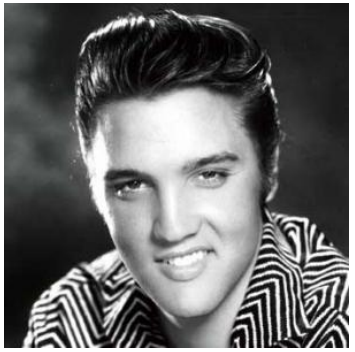
A Hitchhiker's guide to Ontology

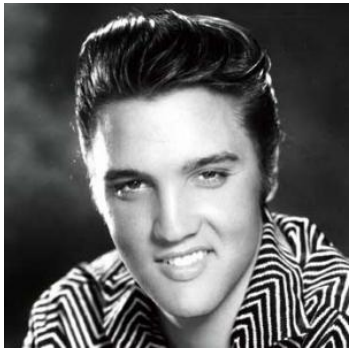
Fabian M. Suchanek
Télécom ParisTech University
Paris, France

I am an Elvis Fan!



1961

Meeting Elvis Presley

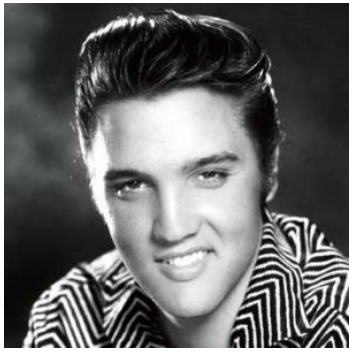


1961

2014

Old Elvis from Sachsmedia

Meeting Elvis Presley



1961

2014

At Elvis' ranch
in New Zealand
(April 2014)

Old Elvis from Sachsmedia

Ranch is Murchison Lodge in New Zealand

Knowledge Cards

Google

Elvis Presley

Elvis Presley - Wikipedia, the free encyclopedia
en.wikipedia.org/wiki/**Elvis_Presley** ▾ Wikipedia ▾
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 ▾
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/ ▾ Elvis Presley ▾
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	1973	Aloha From Hawaii via ...
Jailhouse Rock		
Love Me Tender	1972	As Recorded at Madiso...
Love Me	1972	As Recorded at Madiso...

Knowledge Cards

Google

Elvis Presley

Elvis Presley - Wikipedia, the free encyclopedia
en.wikipedia.org/wiki/**Elvis_Presley** ▾ Wikipedia ▾
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 ▾
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/ ▾ Elvis Presley ▾
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

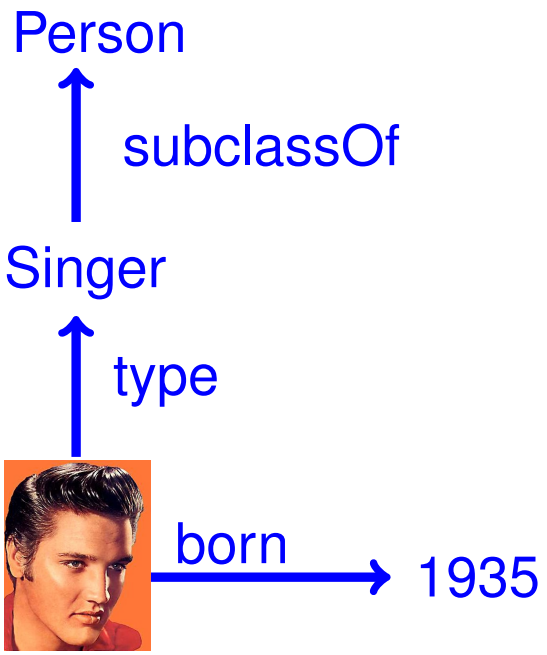
Died: August 16, 1977, Memphis, Tennessee, United States
Spouse: Priscilla Presley (m. 1967–1973)
Children: Lisa Marie Presley

???

Songs

My Way	1973	Aloha From Hawaii via ...
Jailhouse Rock		
Love Me Tender	1972	As Recorded at Madiso...
Love Me	1972	As Recorded at Madiso...

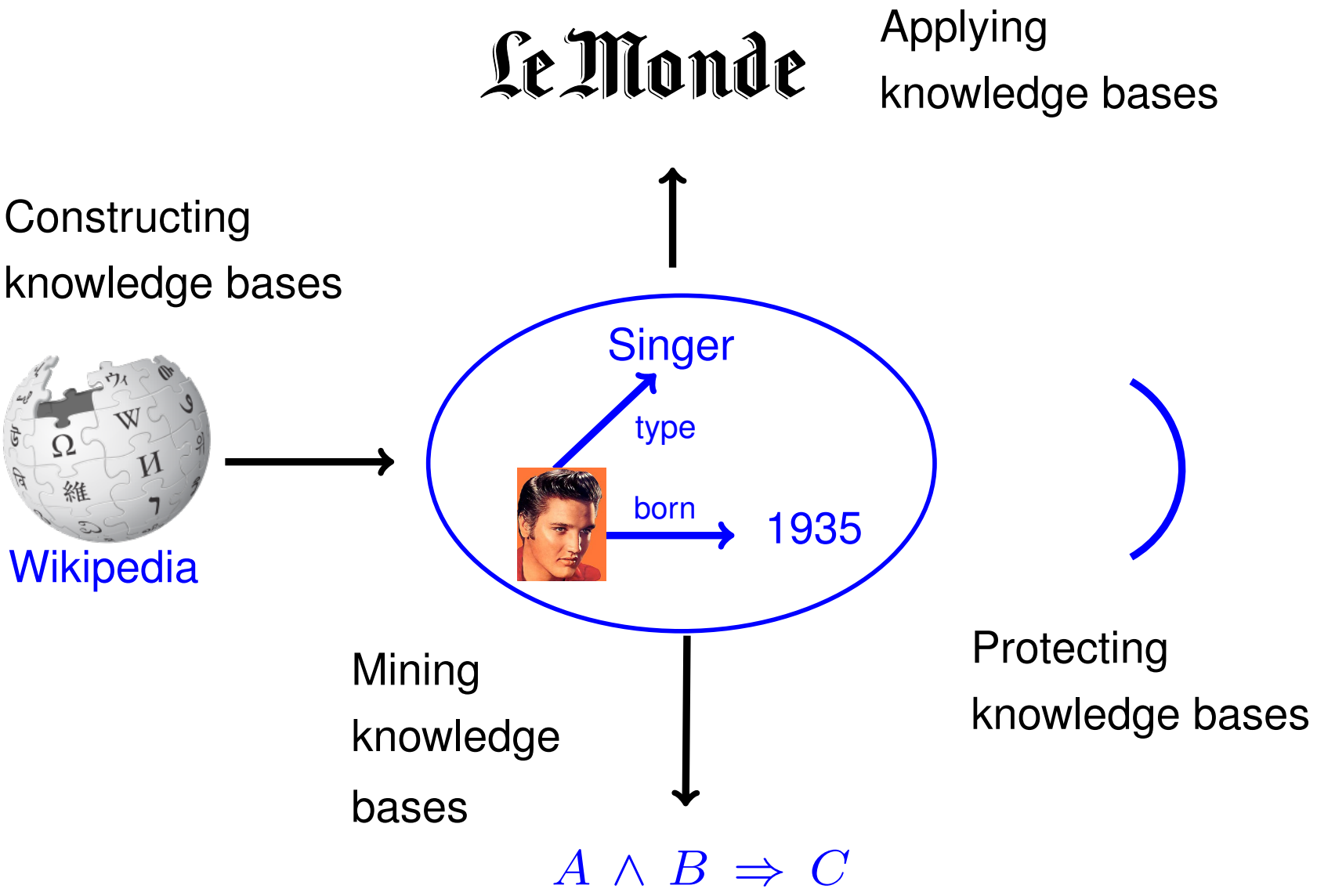
Knowledge Bases



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

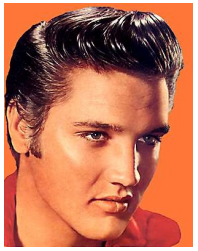


Extracting from Wikipedia

Elvis Presley

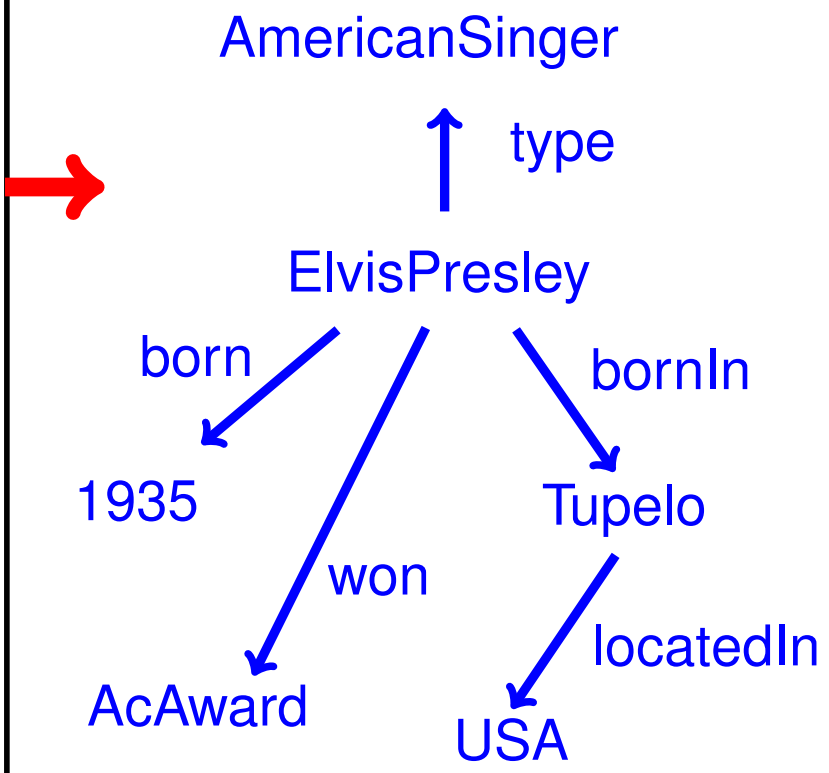
WIKIPEDIA
The Free Encyclopedia

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

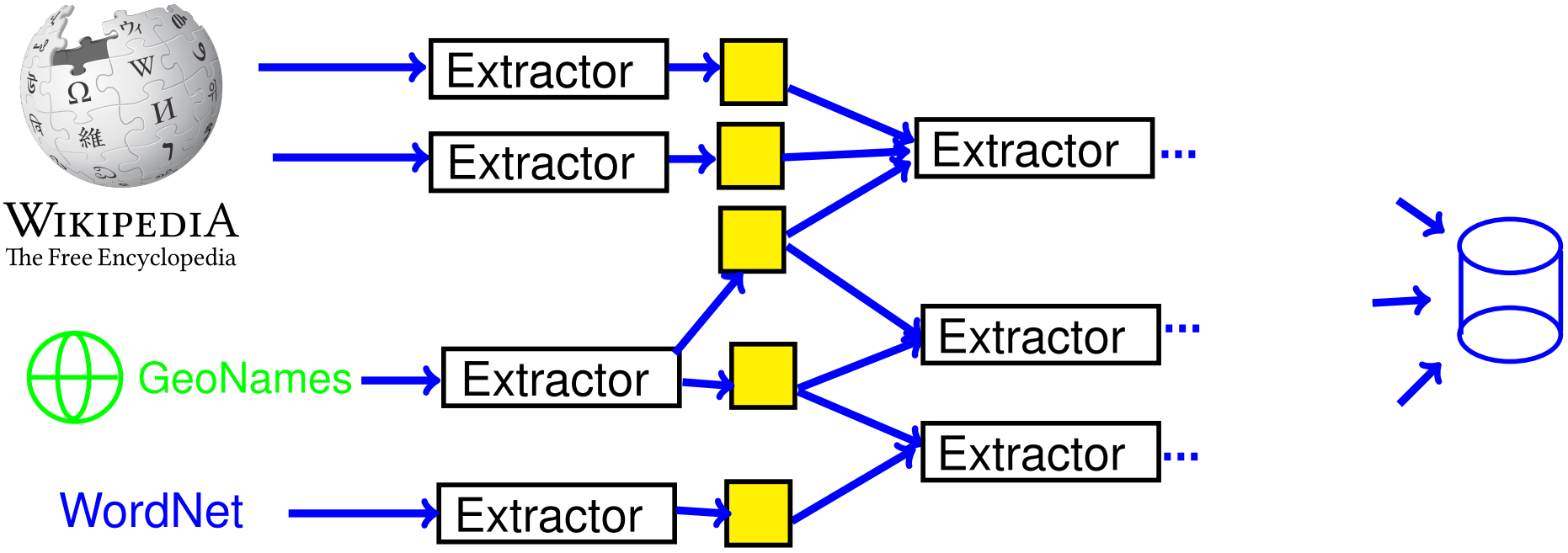


Born: 1935
In: Tupelo
...

