A Hitchhiker’s guide to Ontology

Fabian M. Suchanek
Télécom ParisTech University
Paris, France
I am an Elvis Fan!

1961
Meeting Elvis Presley

1961

2014
Meeting Elvis Presley

1961

At Elvis’ ranch in New Zealand (April 2014)

2014

Ranch is Murchison Lodge in New Zealand
Knowledge Cards

Elvis Presley - Wikipedia, the free encyclopedia
en.wikipedia.org/wiki/Elvis_Presley
Elvis Aaron Presley (January 8, 1935 – August 16, 1977) was an American singer and actor. Regarded as one of the most significant cultural icons of the 20th ...
Priscilla Presley - Albums discography - Singles discography - Michael Lockwood

Elvis Presley Biography - Elvis Life Story - Elvis Presley
www.elvis.com/about-the-king/biography.aspx
Elvis Presley The incredible Elvis life story began when Elvis Aaron Presley was born to Vernon and Gladys Presley in a two-room house in Tupelo, Mississippi, ...

Elvis Presley Official Web Site
www.elvis.com/
Play the Elvis Slots on Slotomania. Play slots like a king with the brand new "ELVIS" Slotomania video slots game now available through Facebook! Prev Next.

In the news
Listening to Elvis Presley's Bizarre Album of Stage Banter
Mental Floss - 20 hours ago
In 1974, Elvis Presley released Having Fun With Elvis on Stage, an album consisting of

Elvis Presley
Singer
Elvis Aaron Presley was an American singer and actor. Regarded as one of the most significant cultural icons of the 20th century, he is often referred to as "the King of Rock and Roll", or simply, "the King". Wikipedia
Died: August 16, 1977, Memphis, Tennessee, United States
Spouse: Priscilla Presley (m. 1967–1973)
Children: Lisa Marie Presley
Songs
My Way
Jailhouse Rock
Love Me Tender
Love Me
1973
1973
1972
1972
Aloha From Hawaii via... As Recorded at Madison... As Recorded at Madison...
Knowledge Cards

Elvis Presley

Elvis Presley - Wikipedia, the free encyclopedia
en.wikipedia.org/wiki/Elvis_Presley
Elvis Aaron Presley (January 8, 1935 – August 16, 1977) was an American singer and actor. Regarded as one of the most significant cultural icons of the 20th century.

Elvis Presley Biography - Elvis Life Story - Elvis Presley
www.elvis.com/about-the-king/biography_aspx
Elvis Presley The incredible Elvis life story began when Elvis Aaron Presley was born to Vernon and Gladys Presley in a two-room house in Tupelo, Mississippi, ...

Elvis Presley Official Web Site
www.elvis.com/
Play the Elvis Slots on Slotomania. Play slots like a king with the brand new “ELVIS” Slotomania video slots game now available through Facebook! Prev Next.

In the news
Listening to Elvis Presley's Bizarre Album of Stage Banter
Mental Floss - 20 hours ago
In 1974, Elvis Presley released Having Fun With Elvis on Stage, an album consisting of

Elvis Presley
Singer
Elvis Aaron Presley was an American singer and actor. Regarded as one of the most significant cultural icons of the 20th century, he is often referred to as "the King of Rock and Roll", or simply, "the King".

Died: August 16, 1977, Memphis, Tennessee, United States
Spouse: Priscilla Presley (m. 1967–1973)
Children: Lisa Marie Presley

Songs
<table>
<thead>
<tr>
<th>Song</th>
<th>Year</th>
<th>Album</th>
</tr>
</thead>
<tbody>
<tr>
<td>My Way</td>
<td>1973</td>
<td>Aloha From Hawaii via...</td>
</tr>
<tr>
<td>Jailhouse Rock</td>
<td>1972</td>
<td>As Recorded at Madison...</td>
</tr>
<tr>
<td>Love Me Tender</td>
<td>1972</td>
<td>As Recorded at Madison...</td>
</tr>
<tr>
<td>Love Me</td>
<td>1972</td>
<td>As Recorded at Madison...</td>
</tr>
</tbody>
</table>
For us, a knowledge base ("KB", "ontology") is a graph, where the nodes are entities and the edges are relations.

