

Linking Video Segments to Relevant Wikipedia Content

Victor Lavrenko, Johanna Moore, Sean Moran

- Help the BBC open up their Archive
 - UK cultural heritage (~10⁶ hours of video)
 - mostly inaccessible (analog, no metadata)
- Turn archive into a set of LEGO bricks
 - short chunks focused on a single subject
 - annotated, searchable, hyper-linked
- Why we are doing this
 - catalyst for creative reuse of the content
 - interactive social platform for studies
 - pilot for proposal on next-generation TV
- Research issues
 - overcoming vocabulary mismatch (ASR ↔ Wikipedia)
 - distance functions for hierarchical alignment
 - architecture: massive but localized computation