Categories:
Rock&Roll, American Singers, Academy Award winners...



Creating a large knowledge base



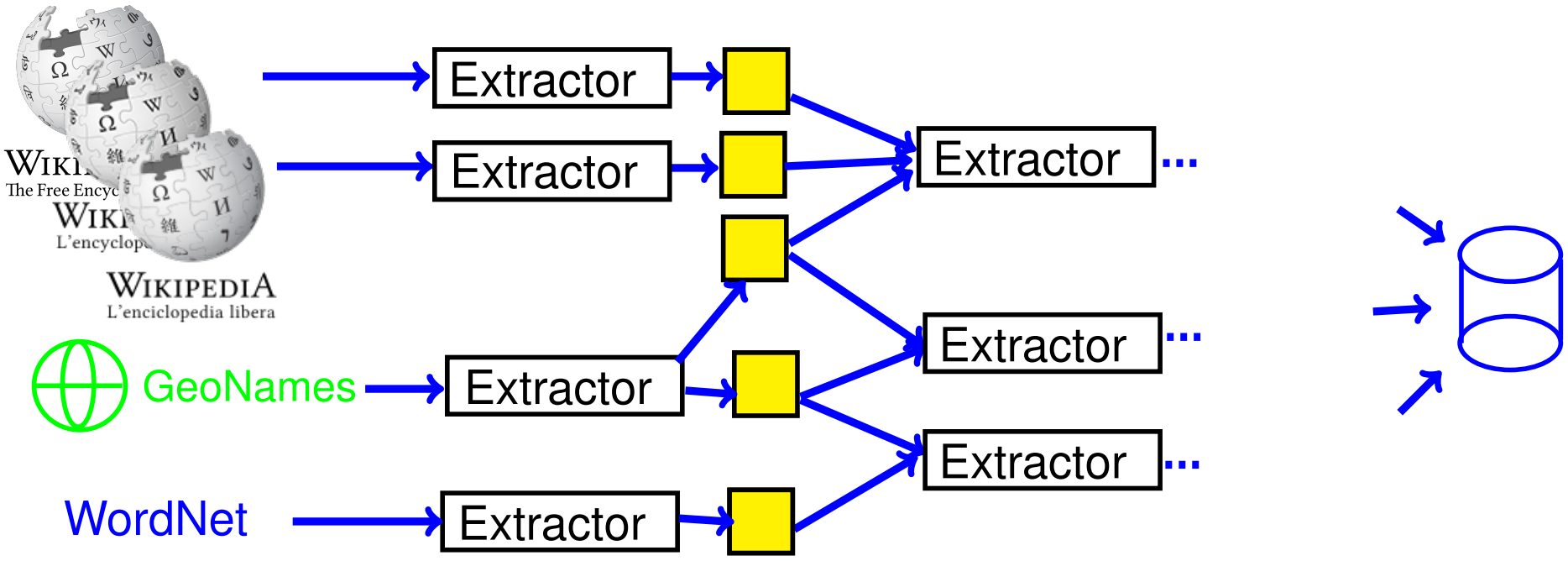
Intermediate extractors

- clean facts
- deduplicate facts and entities
- check consistency

ensuring high quality (95%)

See it online See it locally

Creating a large knowledge base

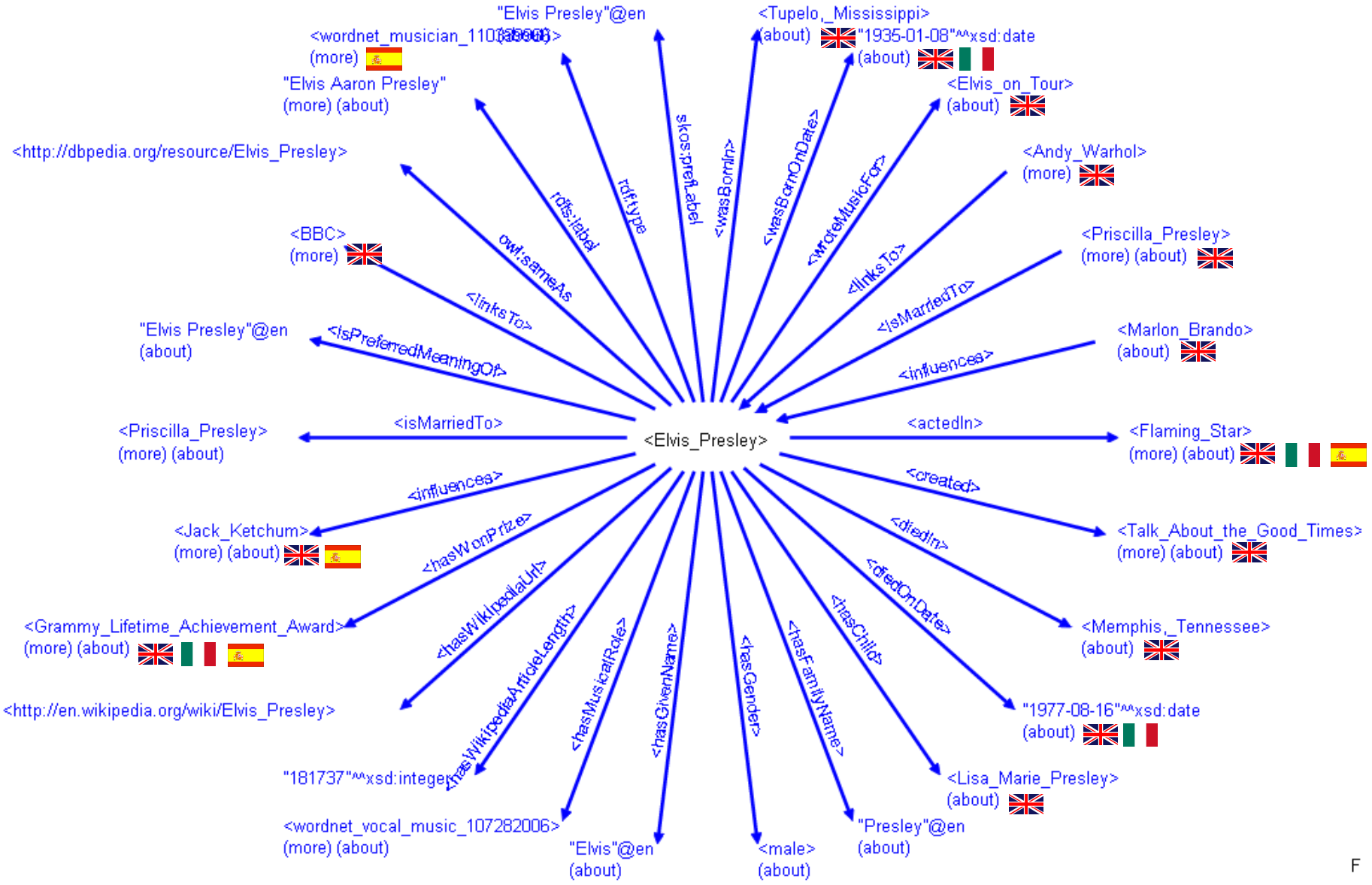


Intermediate extractors

- clean facts
- deduplicate facts and entities
- check consistency

ensuring high quality (95%)

Example: YAGO about Elvis



YAGO: a large knowledge base

http://yago-knowledge.org

Wikipedia + WordNet

time and space

multilingual (soon)

100 relations

100m facts

10m entities

95% accuracy

used by DBpedia

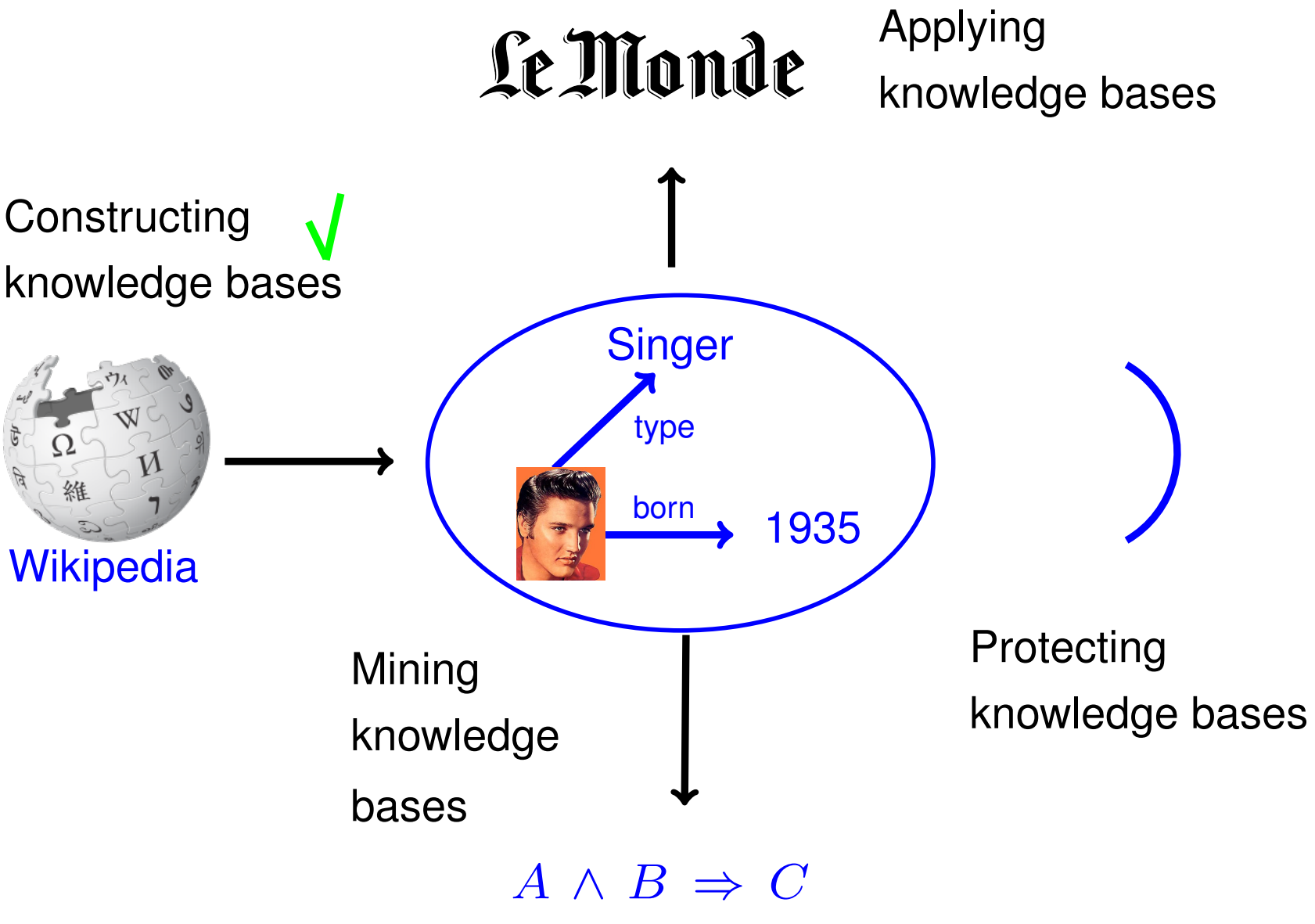
Caveat:

