

Named Entity Extraction

What does it do? Given a text, it gives you a list of people and things mentioned in it. The most popular implementation extracts names of people, places, and organizations.

What's so great about it? Sometimes you don't really want to work with text: you just want to Hoover out specific details. Maybe you want to make a map of places mentioned in a book. Maybe you want to make a social network based on who shows up in social registers.

What choices do I have to make?

Are there alternate implementations? Sort of. The Stanford Natural Language Toolkit contains