader biosecurity/biosafety community to identify and assess the biological research potential, as well as likelihood of misuse, by nations with both friendly and adversarial relationships to the United States and provides a framework for similar analyses to quantitatively explore additional variables related to biological research prevalence and quality.