focus on

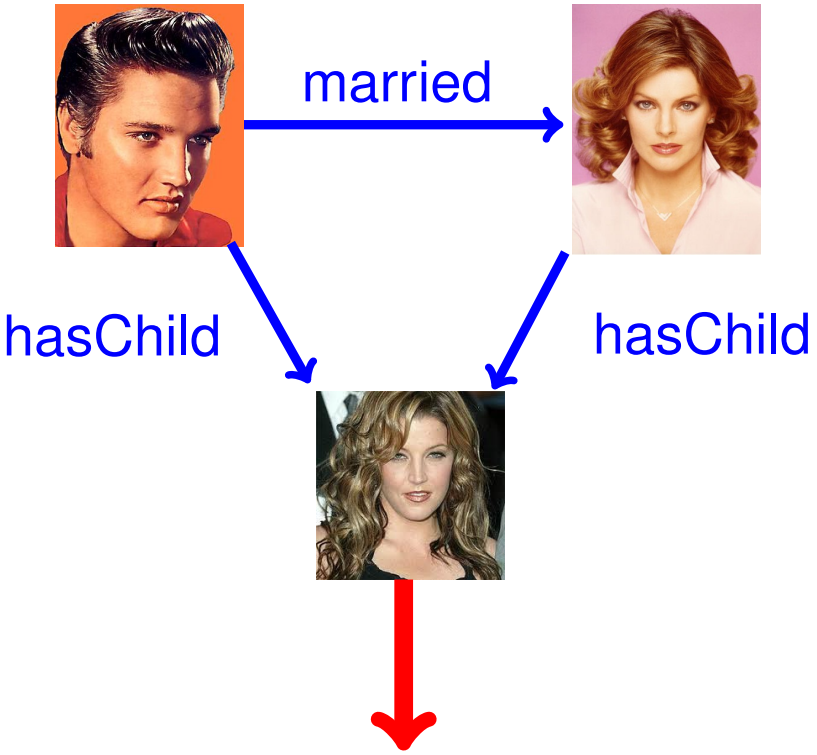
precision!

[WWW2007,JWS2008,WWW2011 demo,AIJ2013,WWW2013 demo,CIDR20

Knowledge Base Life Cycle

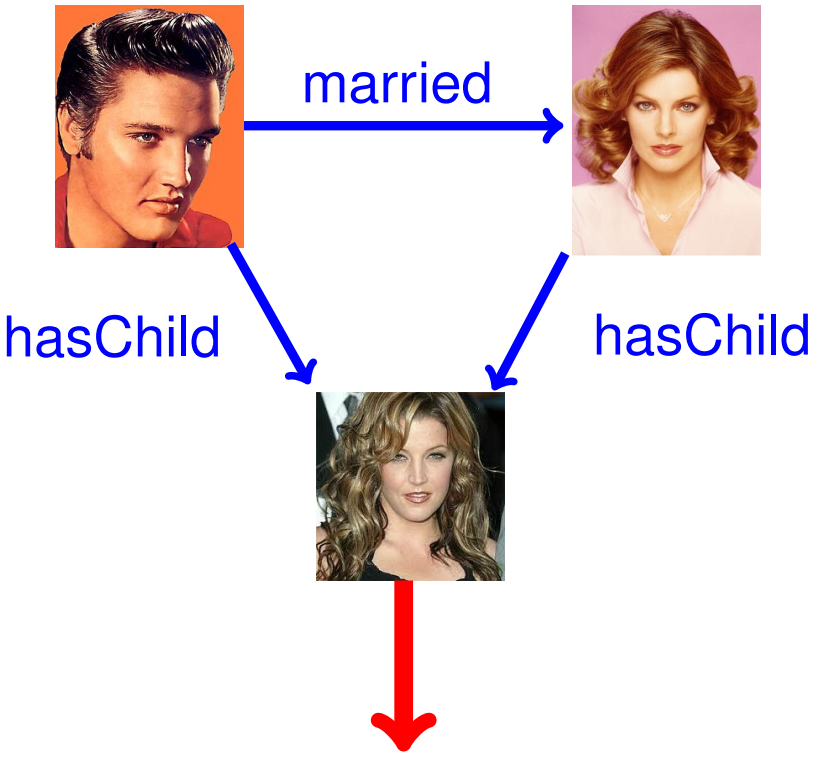


Rule Mining finds patterns



$$married(x, y) \wedge hasChild(x, z) \Rightarrow hasChild(y, z)$$

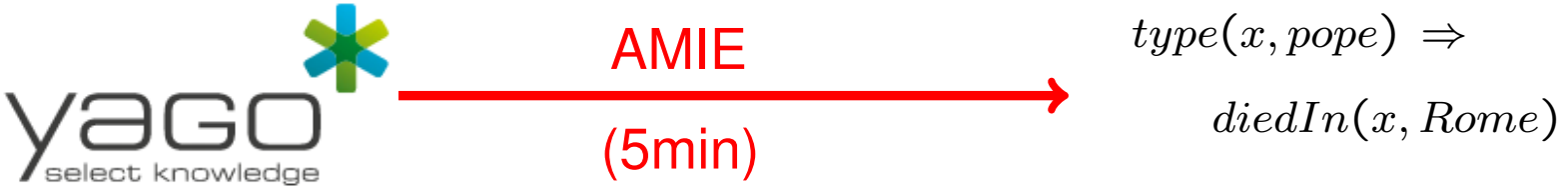
Rule Mining finds patterns



$$married(x, y) \wedge hasChild(x, z) \Rightarrow hasChild(y, z)$$

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

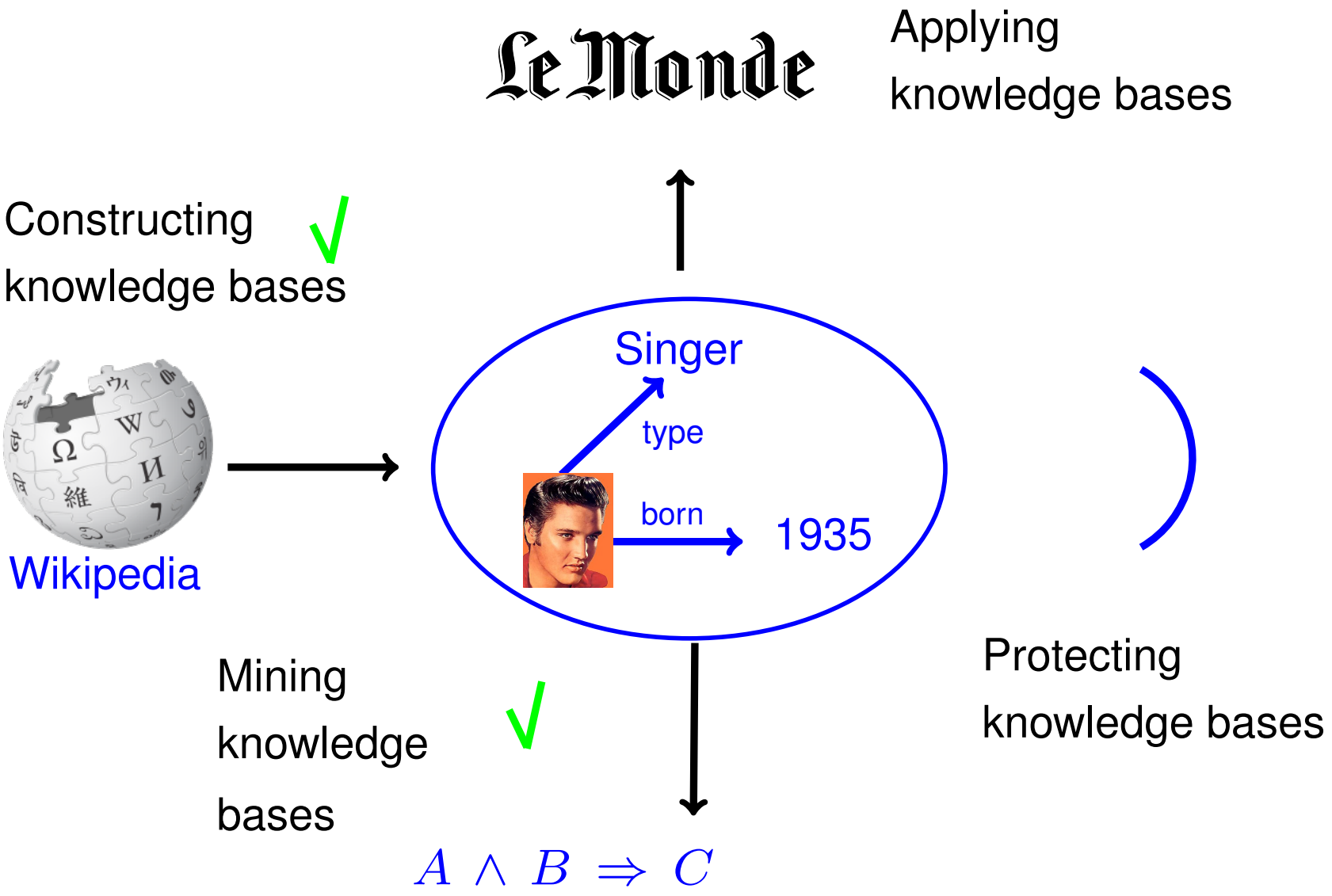
AMIE

(5min)

$$type(x, pope) \Rightarrow$$
$$diedIn(x, Rome)$$

[WWW 2013]

