

# The Principles of Depth, Breadth, and Precarity in Developing Metrics for Information Retrieval Evaluation

Fernando Diaz  
CMU

## ABSTRACT

In this perspective presentation, I will review two core methods of improving the evaluation of information retrieval systems. Innovation in depth refers to deepening our empirical and theoretical understanding of effectiveness within a class of information retrieval problems. Innovation in breadth refers to understanding effectiveness in novel or underrepresented information access patterns. I will describe how different styles of evaluation research fit into these classes, connecting recent work in fairness and preference-based evaluation to a longer tradition of information retrieval research. I will close by reflecting on the value of seeing evaluation research as brittle and precarious.

## BIOGRAPHY

Fernando Diaz is an Associate Professor in the Language Technologies Institute at CMU and a researcher at Google Research. Fernando designs and evaluates search engines and recommender systems, covering three themes: quantitative evaluation of AI systems, retrieval-enhanced AI, and assessment of the cultural impact of AI in domains like music and literature. He has previously held both individual contributor and senior research leadership roles at Microsoft Research, Spotify, and Yahoo Research. He is particularly interested in evaluating and optimizing deployment of real AI systems.