The Web Engineering 3

Introduction to Semantic Web

Lecture 2
Introduction to Semantic Web (Cont.)

Presented by
Prof. Khaled Wassif

Course Topics

➢ Introduction to the semantic web.

➢ Semantic web technologies and layered approach.

➢ Structured web documents in XML.

➢ Describing web resources in basic elements of Resource Description Framework (RDF).

➢ Web Ontology Language: OWL.

➢ Ontologies Applications.

Course References

1. Grigoris Antoniou, Paul Groth, Frank van Harmelen, Rinke Hoekstra, "A Semantic Web Primer", 2012.

2. John Domingue, Dieter Fensel, James A. Hendler, "Introduction to the Semantic Web Technologies", 2011.

Lecture 2 Outlines

➢ Review: what is the semantic web?

➢ The Semantic Web impact in B2C E-Commerce.

➢ The Semantic Web impact in Wikis.

➢ Design decisions for the Semantic Web.

➢ Example of How Semantic Web is working.

What is the Semantic Web? (Review)

➢ The Semantic Web is an extension of the World Wide Web through standards set by the World Wide Web Consortium (W3C).

 – The World Wide Web Consortium (W3C): is an international community where member organizations is a full-time staff, and the public work together to develop Web standards.

➢ The goals of the Semantic Web:

 1 – Is to make internet data content machine-readable and detailed accessible.

 2 – Use intelligent techniques to take advantage of these representations.

 3 – Gradually evolve out the existing Web by the Semantic Web, it is not a race to the current WWW

Lecture 2 Outlines

➢ Review: what is the semantic web?

➢ The Semantic Web impact in B2C E-Commerce.

➢ The Semantic Web impact in Wikis.

➢ Design decisions for the Semantic Web.

➢ Example of How Semantic Web is working.

Example: The Semantic Web impact in B2C E-Commmerce

➢ B2C (Business-to-Consumer e-Commerce), also called retail e-commerce, is a business model that involves sales between online businesses and consumers. Ex. Jumia and Amazon

➢ A typical scenario: user visits one or several online shops, browses their offers, selects and orders products.

➢ Ideally humans would visit all, or all major online stores; but too time consuming

➢ **Shopbots** are a useful tool

– Shopbots are a software program or script designed to search on the internet or other forms of data and obtain pricing information on products and services.

Limitations of Shopbots Tools

القرآن

➢ They based on wrappers (covers): required extensive and complex programming. (Covering the retrieved data in a specific interface shape to the user)

  ❖ Wrapper is any entity that encapsulates (wraps around) another data item.

  ❖ Wrappers are used for two primary purposes: to convert data to a suitable format or to hide the complexity of the underlying entity using abstraction.

➢ Wrappers extract information based on textual (Keyword
➢ based) analysis.

  ❖ Error-prone

  ❖ Limited information extracted

➢ Shopbots don't yet "understand" the meaning of information that they retrieve.
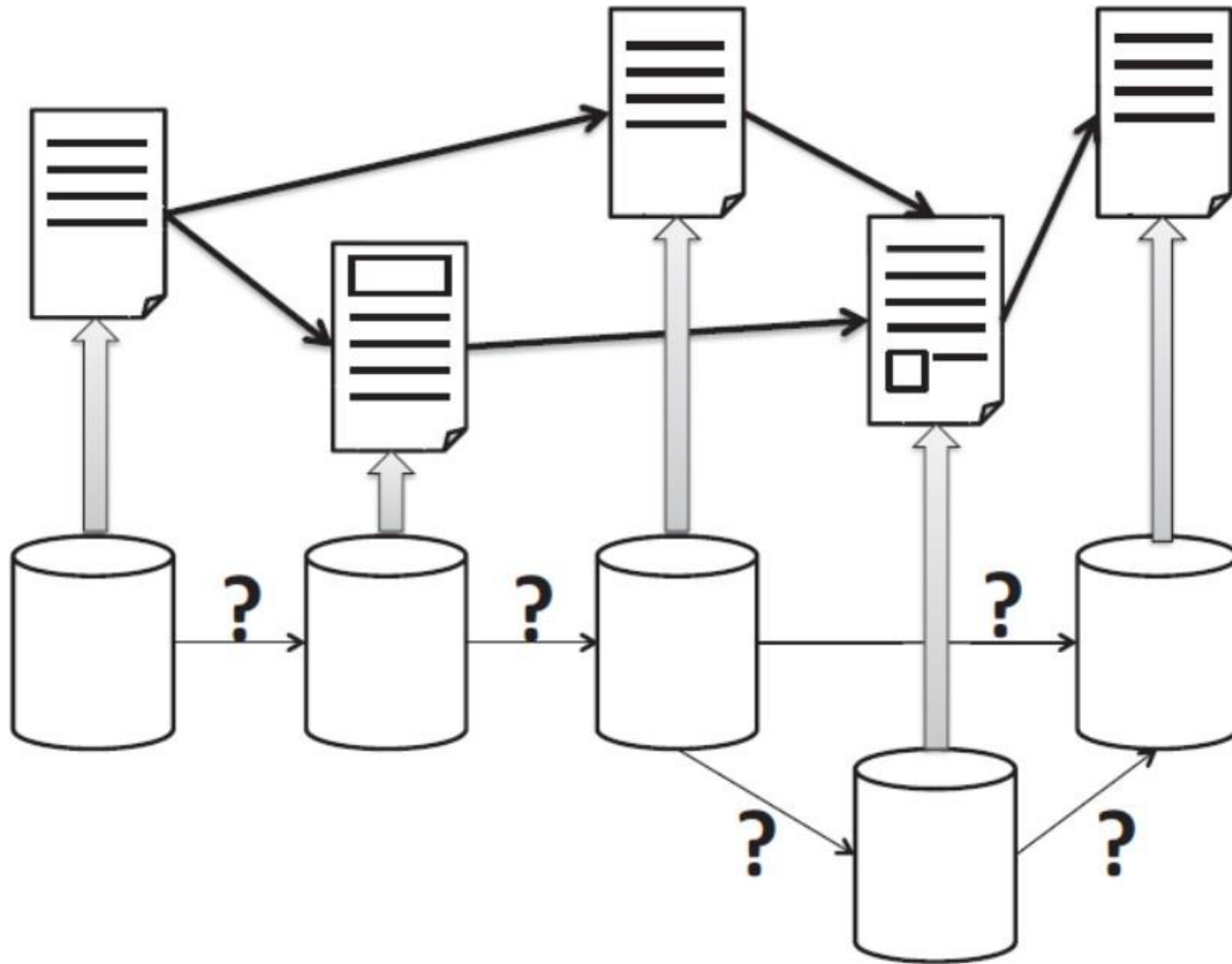
8

Lecture 2 Outlines

➢ Review: what is the semantic web?