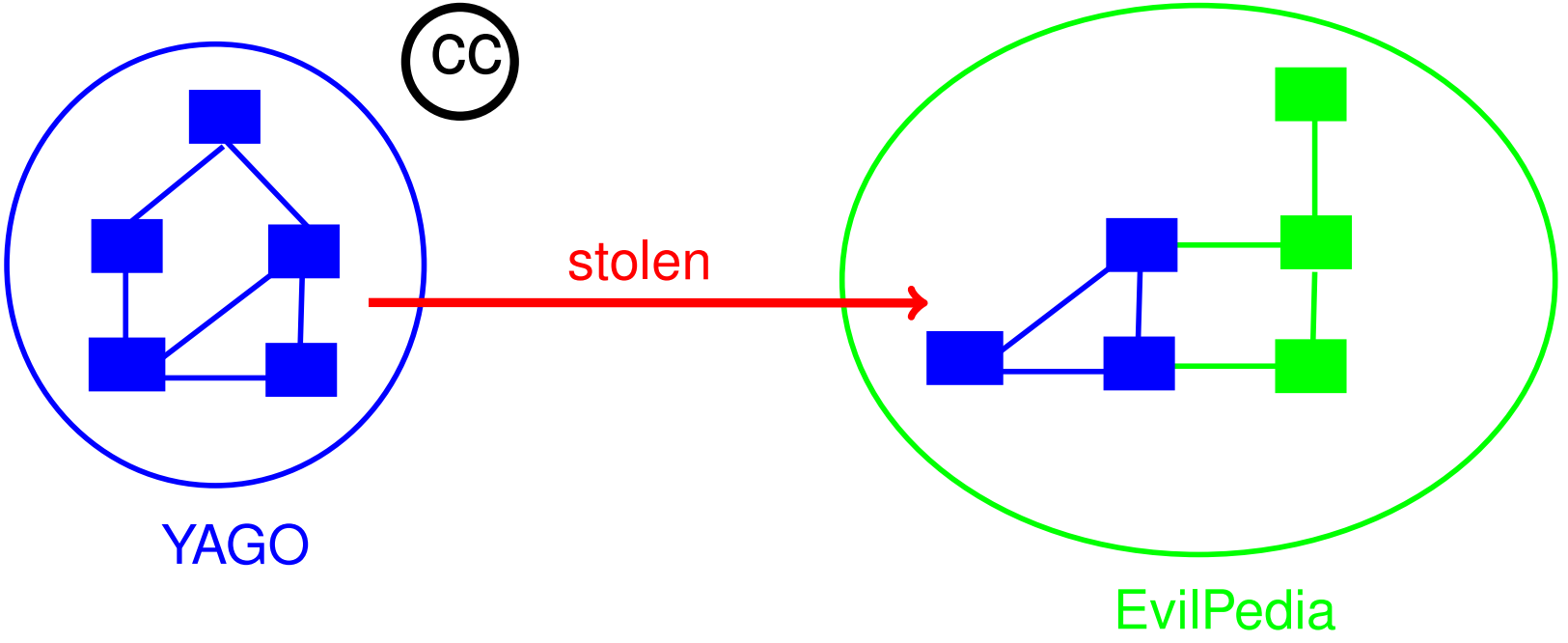
Knowledge Base Life Cycle



Plagiarism

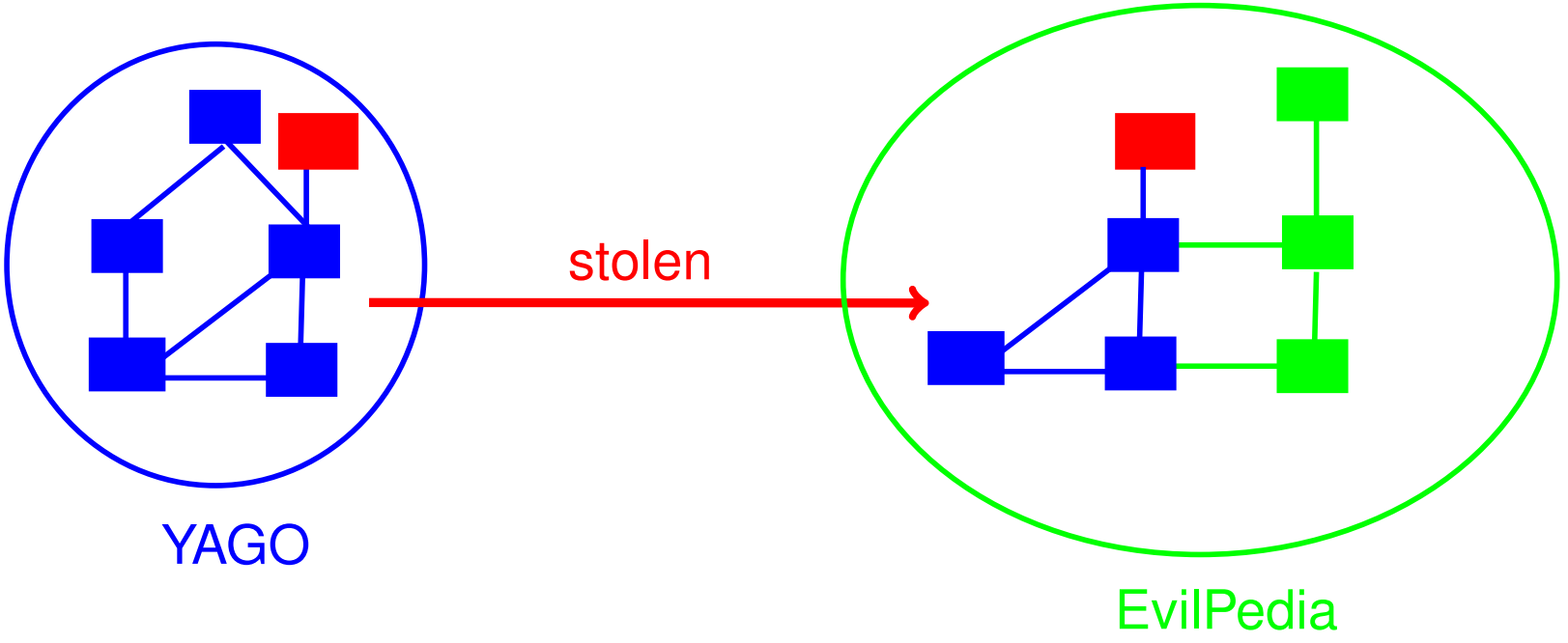
People may “steal” from other ontologies without giving due credit.

Most ontologies have licenses that require attribution.



Additive Watermarking

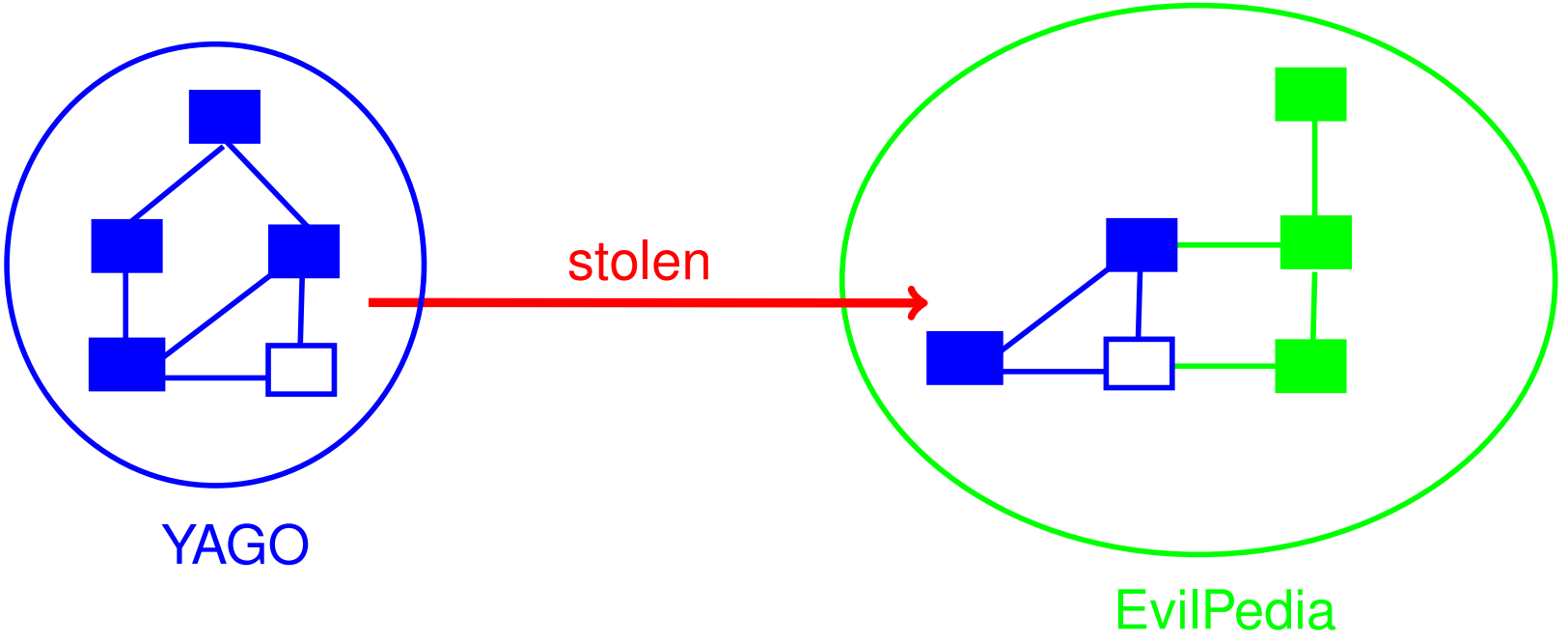
By adding a few fake facts to the source ontology,
one can prove theft in the target ontology.



[WWW 2012 demo]

Subtractive Watermarking

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



[ISWC 2011]