(We do not distinguish T-Box and A-Box.)
Knowledge Base Life Cycle

Le Monde

Applying knowledge bases

Constructing knowledge bases

Wikipedia

Mining knowledge bases

Singer

Born 1935

Type

\[ A \land B \Rightarrow C \]
Elvis Presley

Elvis Presley was one of the best blah blah blub blah don’t read this, listen to the speaker! blah blah blah blubl blah you are still reading this! blah blah blah blah blabbel blah

Categories:
Rock&Roll, American Singers, Academy Award winners...
Creating a large knowledge base

Intermediate extractors
- clean facts
- deduplicate facts and entities
- check consistency

Ensuring high quality (95%)
Creating a large knowledge base

Intermediate extractors
- clean facts
- deduplicate facts and entities
- check consistency

ensuring high quality (95%)
Example: YAGO about Elvis
YAGO: a large knowledge base

Wikipedia + WordNet

time and space

multilingual (soon)

100 relations

100m facts

10m entities

95% accuracy

used by DBpedia

Caveat:

focus on precision!

Knowledge Base Life Cycle

LeMonde

Wikipedia

Constructing knowledge bases

Minning knowledge bases

Applying knowledge bases

A ∧ B ⇒ C
Rule Mining finds patterns

\[\text{married}(x, y) \land \text{hasChild}(x, z) \Rightarrow \text{hasChild}(y, z)\]
Rule Mining finds patterns

married(x, y) \land hasChild(x, z) \Rightarrow hasChild(y, z)

But: Rule mining needs counter examples and RDF ontologies are positive only
AMIE finds rules in ontologies

AMIE is based on an efficient in-memory database implementation.

Caveat: rules cannot predict the unknown with high precision.
AMIE finds rules in ontologies

\[ \text{AMIE finds rules in ontologies} \]

\[ \text{AMIE} \quad \text{(5min)} \quad \Rightarrow \quad \text{type}(x, \text{pope}) \Rightarrow \text{diedIn}(x, \text{Rome}) \]

[WWW 2013]
Knowledge Base Life Cycle

Le Monde

Applying knowledge bases

Constructing knowledge bases

Wikipedia

Mining knowledge bases

A ∧ B ⇒ C

Singer

1935

Protec ting knowledge bases
Plagiarism

People may “steal” from other ontologies without giving due credit. Most ontologies have licenses that require attribution.
Additive Watermarking

By adding a few fake facts to the source ontology, one can prove theft in the target ontology.

[WWW 2012 demo]
Subtractive Watermarking

One can also prove theft by selectively removing facts from the source ontology.

YAGO

stolen

EvilPedia

[ISWC 2011]
Protecting the user

Mat Honan
Protecting the user

Mat Honan
Protecting the user

Mat Honan
Protecting the user

Mat Honan
Protecting the user

Mat Honan
Protecting the user

Mat Honan
Protecting the user

Mat Honan

Read Mat's story on Wired
Protecting the user

Mat Honan

Owen Williams

Read Mat's story on Wired
Protecting the user

Mat Honan

Two-factor authentication!

Owen Williams

Read Mat's story on Wired
Protecting the user

Mat Honan

Two-factor authentication!

Owen Williams

Read Mat’s story on Wired
Protecting the user

Mat Honan

Recovery codes!

Two-factor authentication!

Owen Williams
Protecting the user

Mat Honan

Recovery codes!

Owen Williams

Two-factor authentication!

Read Mat’s story on Wired

Read Owen’s story on The Next Web
Account Dependencies

- gmail-access
  - gmail-data
  - gmail-destruction
  - gmail-1stfactor
    - gmail-password
    - yahoo-data
  - gmail-2ndfactor
    - phone-authentication
      - phone-app
        - phone-password
      - phone-SMS
        - phone-possession
        - phone-simPin
    - phone-recovery
      - gmail-trusted-device
        - mac-access
        - dell-access
Account Dependencies

![dependencies diagram]

Safety: How many keys can I afford to lose, and still have access to my data?

