SEARCH AND SEMANTIC SEARCH

Indian Institute of Technology Kanpur
Commonwealth of Learning Vancouver
T.V. Prabhakar
IIT Kanpur
Search Engines

• Do not search the web in real time
• Search an index of the web
• Created by crawling with spiders
• Searches the documents for the occurrence of the “search string” - substring search
• After retrieving the documents does ranking
• Ranking is critical and complex
  – Google uses 200+ factors for ranking
Google Webmasters Tools

• How search works
  https://youtu.be/BNHR6IQJGZs

• More detailed
  https://youtu.be/KyCYyoGusqs
But

Holstein

Jersey
How do we do this?

• The system needs to know that Holstein is a breed of cow and Jersey is another breed.
We call these

• Knowledge Models
• Useful in
  – Semantic Search
  – Related Documents
  – Inference
  – Teaching
  – ..
THANK YOU