Protecting the user

Mat Honan

stories>88

©Ethan Hill CC by-sa VectorPortal

Protecting the user

add
credit
card

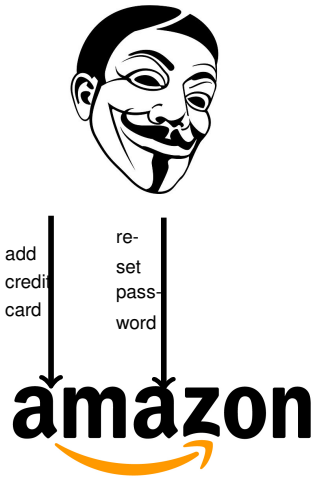
amazon

Mat Honan

©Ethan Hill CC by-sa VectorPortal

Protecting the user

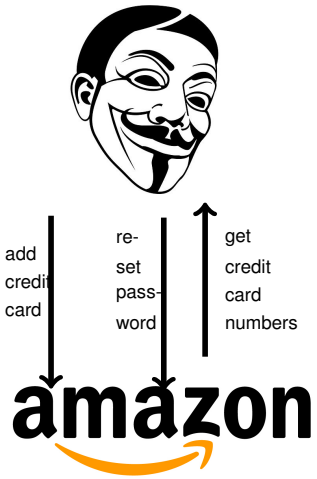
Mat Honan



©Ethan Hill CC by-sa VectorPortal

Protecting the user

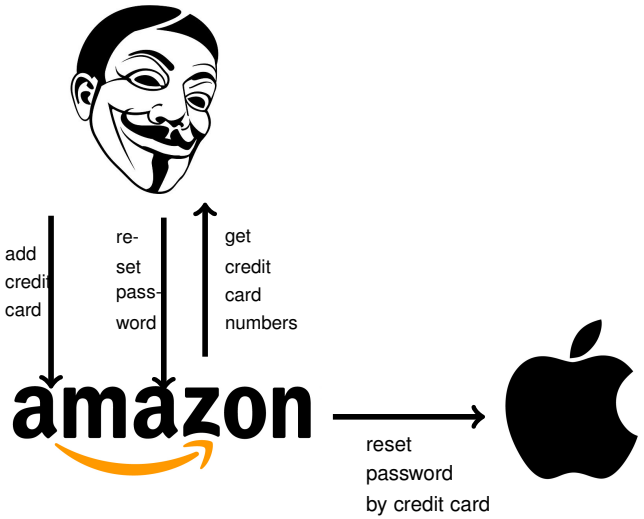
Mat Honan



©Ethan Hill CC by-sa VectorPortal

Protecting the user

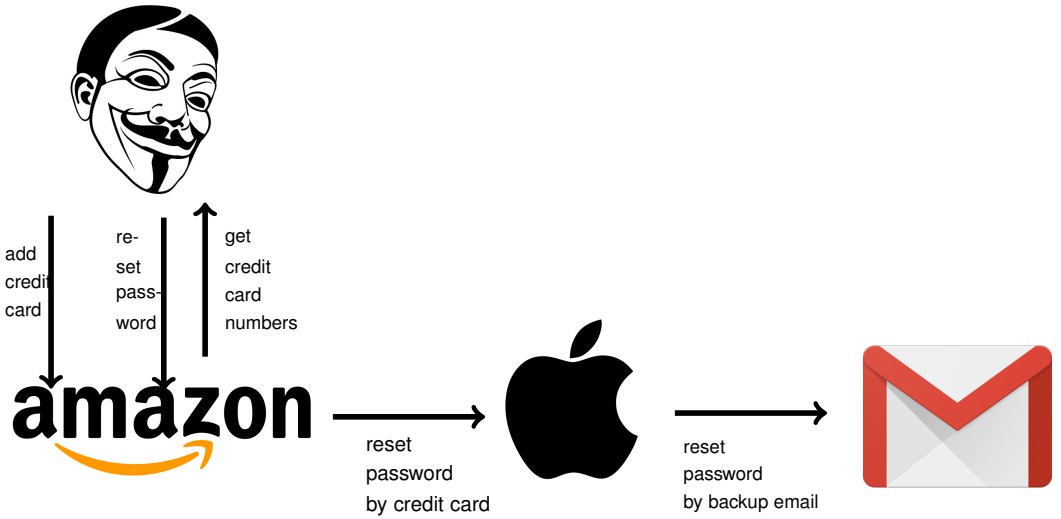
Mat Honan



©Ethan Hill CC by-sa VectorPortal

Protecting the user

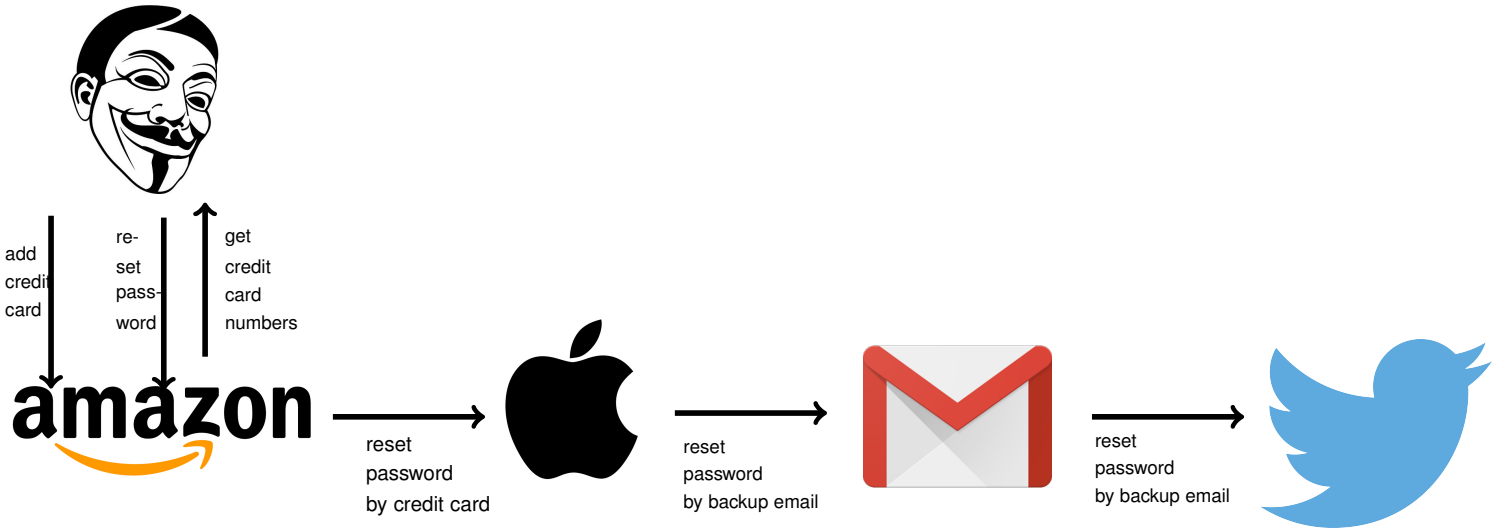
Mat Honan



©Ethan Hill CC by-sa VectorPortal

Protecting the user

Mat Honan

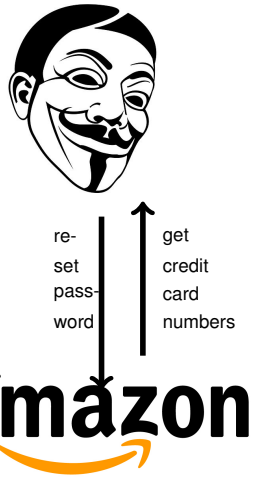


Read Mat's story on Wired

©Ethan Hill CC by-sa VectorPortal

Protecting the user

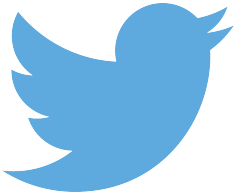
Mat Honan



reset password by credit card

reset password by backup email

reset password by backup email



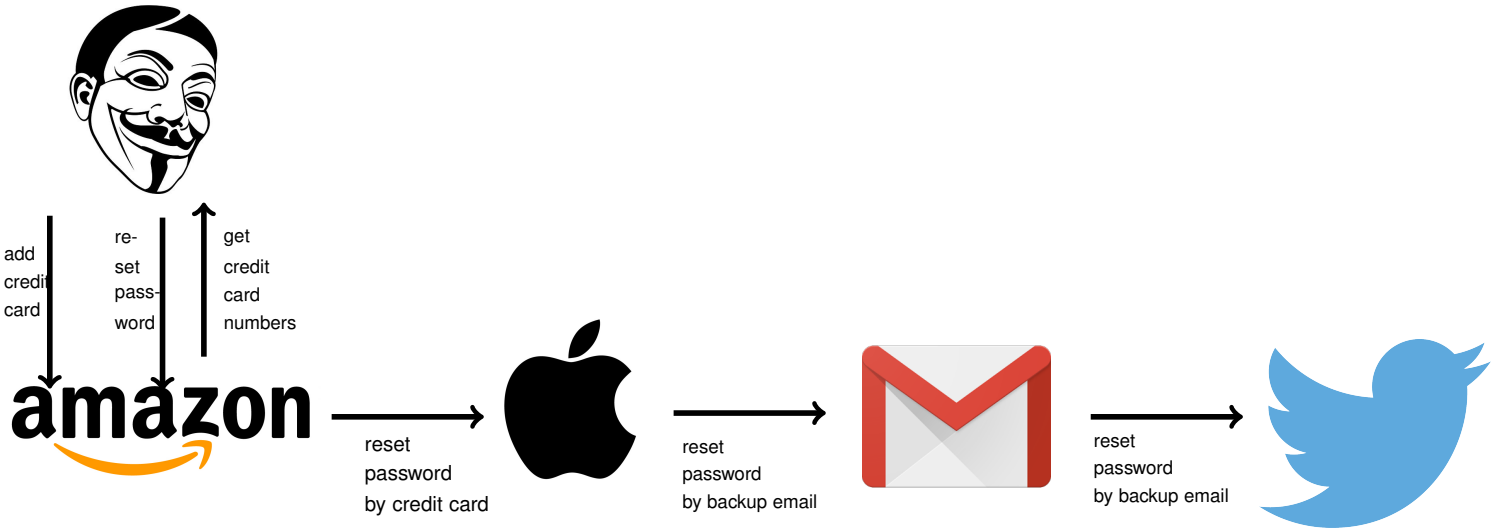
Owen Williams

©Ethan Hill CC by-SA VectorPortal

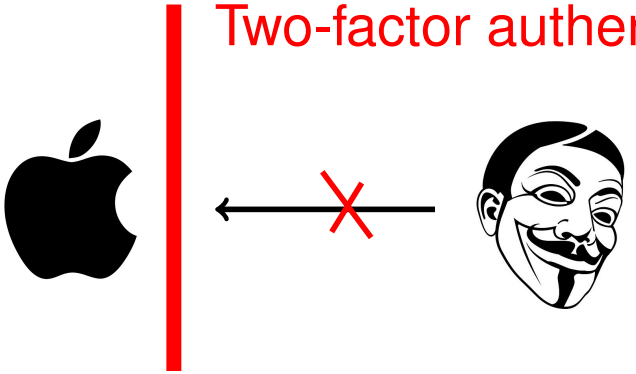
Read Mat's story on Wired

Protecting the user

Mat Honan



Owen Williams



©Ethan Hill CC by-SA VectorPortal

Read Mat's story on Wired

Protecting the user

Mat Honan

add credit card

re-set pass-word

get credit card numbers

amazon

reset password by credit card

reset password by backup email

reset password by backup email

Owen Williams

Two-factor authentication!

Protecting the user

Mat Honan

add credit card

re-set pass-word

get credit card numbers

amazon

reset password by credit card

reset password by backup email

reset password by backup email

Recovery codes!

Owen Williams

Two-factor authentication!

©Ethan Hill CC by-SA VectorPortal

Read Mat's story on Wired

Protecting the user

Mat Honan

add credit card

re-set pass-word

get credit card numbers

amazon

reset password by credit card

reset password by backup email

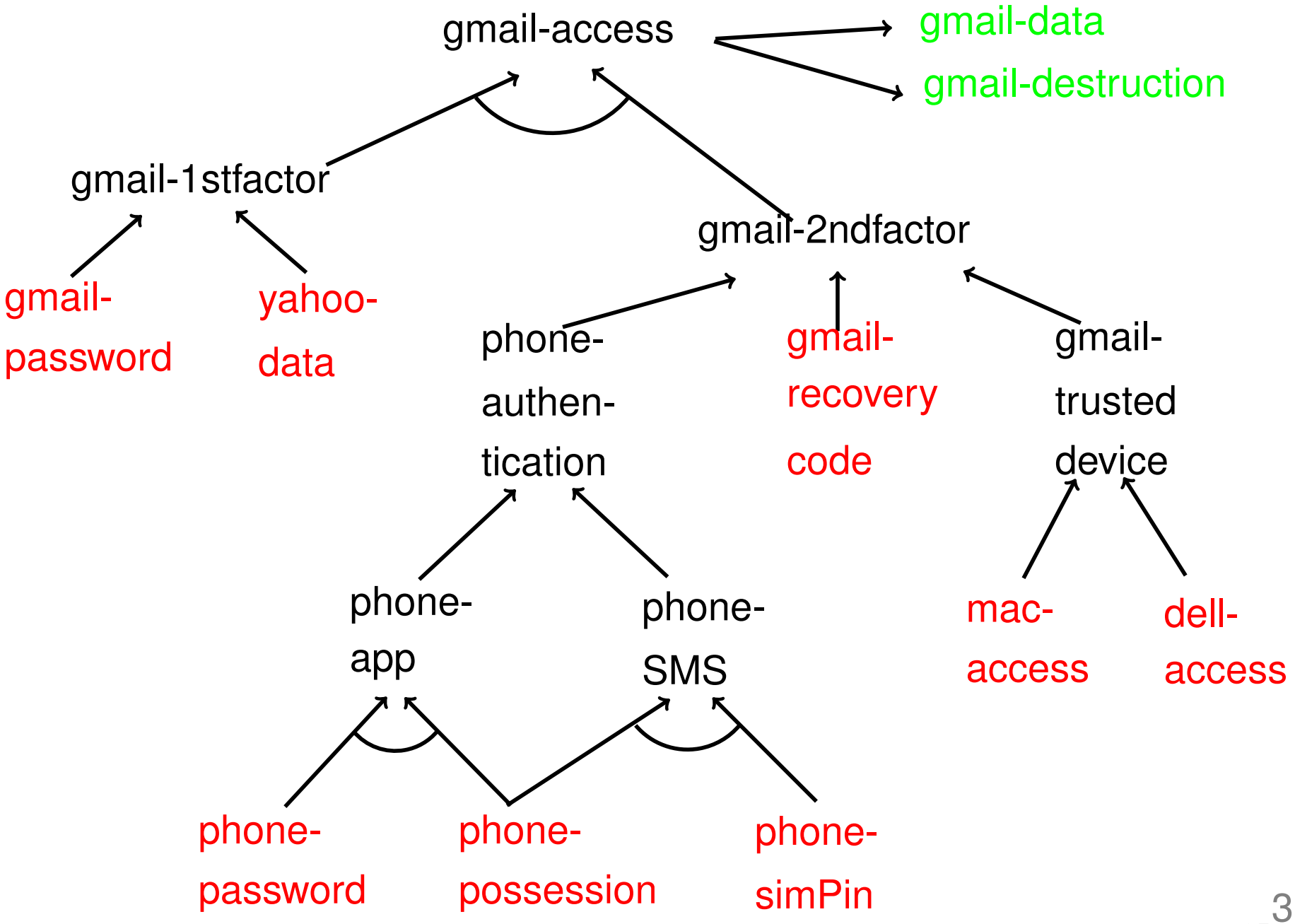
reset password by backup email

Recovery codes!