➢ The Semantic Web impact in B2C E-Commerce.

➢ The Semantic Web impact in Wikis.

➢ Design decisions for the Semantic Web.

➢ Example of How Semantic Web is working.

Wikis as a Different Shape of Data Collaboration

Properties of Wikis

➢ Collections of web pages that allow users to add content via a browser interface.

➢ Wiki systems allow for collaborative knowledge.

➢ Users are free to add and change information without ownership of content, access restrictions, or hard workflows

Some Uses of Wikis

➤ Development of knowledge in a human community effort, with contributions from a wide range of users (e.g. Wikipedia).

➤ Knowledge management of an activity or a project (e.g. brainstorming and exchanging ideas, coordinating activities, exchanging records of meetings).

The Semantic Web Impact in Wikis

Semantic Web impacts more strength for Wikis:

➢ The Wiki tools method:

- ❖ the inherent structure of a wiki, given by data linking between pages, gets accessible to machines more meaningful than simple navigation.

➢ Semantic Web structured can impact by strong annotations referring to the real meaning of the knowledge captured by the wiki.

Lecture 2 Outlines

➢ Review: what is the semantic web?

➢ The Semantic Web impact in B2C E-Commerce.

➢ The Semantic Web impact in Wikis.

➢ Design decisions for the Semantic Web.

➢ Example of How Semantic Web is working.

Design decisions for Semantic Web

The Semantic Web uses different design principles (steps):

1. Make structured and semi-structured data available in standardized formats on the web.

2. Make datasets, the individual data-elements, and their relations accessible on the web.

3. Describe the semantics of this data in a formalism that allow machines to processing it.

Design decisions for Semantic Web

Structured and unstructured data on the web

Design decisions for Semantic Web

How to perform this processing?

❑ The key for performing the vision of Semantic Web is publish and interlink the underlying structured datasets (instead of just publishing and interlinking the HTML pages after much of the data basic structure has been lost).
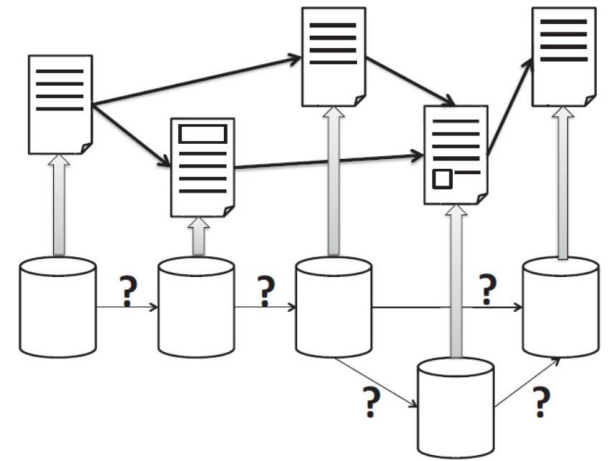
Lecture 2 Outlines

➢ Review: what is the semantic web?

➢ The Semantic Web impact in B2C E-Commerce.

➢ The Semantic Web impact in Wikis.

➢ Design decisions for the Semantic Web.

➢ Example of How Semantic Web is working.

A simplified bookstore data (Dataset "A")

ID	Author	Title	Publisher	Year
ISBN0-00-651409-X	id_xyz	The Glass Palace	id_qpr	2000

ID	Name	Home Page
id_xyz	Ghosh, Amitav	http://www.amitavghosh.com

ID	Publ. Name	City
id_qpr	Harper Collins	London

1st: export your data as a set of relations
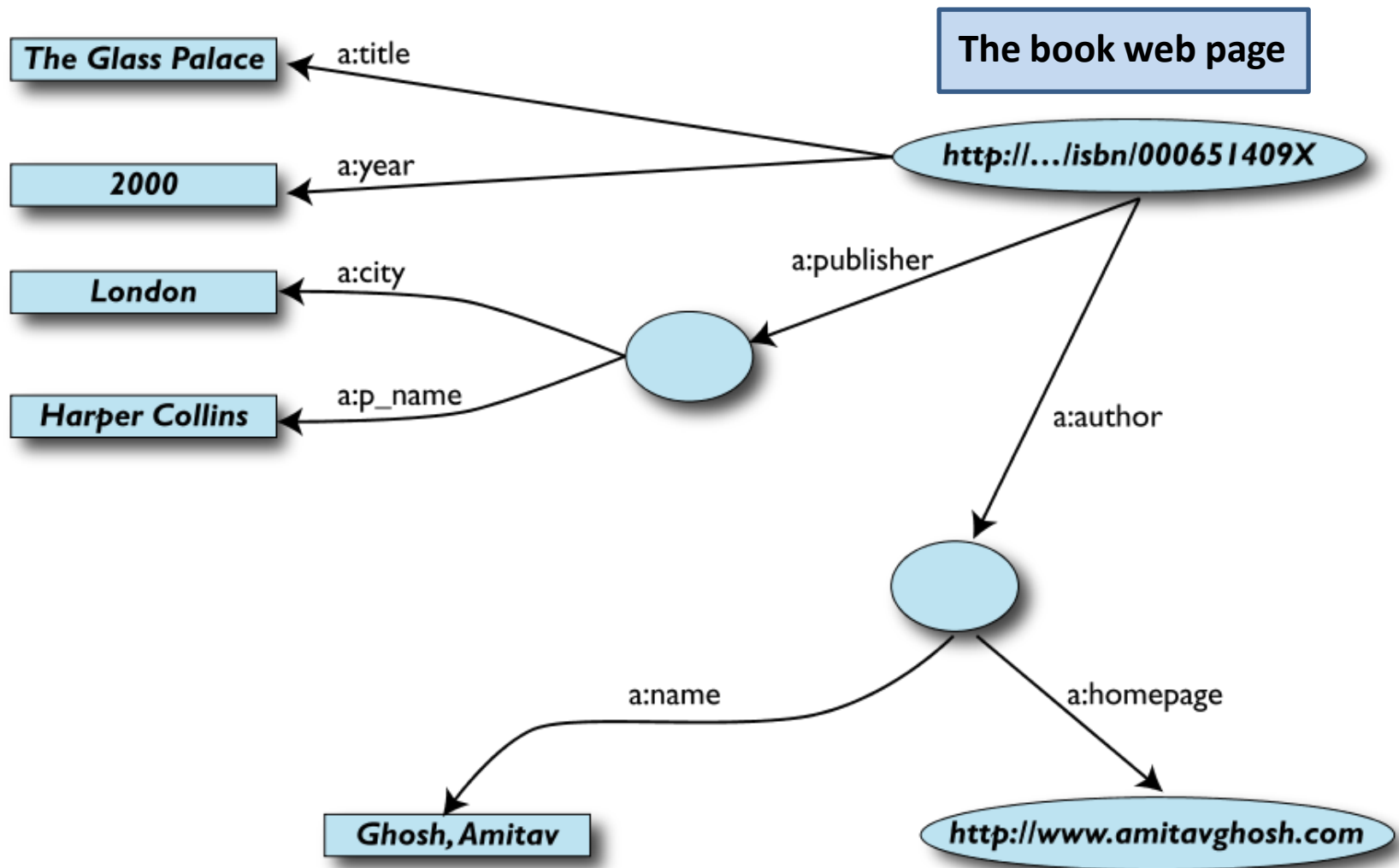
Database A