Security: How many keys does a hacker need to access my data?

Protection: How many keys are needed to destroy my data?
DIVINA
Discovering Vulnerabilities of Internet Accounts

Dropbox

On which computers do you have Dropbox installed?
mac, dell

Do you have two-factor authentication enabled?  
Yes  No
DIVINA
Discovering Vulnerabilities of Internet Accounts

Dropbox

On which computers do you have Dropbox installed?
mac, dell

Do you have two-factor authentication enabled?  
☐ Yes  ☐ No


dropbox

phone

password1

password2
DIVINA
Discovering Vulnerabilities of Internet Accounts

<table>
<thead>
<tr>
<th>Safety</th>
<th>Security</th>
<th>Protection</th>
</tr>
</thead>
</table>
| • gmail-password yahoo-access ✓ | • gmail-password phone-access ✓ | X

For online accounts: Consider two-factor authentication, lock codes on your phone, or encryption of your trusted devices.
**DIVINA**

Discovering Vulnerabilities of Internet Accounts

[WWW 2015 demo]

http://suchanek.name/programs/divina

<table>
<thead>
<tr>
<th>Safety</th>
<th>Security</th>
<th>Protection</th>
</tr>
</thead>
<tbody>
<tr>
<td>![Image]</td>
<td>![Image]</td>
<td>![Image]</td>
</tr>
</tbody>
</table>
| gmail-password yahoo-access | gmail-password phone-access | ![Image]  

**For online accounts:** Consider two-factor authentication, lock codes on your phone, or encryption of your trusted devices.
Knowledge Base Life Cycle

Le Monde

Applying knowledge bases

Constructing knowledge bases

Wikipedia

Mining knowledge bases

Singer

type

born

1935

A \land B \Rightarrow C

Protecting knowledge bases

\[ A \land B \Rightarrow C \]
Mining Le Monde

Le Monde

1944

2013
Mining Le Monde

Le Monde

1944 2013

yago
select knowledge
Mining Le Monde

Le Monde

1944 2013

yago
select knowledge
<table>
<thead>
<tr>
<th>time</th>
<th>place</th>
<th>entity</th>
</tr>
</thead>
<tbody>
<tr>
<td>1967</td>
<td>USA</td>
<td>Elvis</td>
</tr>
</tbody>
</table>
Mining Le Monde

Average age of people mentioned
Presence of foreign companies

red: many foreign companies mentioned
blue: few foreign companies mentioned
Mining Le Monde

[AKBC 2013]

[VLDB 2014 vision]
Knowledge Base Life Cycle

Knowledge bases

Constructing knowledge bases

Wikipedia

Mining knowledge bases

A ∧ B ⇒ C

Singer

type

born

1935

Applying knowledge bases

Protecting knowledge bases

A ∧ B ⇒ C
Is Elvis dead?

Elvis Presley
Singer

Elvis Aaron Presley was an American singer and actor. Regarded as one of the most significant cultural icons of the 20th century, he is often referred to as "the King of Rock and Roll", or simply, "the King". Wikipedia

Died: August 16, 1977, Memphis, Tennessee, United States
Spouse: Priscilla Presley (m. 1967–1973)
Children: Lisa Marie Presley

Elvis died 1977
Is Elvis dead?

Elvis Presley
Singer

Elvis Aaron Presley was an American singer and actor. Regarded as one of the most significant cultural icons of the 20th century, he is often referred to as "the King of Rock and Roll", or simply, "the King". Wikipedia

**Died:** August 16, 1977, Memphis, Tennessee, United States
**Spouse:** Priscilla Presley (m. 1967–1973)
**Children:** Lisa Marie Presley

Elvis died 1977

100m statements
95% precision
-> 5m wrong statements
Thank you for your attention!

Constructing knowledge bases

Wikipedia

Mining knowledge bases

A \land B \Rightarrow C

Singer

1935

Applying knowledge bases

Le Monde

Protecting knowledge bases

Slides done with Powerline, my free SVG slide editor with LaTeX support