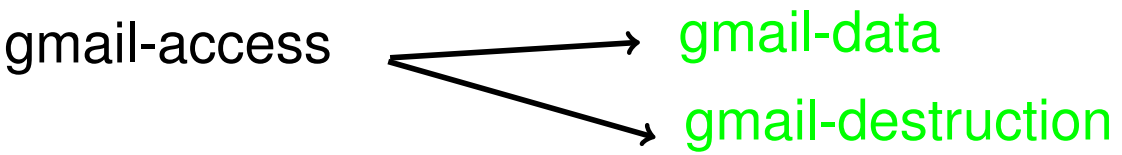
Owen Williams

Two-factor authentication!

Account Dependencies



Account Dependencies



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?

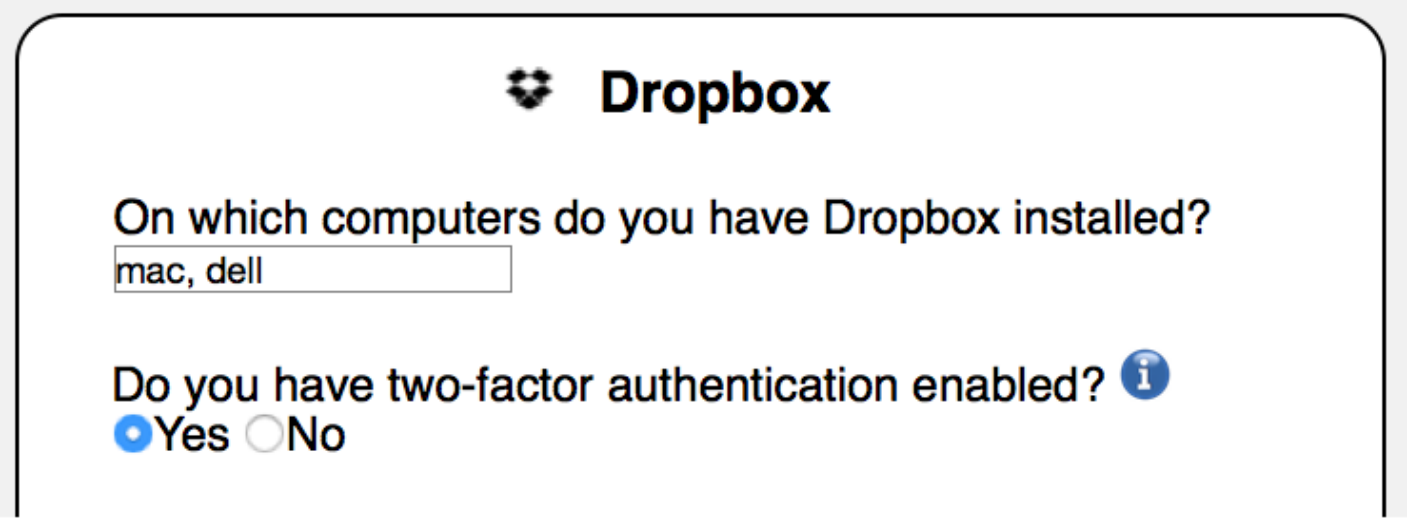
phone-password

phone-possession

phone-simPin

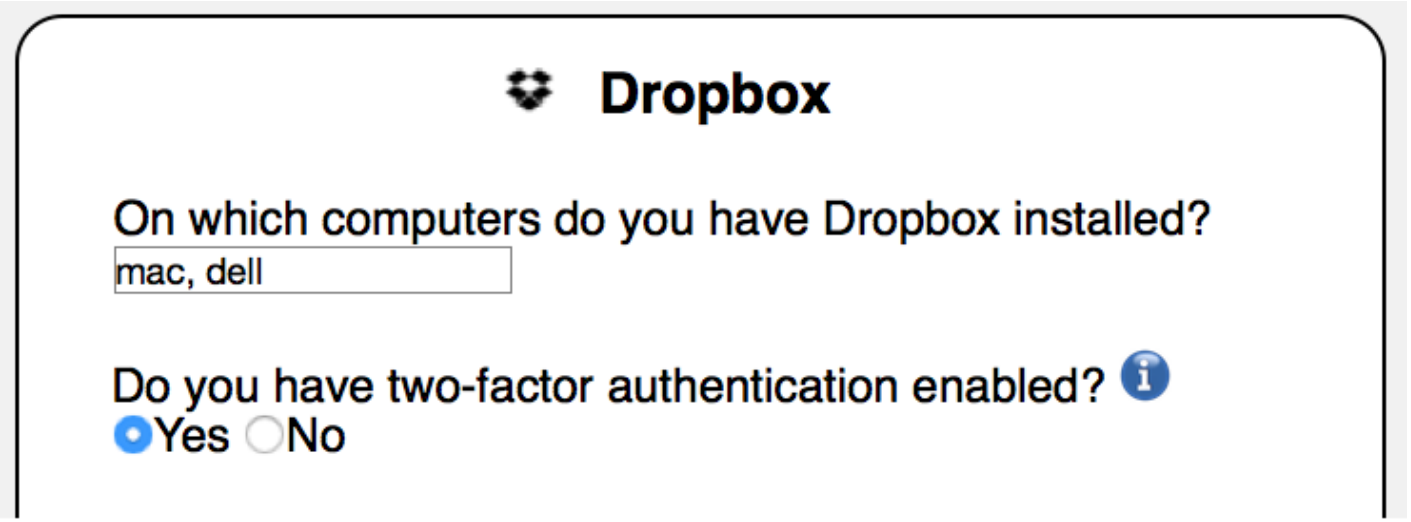
DIVINA

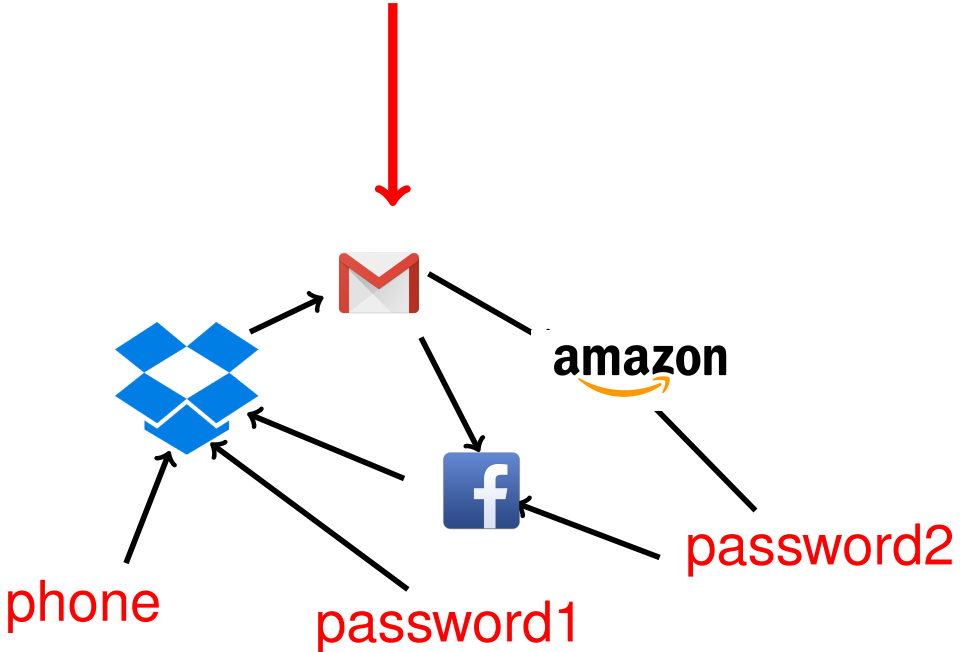
Discovering Vulnerabilities of Internet Accounts



DIVINA

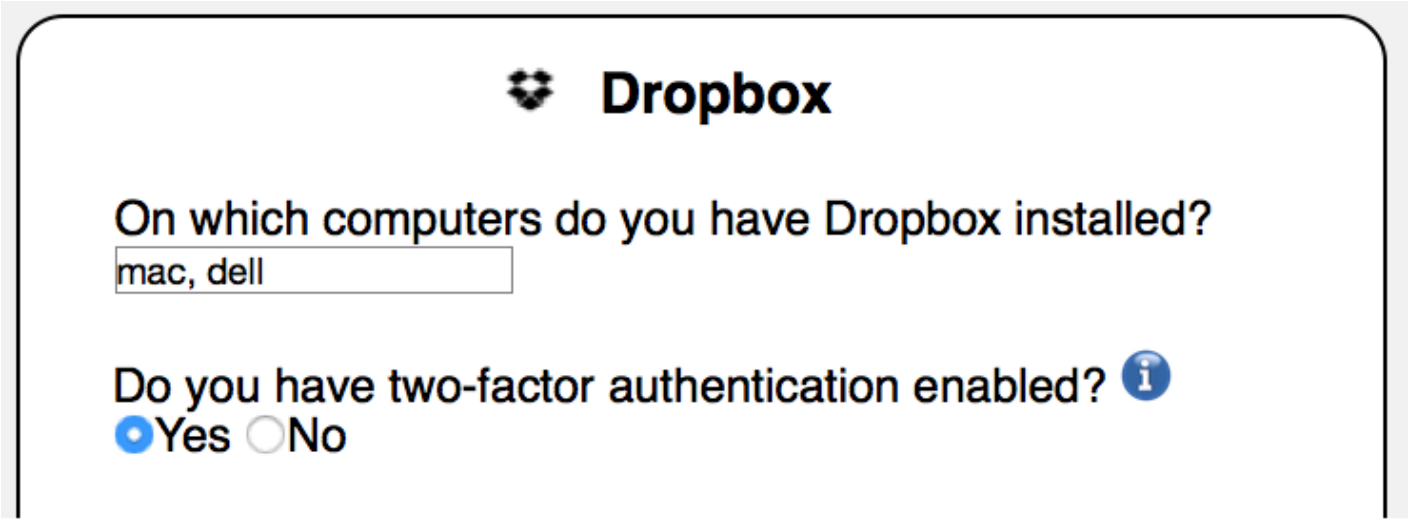
Discovering Vulnerabilities of Internet Accounts





DIVINA

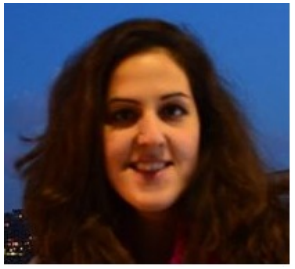
Discovering Vulnerabilities of Internet Accounts



	Safety	Security	Protection
	✓	✓	✗
	• gmail-password yahoo-access	• gmail-password phone-access	**For online accounts:** Consider two-factor authentication, lock codes on your phone, or encryption of your trusted devices.

