Disambiguating Web Tables using Partial Data

Web table disambiguation
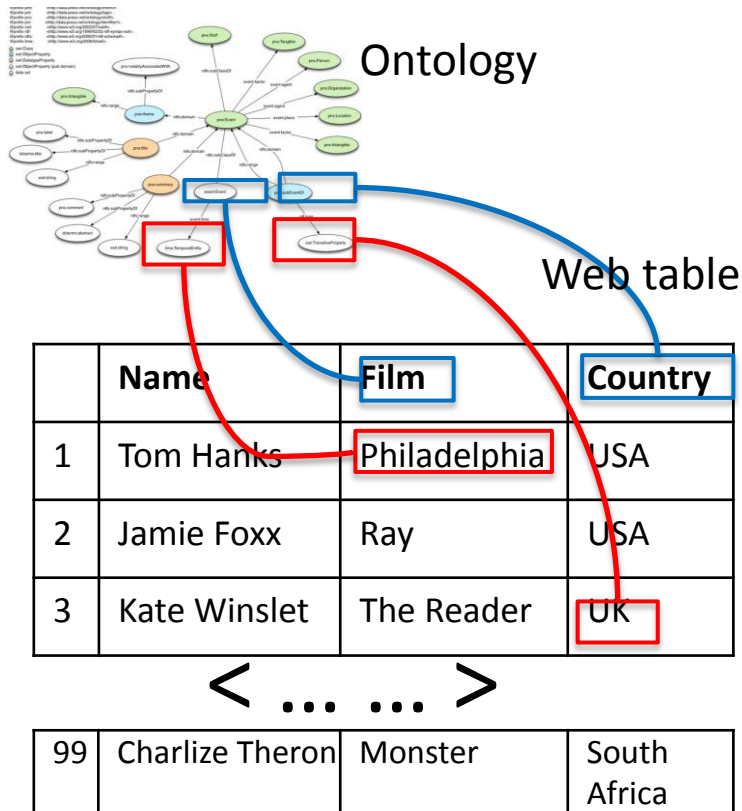

Ontology

Web table

	Name	Film	Country
1	Tom Hanks	Philadelphia	USA
2	Jamie Foxx	Ray	USA
3	Kate Winslet	The Reader	UK

< ... ... >

| 99 | Charlize Theron | Monster | South Africa |

Table of Best Actor/Actress

Partial data

Linking concepts to headers: can we do this without reading the entire data in the table, but only analysing a **sample**?

Sample selection: does the choice of sample matter? If so how to make the right choice?