Some notes on exporting the data

❑ Relations form a graph

  ➢ The nodes (each one) refer to the "real" data.

  ➢ How the graph (relationships) is represented in machine in semantic web.

❑ Data export does not mean physical conversion of the data

  ➢ relations can be generated online at query time via SQL "bridges" for example.

  ➢ extracting HTML pages.

  ➢ extracting data from Excel sheets.

  ➢ etc.

Another bookstore data (Dataset "B")

❑ Another Structure with another language "French"

	A	B	D	E
1	**ID**	**Titre**	**Traducteur**	**Original**
2	ISBN0 2020386682	Le Palais des miroirs	A13	ISBN-0-00-651409-X
3				
6	**ID**	**Auteur**		
7	ISBN-0-00-651409-X	A12		
11	**Nom**			
12	Ghosh, Amitav			
13	Besse, Christianne			

The Translator

The Translating Version ISBN

2ⁿᵈ: export your second set of data
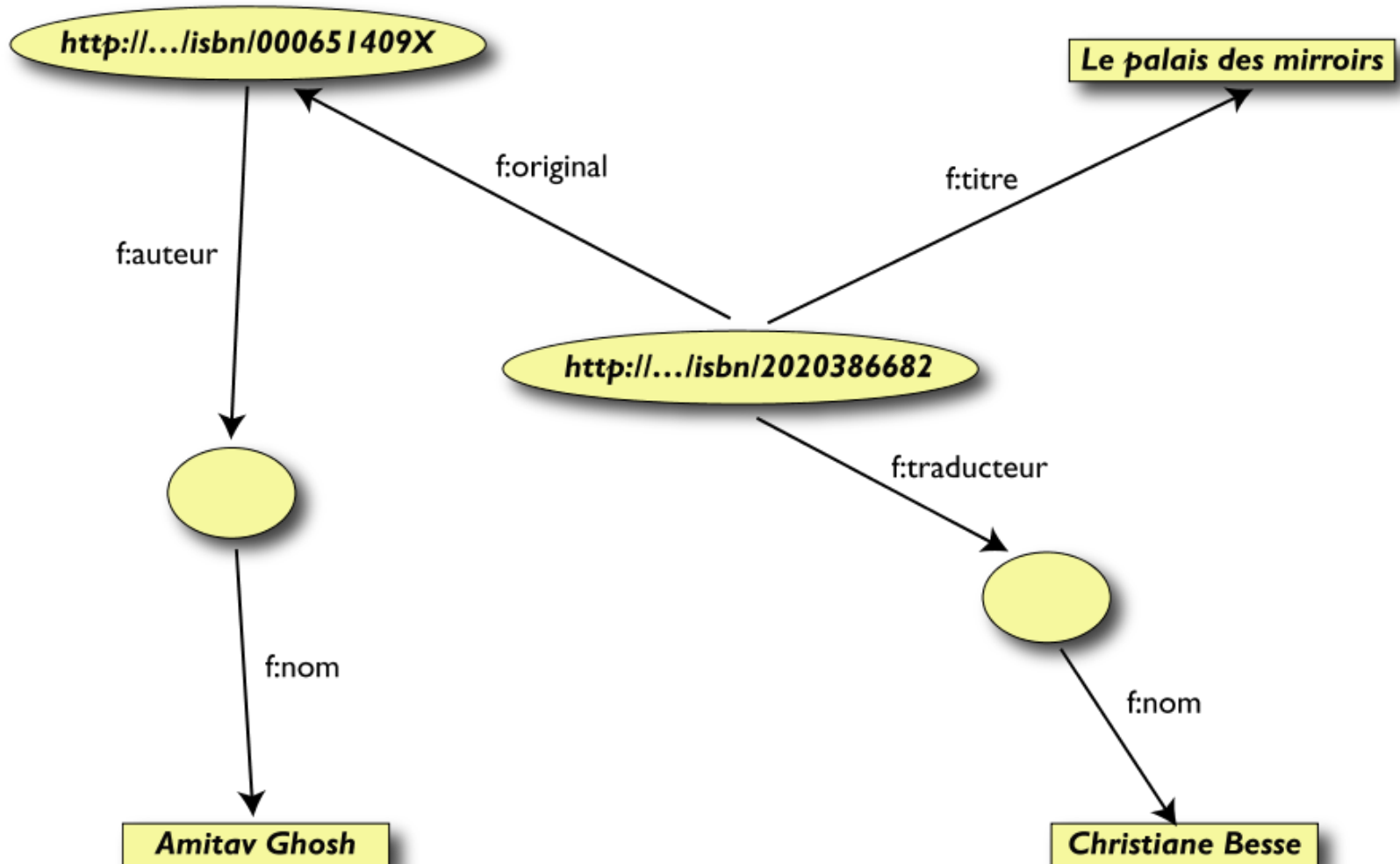
Database B
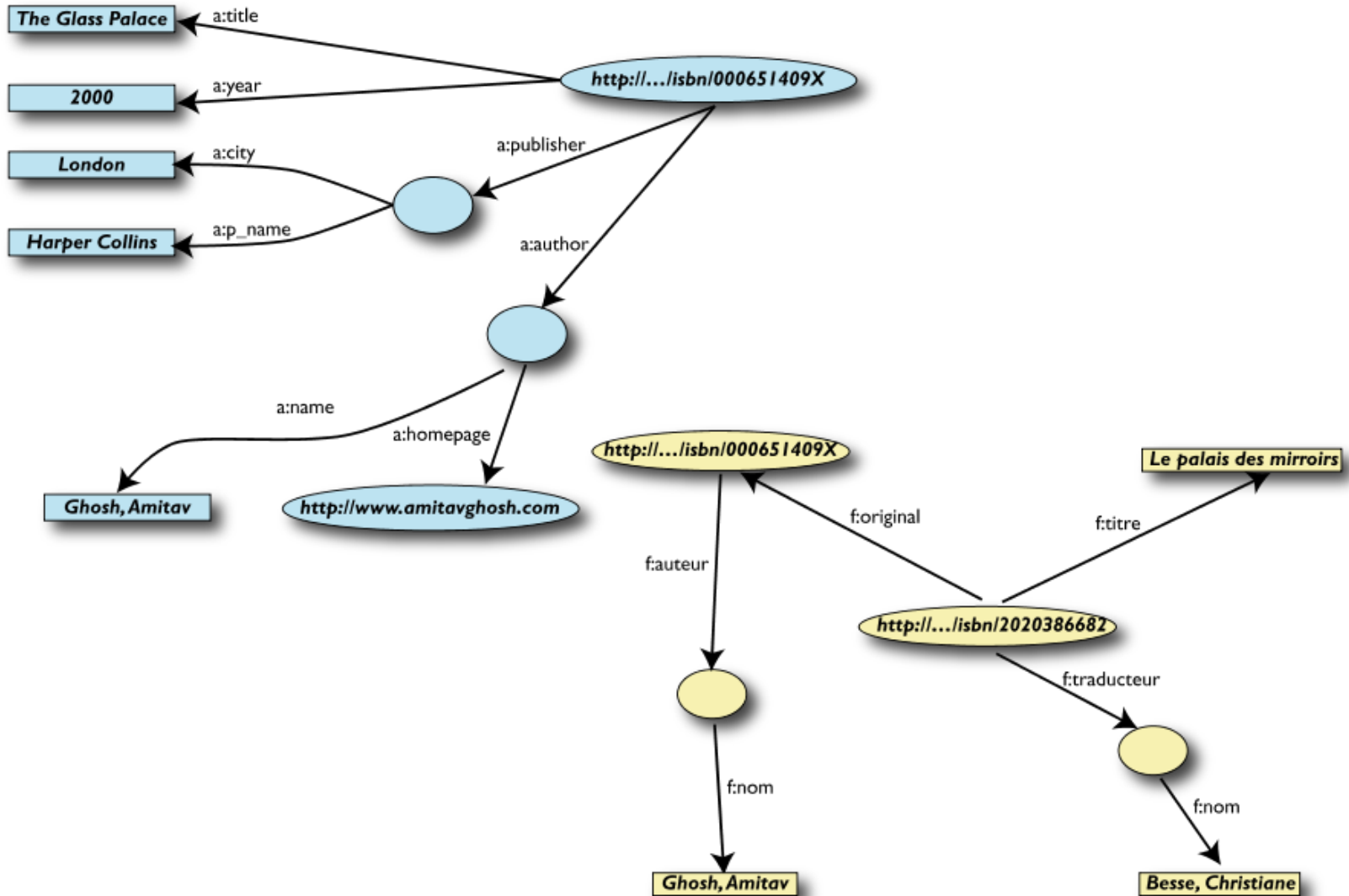
Start merging your data