DIVINA

Discovering Vulnerabilities of Internet Accounts

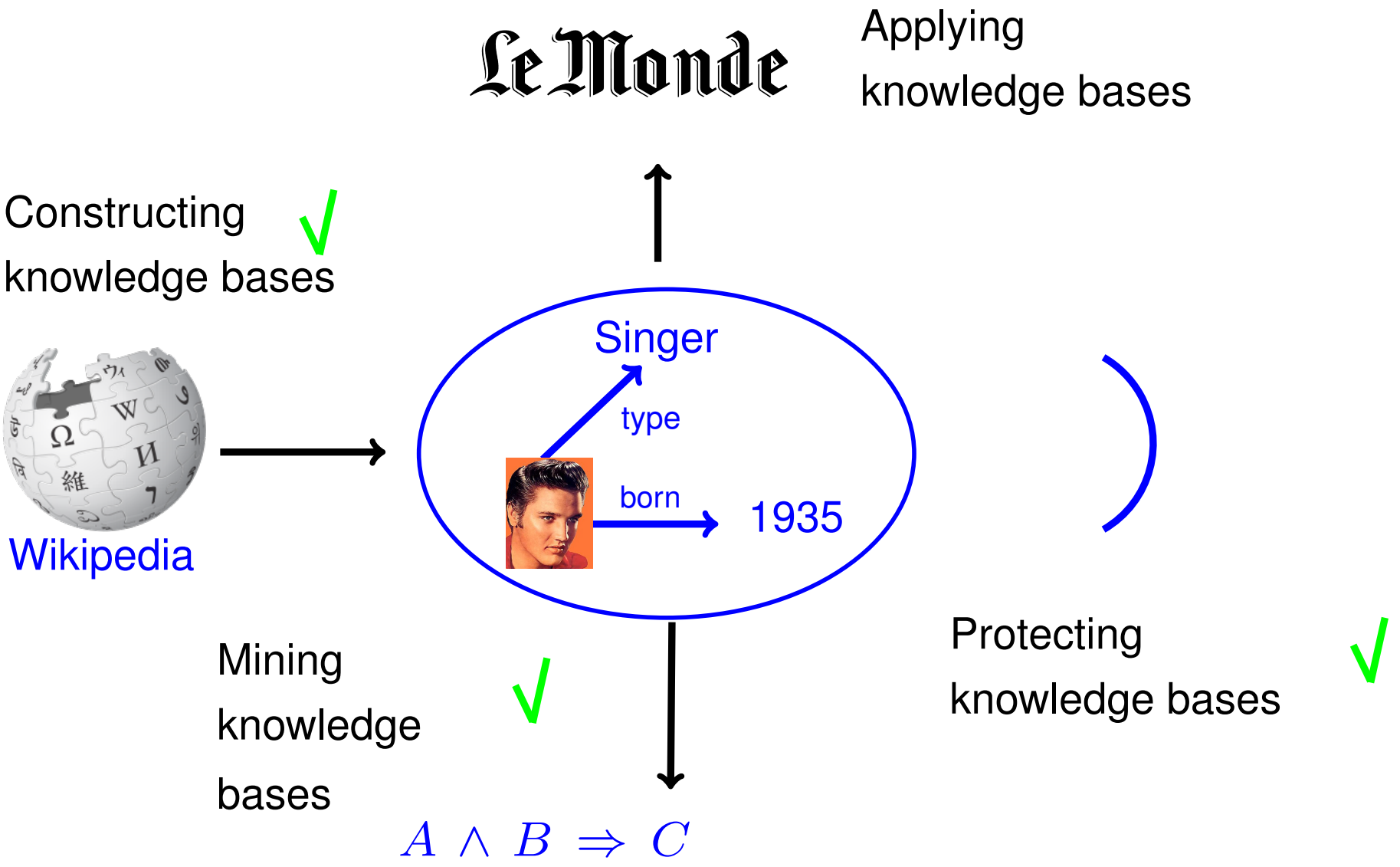


[WWW 2015 demo]

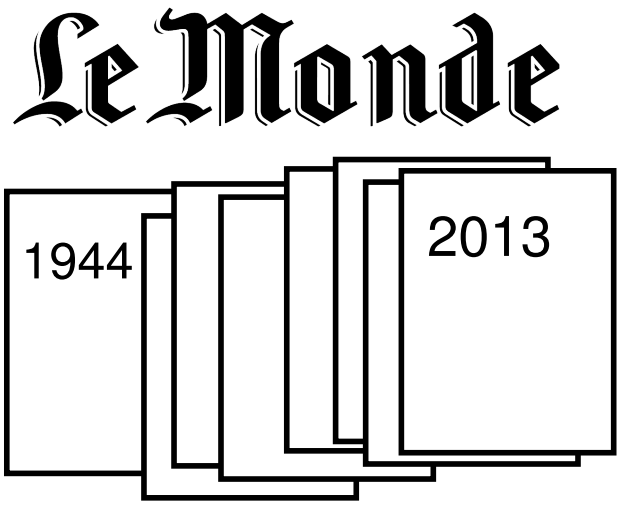
http://suchanek.name/programs/divina

	Safety	Security	Protection
	✓	✓	✗
	• gmail-password yahoo-access	• gmail-password phone-access	**For online accounts:** Consider two-factor authentication, lock codes on your phone, or encryption of your trusted devices.

