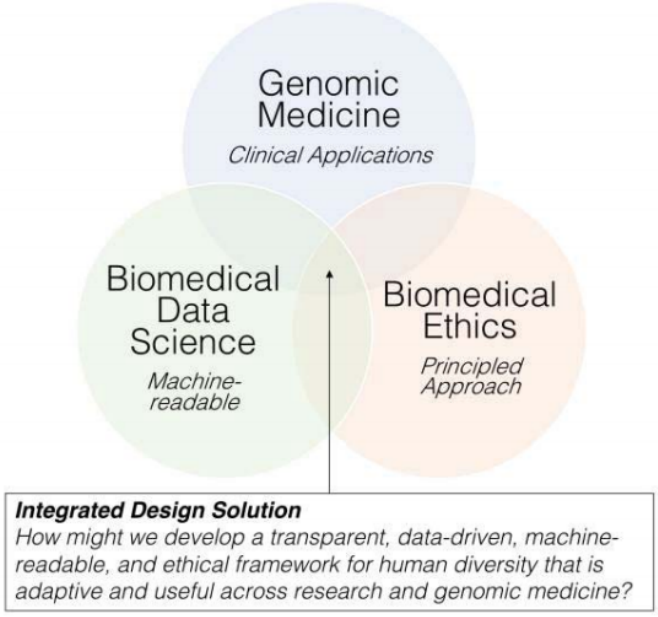
Docket #: S20-484

Methods for Measuring Diversity using Multi-Dimensional Networks

Researchers at Stanford have developed an adaptive, multi-dimensional process and software tool for characterizing human socio-demographic diversity. Current methods to characterize human diversity are rife with both conceptual and logistical issues including: missing data, misclassification errors, lack of validity and utility of labels, and data privacy. **The new, alternative approach offers a substantive improvement by allowing people to self-identify on the basis of their own words and language**. Utilizing the power of natural language processing and neural networks improves the validity and potential utility of socio-demographic data by allowing for open-ended responses. It expands the scope of data types and inputs, enabling individuals to truly self identify. The new approach solves multiple methodological issues across biomedical research and has the potential to help solve one of the most challenging aspects of genomics research: teasing apart genetic, socio-demographic, and environmental effects on phenotypes.



Trans-disciplinary approach integrating genomic medicine, biomedical data science, and biomedical ethics (image credit: Alice Popejoy)

Applications

- Research
- Healthcare systems
- Pharmaceutical companies
- Direct-to-consumer genetic testing

Advantages

- Allows people to self-identify on the basis of their own words and language
- Addresses longstanding misclassification issues
- Enables a secure, de-identified, scalable, and machine-readable method to collect and store free-text diversity information

Innovators

- Alice Popejoy

Licensing Contact

Kimberly Griffin

Technology Licensing and Strategic Alliances Manager

Email