

The Promise and Perils of Found Data

Moderator:

Jonathon Phillips

Panelists:

Ross Beveridge

Kevin Bowyer

Jeff Cohn

The Audience



Definitions

- Laboratory: Data collected in a laboratory, usually following a protocol.
- Found: Data obtained opportunistically
- Designed: Data collected to answer a question
- Undesigned: Generally for exploratory analysis

Background

- Until about ten years ago, virtually all publicly available face and gesture data sets were collected in a laboratory.
- The Labeled Faces in the Wild (LFW) data set popularized data sets from the web.
- Strength and weakness of each data
- Role in experiments, challenges, and evaluations

Guidance

- I would like to focus the panel on a discussion of the future. I do not want to focus on the shortcomings of previous data sets.