

Making the Best Use of Textual ED Data for Syndromic Surveillance

Julia Gunn

Boston Public Health Commission

Data Quality

A concept in need of an adjective

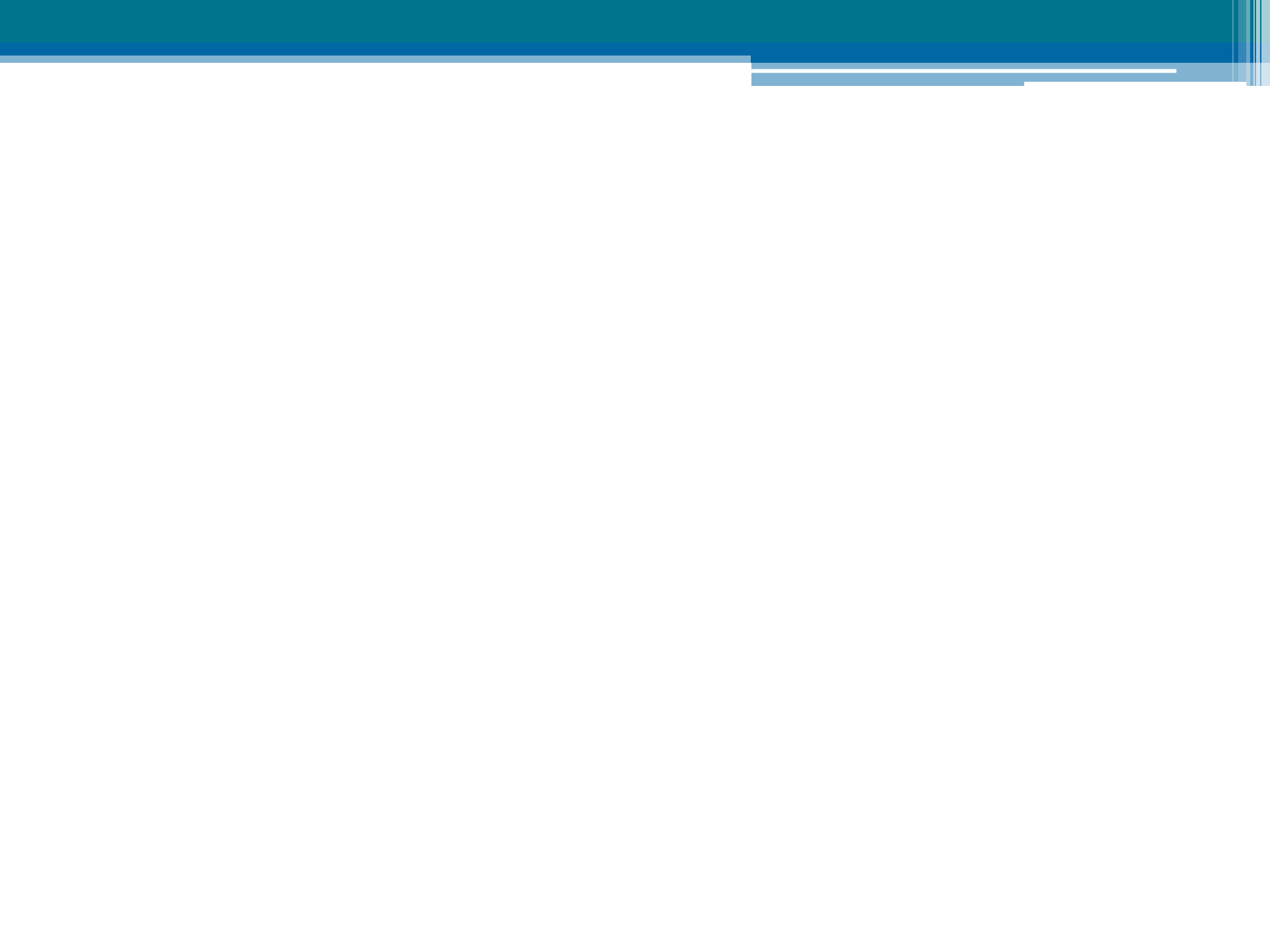
Data Quality Challenges

- **Technical**

- Transmission
 - Systems offline
- Messaging
 - Errors
 - Changing standards
- Structural
 - Recoded fields
 - Restructured

- **Linguistic**

- Terminology
 - Spelling
 - Abbreviations
 - Shift
 - Haiti earthquake
 - Hurricanes
 - Drift
 - New issues
 - MERS, Ebola
 - Changing practice
 - Narcan
- Contextual
 - Negation – No , denies,



EMT-P and UMLS: The Boston Experience

- Training – Not all terms map
 - Hd = dialysis
 - scd = sickle cell disease
 - ooc = out of control
- One to Many (EMT-P)
 - Boston - Many to One
- Terms with multiple meanings
 - OD: Overdose vs Right eye
 - ST: Sore throat vs Sinus tachycardia
 - Context matters

Impact

- Decrease processing time
- Fewer unique terms
 - Standardized: 14,893 unique words
 - Original (unaltered chief complaints): 22,748 unique words.
- Facilitates identification of new terms – what are you missing

Language processing tools will be critical for public health to efficiently and accurately process electronic health record data.