Start merging your data

Same URI = Same Resources
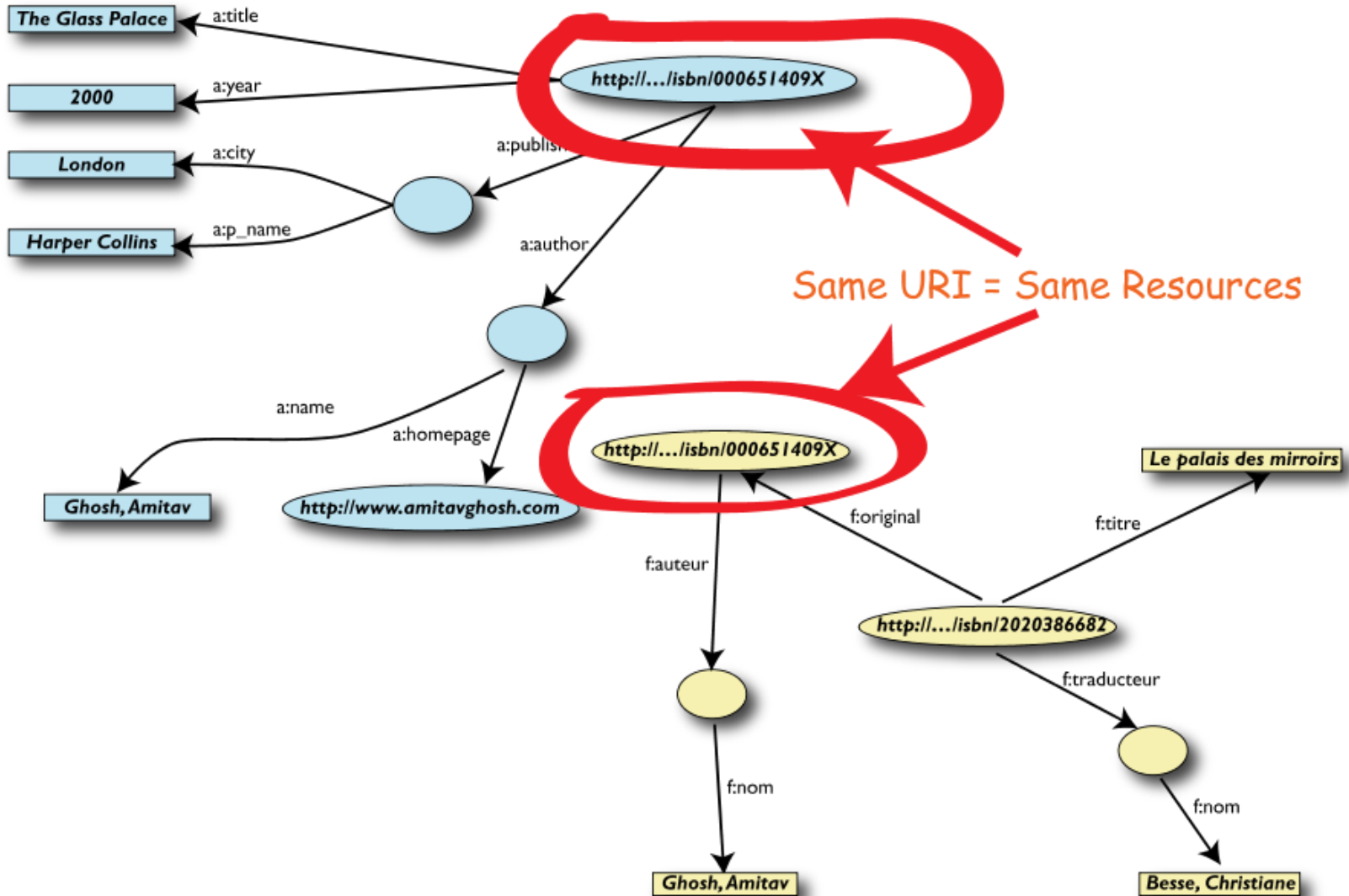
Start merging your data
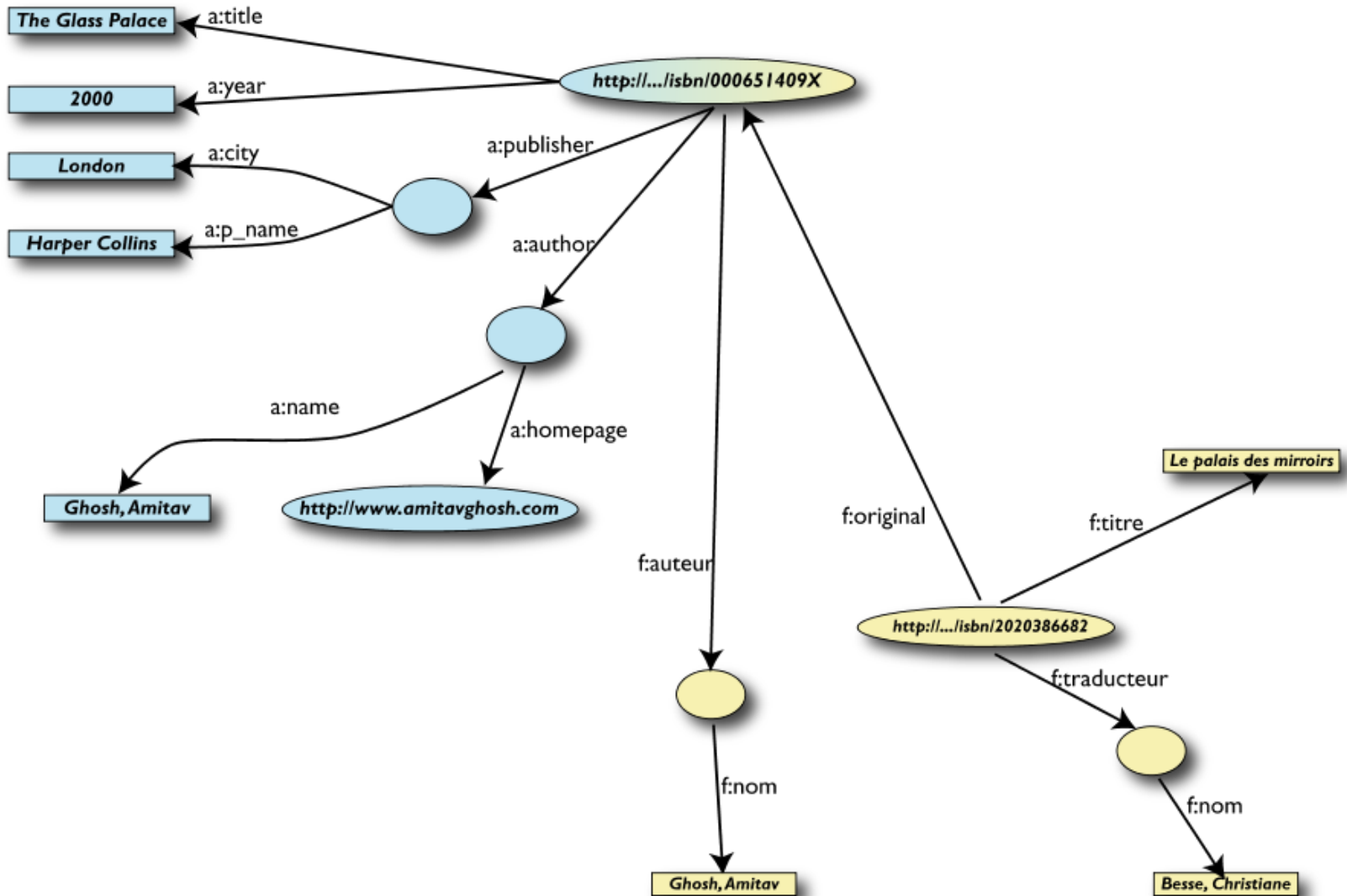
Start making queries…

❑ User of dataset "B" can now ask queries like:

  ➢ « give me the title of the original » → in English and by a way.

  ➢ « donnes-moi le titre de l'original » → in French and by another way.