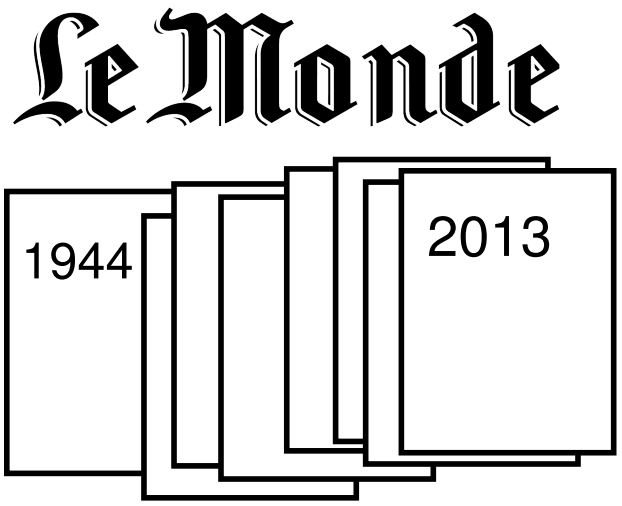
Knowledge Base Life Cycle



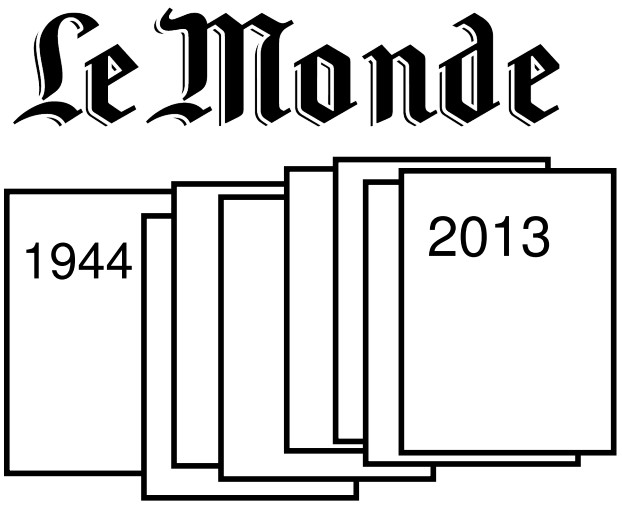
Mining Le Monde

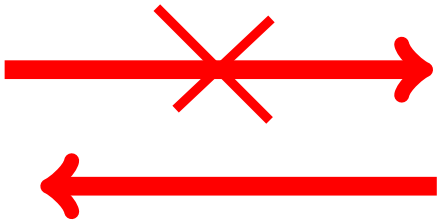


Mining Le Monde



Mining Le Monde





Mining Le Monde

Le Monde

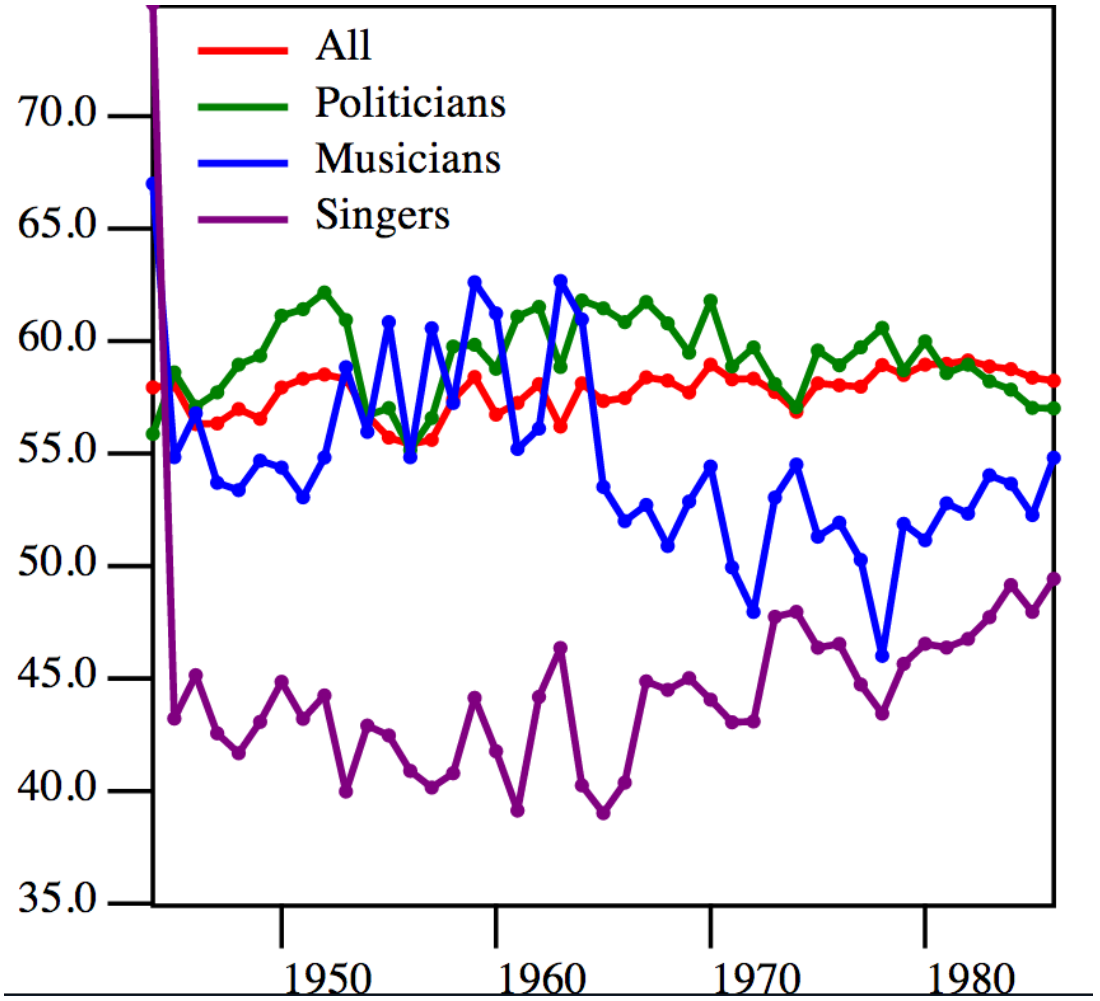
time	place	entity
1967	USA	

yago
select knowledge

Mining Le Monde

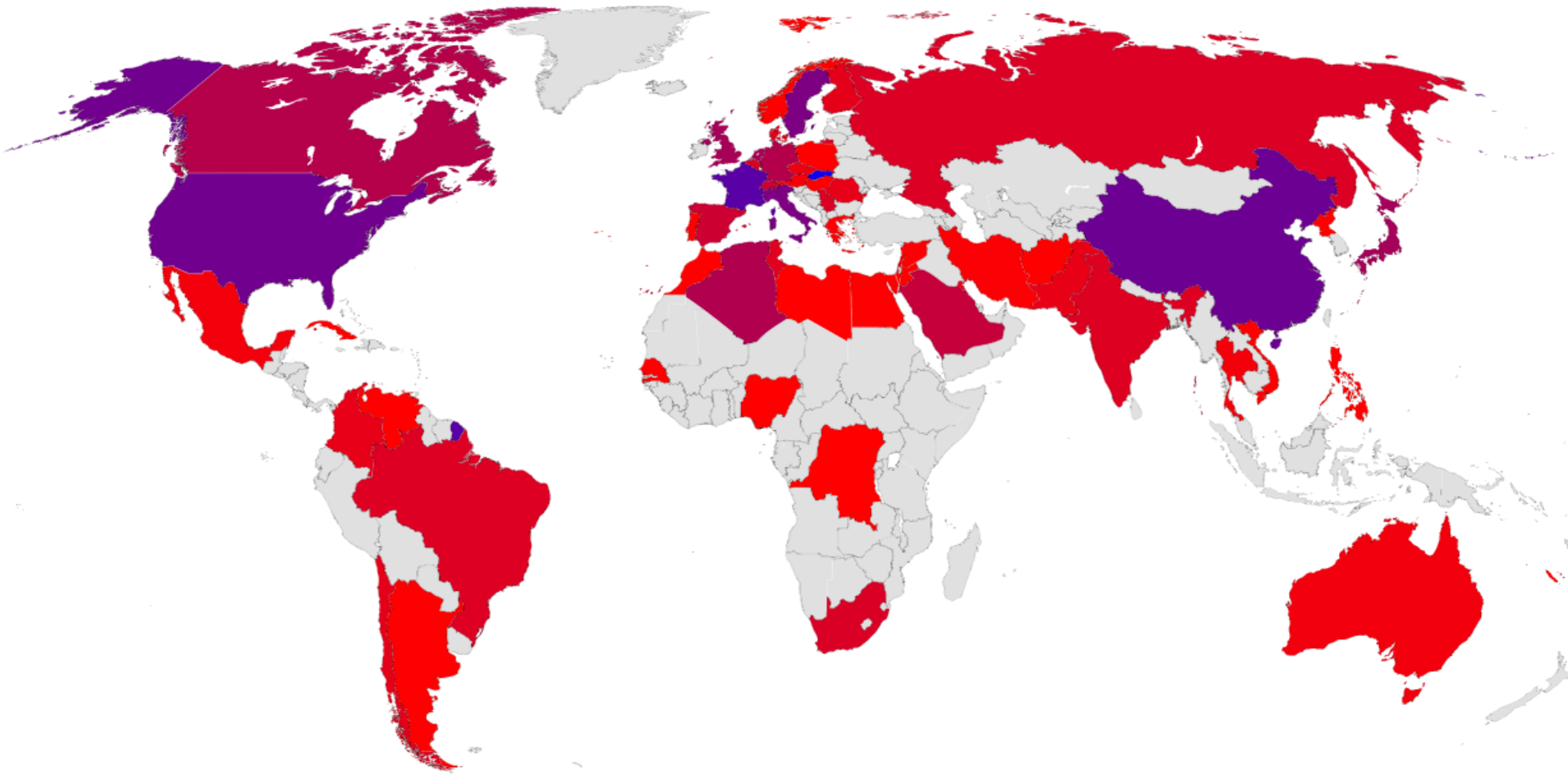
Le Monde

yago
select knowledge



Average age of people mentioned

Presence of foreign companies



red: many foreign companies mentioned

blue: few foreign companies mentioned

Mining Le Monde

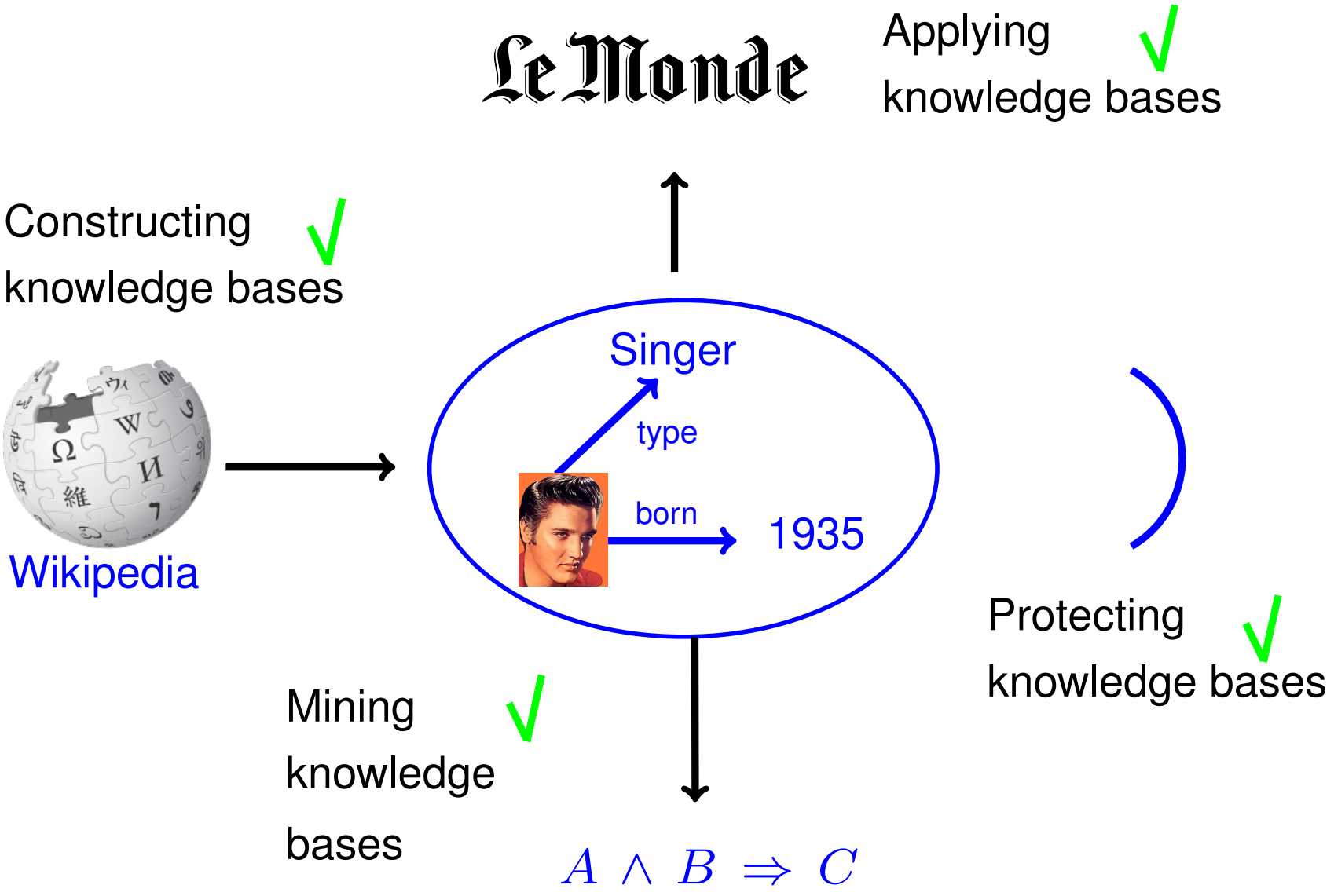
Le Monde

yago
select knowledge

[AKBC 2013]

[VLDB 2014 vision]

Knowledge Base Life Cycle



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

yago
select knowledge

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

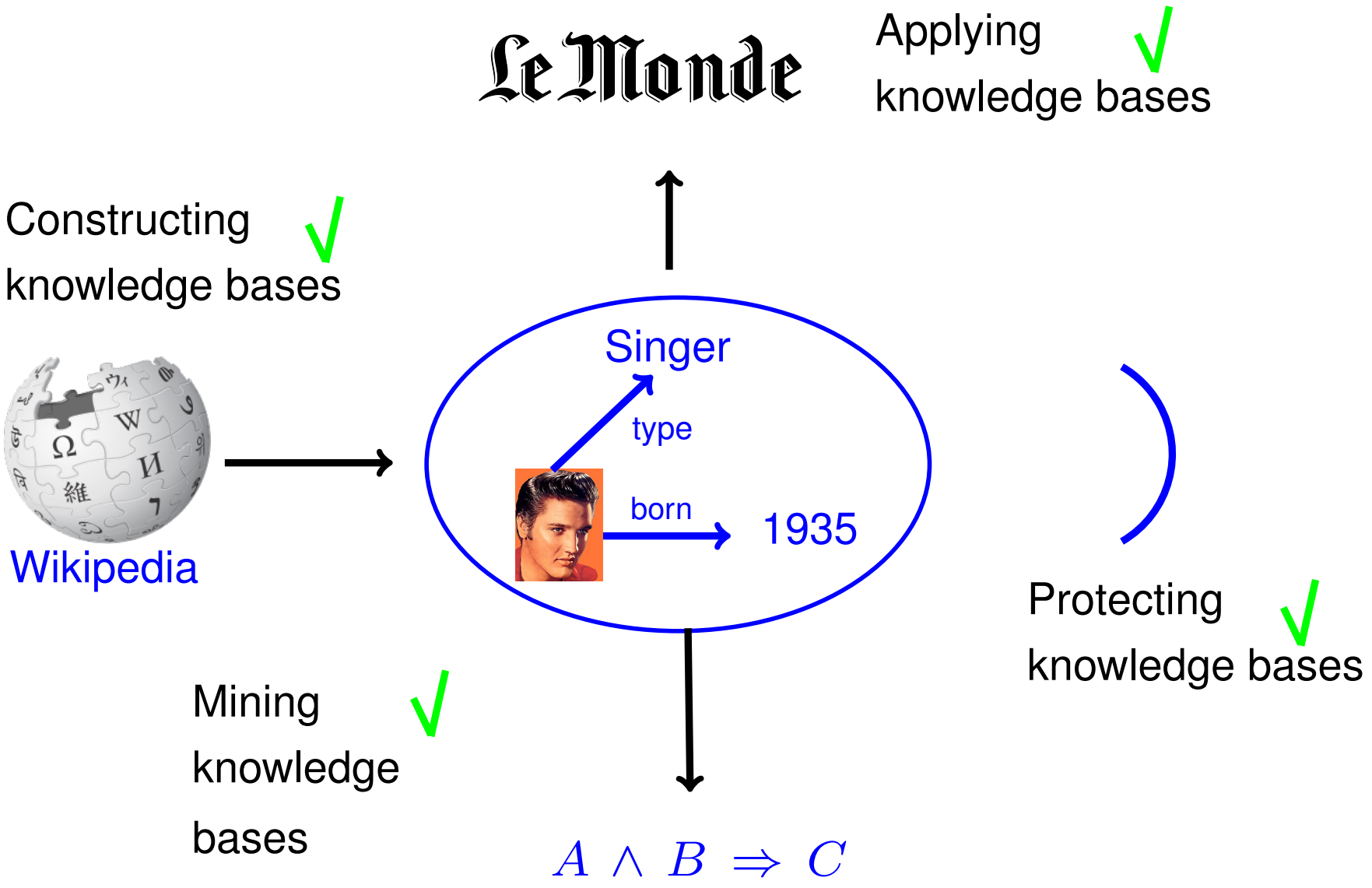
yago
select knowledge

100m statements

95% precision

-> 5m wrong statements

Thank you for your attention!



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