

RON BROWN EXCELLENCE IN INNOVATION AWARD



Dr. Ellen M. Voorhees

*Information Technology Laboratory
National Institute of Standards and Technology*

Dr. Voorhees led the design and implementation of the NIST TREC COVID Challenge, an evaluation challenge for information retrieval that provided unprecedented capability in accurately retrieving reliable, contemporaneous information for the global fight against the novel coronavirus, COVID-19. The resulting search technology developed from the TREC COVID Challenge under Dr. Voorhees's leadership was quickly deployed by researchers worldwide to improve search technology in the fight against the ongoing pandemic.

Within days after the White House call to action to the artificial intelligence (AI) community in response to the COVID-19 virus, Dr. Voorhees built a coalition with the National Institutes of Health (NIH), the Allen Institute for AI, Kaggle, and two universities. Dr. Voorhees designed and led the TREC COVID Challenge to evaluate information retrieval (IR) technologies on scientific literature, leading five evaluation challenge cycles from April–July 2020. Each successive cycle measured progress and provided feedback to improve system development for the next cycle. Participation in TREC COVID grew to 92 unique teams, including Google and Microsoft, and 556 algorithms were created and tested.

Dr. Voorhees proposed to enhance COVID Open Research Dataset (CORD19), a large set of scientific papers on virology CORD, into a "test collection," with search queries and ground truth relevance judgments indicating which papers should be retrieved for each query. The test collection would enable measurements to drive the research, giving birth to the NIST TREC COVID Challenge. Typically, the creation of a test collection can take up to a year, but this challenge demanded a faster outcome. Dr. Voorhees designed a process with multiple cycles that incorporated relevance feedback, so that developed algorithms could learn and improve at each stage. Dr. Voorhees mobilized the TREC community to rapidly iterate on the data. This combination of iterative learning with a strong research community that understands search technology and the subject domain produced a large, reusable resource in record time.

Dr. Voorhees produced the blueprint for IR evaluations' methods, and built trust in information returned by search engines in support of fighting COVID and other pandemics—a lasting resource for biomedical research. Her critical work and remarkable achievement have directly contributed to the acceleration of lifesaving research in the fight against COVID-19.