❑ This information is not in the dataset "B"…

❑ …but can be retrieved by merging with dataset "A"!
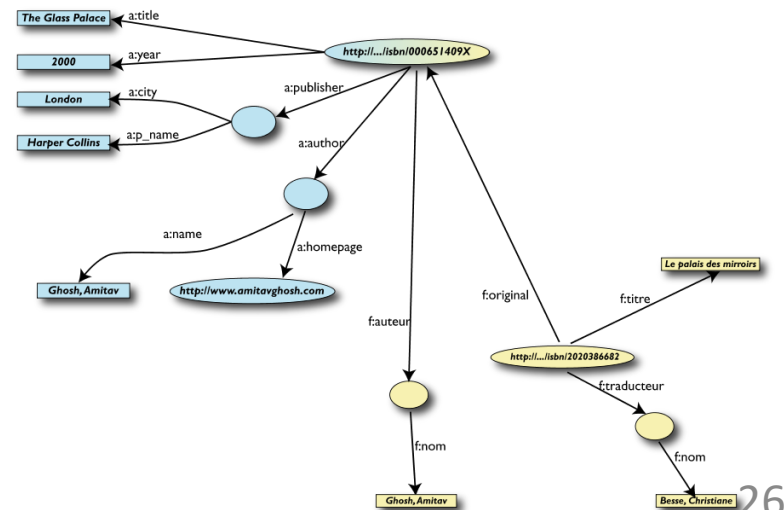
However, more can be achieved…

❑ We "feel" that a:<u>author</u> and b:auteur (translator) should be the same.

❑ But an automatic merge does not know that!

❑ Let us add some extra information to the merged data:

➢ **`a:author`** same as **`f:auteur`**

➢ both identify a "<u>Person</u>"

➢ a term that a community may have already defined:

➢ a "Person" is uniquely identified by his/her name and, say, homepage

➢ it can be used as a "<u>category</u>" for certain type of resources

Obtain the extra knowledge
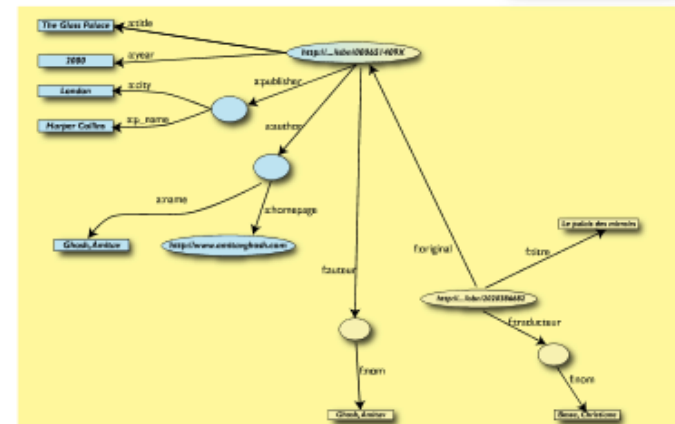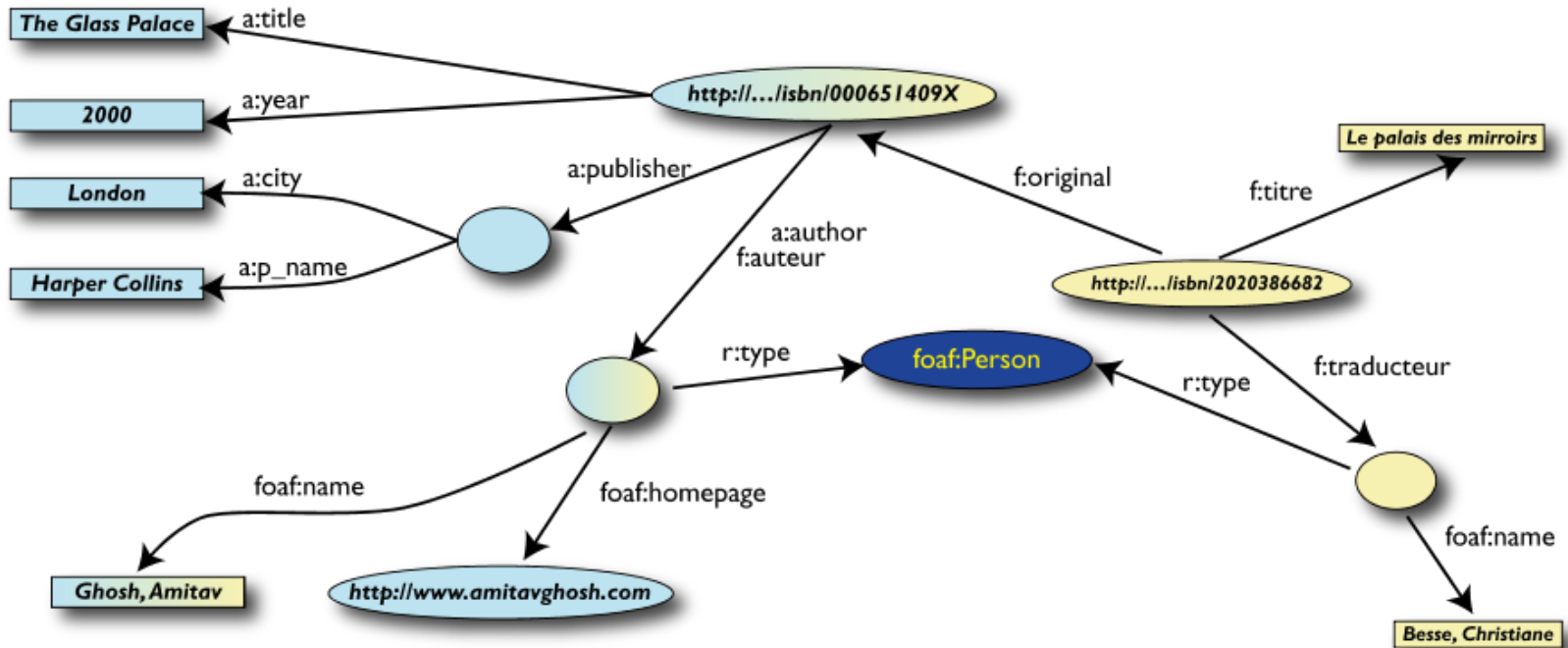
Start making richer queries!

❑ User of dataset "B" can now query:

➢ "donnes-moi la page d'accueil de l'auteur de l'originale"
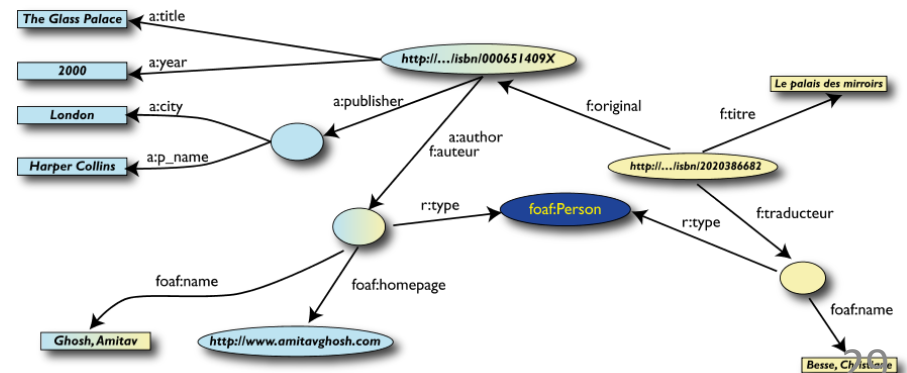
OR

➢ "give me the home page of the original's **'auteur'**

❑ The information is not in datasets "B" or "A"…

❑ …but was made available by:

➢ merging datasets "A" and datasets "B"

➢ adding three simple extra statements as an extra elements
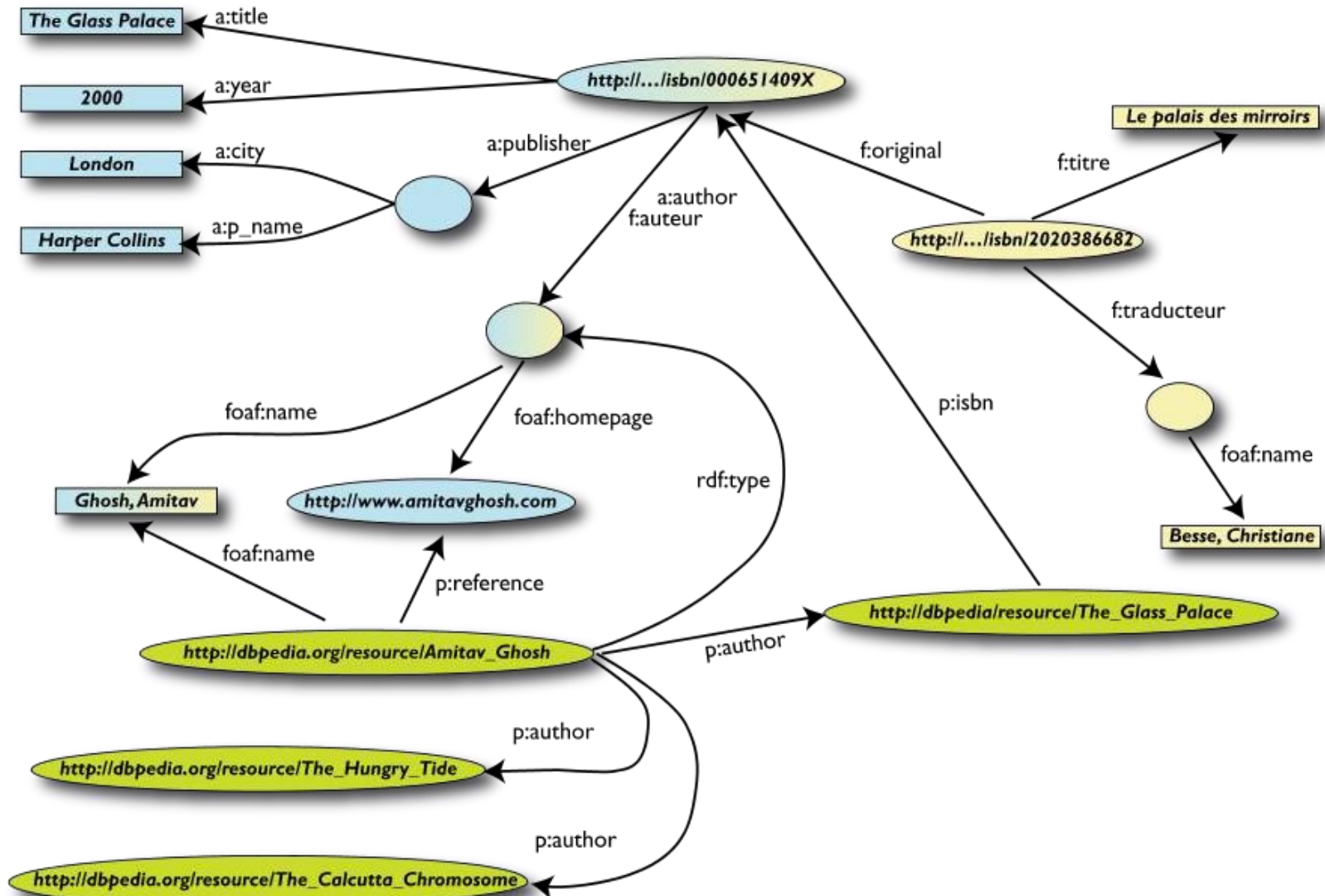
Combine with different datasets

❑ Different dataset can be combined with other sources.

❑ For example, data in Wikipedia can be extracted using dedicated tools.

➢ e.g., the "dbpedia" project can extract the "infobox" information from Wikipedia already…
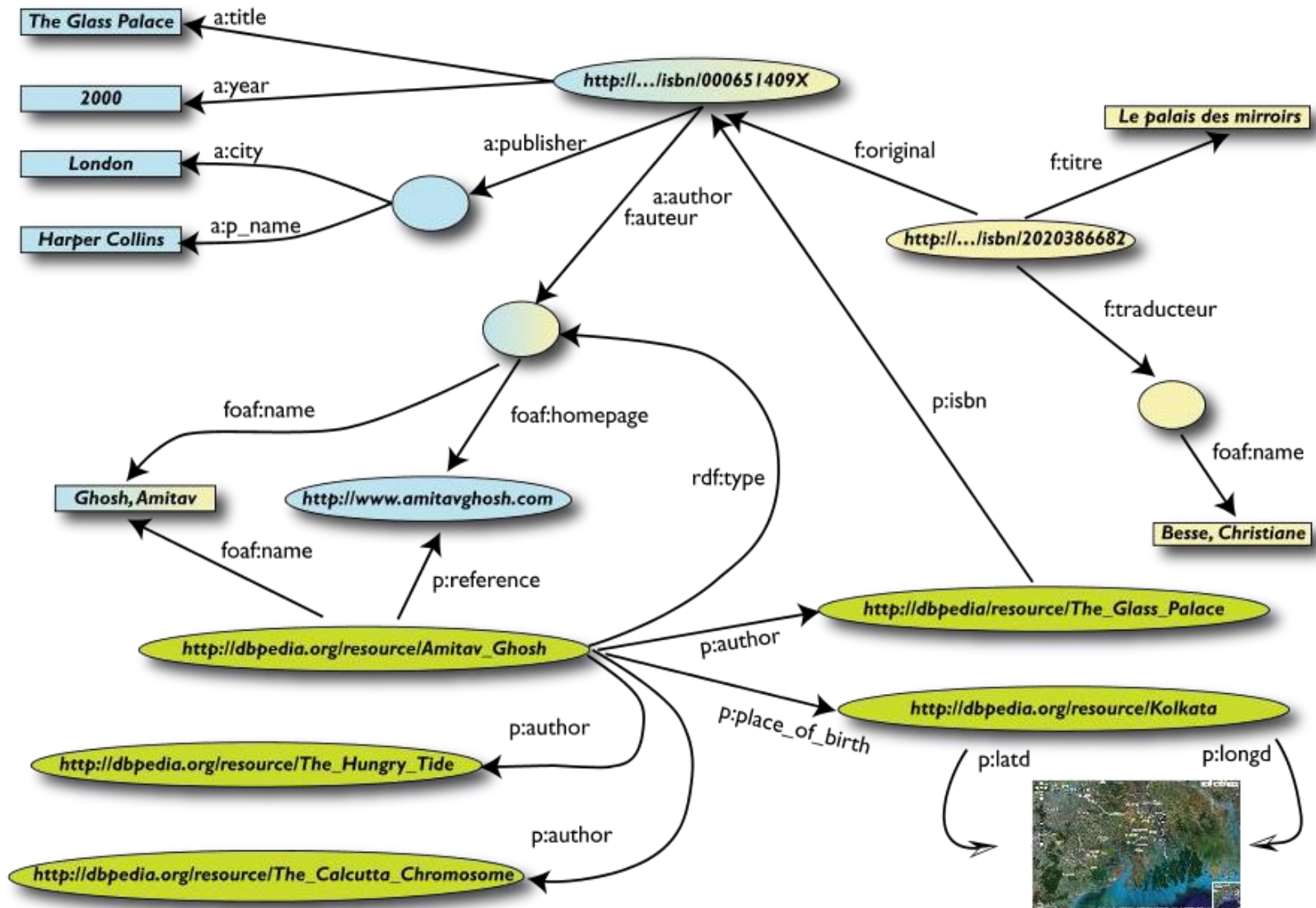
Merge with Wikipedia data
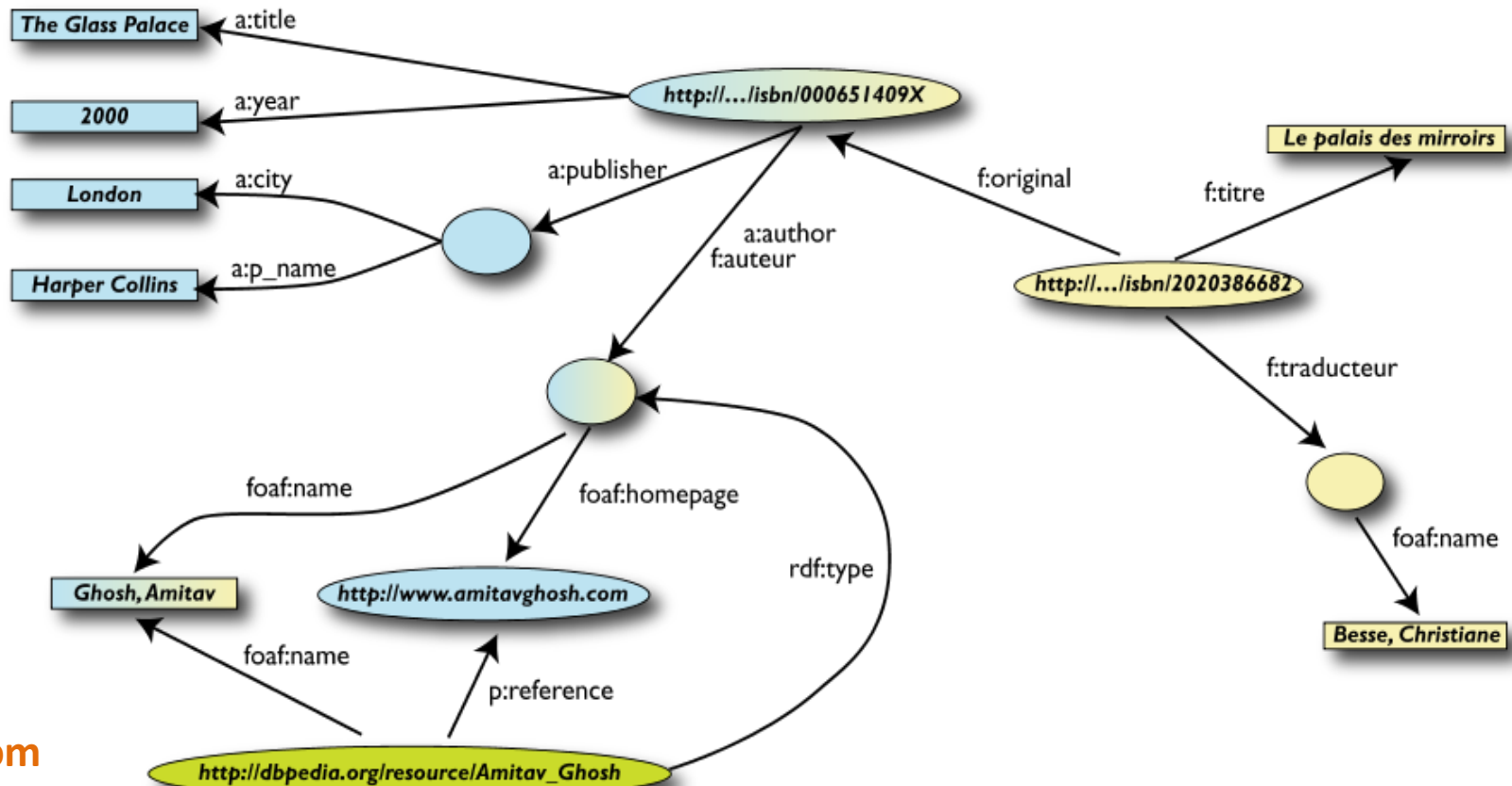
From Wikipedia

Merge with Wikipedia data

Merge with Wikipedia data

Thank you

Prof. Khaled Wassif