

# Simple Knowledge Organisation Systems (SKOS) Requirements for Standardization

Alistair Miles  
CCLRC Rutherford Appleton Laboratory

Dublin Core 2006, Manzanillo

# Theme of this Presentation

- SKOS ... it will all end in tears! (joke)
- SKOS ... what problem are we trying to solve?

# “iSKOS!”

- **S**imple
- **K**nowledge
- **O**rganisation
- **S**ystems
  
- Say “SKOS”

# “¿SKOS?”

- ~~What is it?~~
- What do you think it should be?
- (If you are new to SKOS, try the SKOS Core Tutorial from DC 2005, or the SKOS Core Guide.)

# Past & Present

- W3C Working Drafts (2005) ...
  - “SKOS Core Guide”,
  - “SKOS Core Vocabulary Specification”,
  - “Quick Guide to Publishing a Thesaurus on the Semantic Web”.

# Immediate Future ...

- W3C “**Semantic Web Deployment**” Working Group ...
  - Begins 10 October 2006,
  - Tom Baker & Guus Schreiber to co-chair,
  - **SKOS to W3C Recommendation.**

# Standardization ...

- The purpose of this presentation is to prepare for the standardization of SKOS.

# ... isn't Easy!

- Finding solutions to problems is relatively easy, but ...
- Agreeing on which problem we are all trying to solve is not!

# ¡Requirements!

- **What is SKOS for?**
- I.e. What problem(s) is SKOS supposed to solve?
  
- **How can we state the requirements?**
- I.e. How do we know when SKOS is good enough?

# Outline of this Presentation

- **Part 1 – Suggestions for ...**
  - What requirements should be,
  - How to state them.
  
- **Part 2 – Design Constraints ...**
  - Discuss social and technological trends.
  - (Hitting a moving target!)

# Part 1

## Requirements

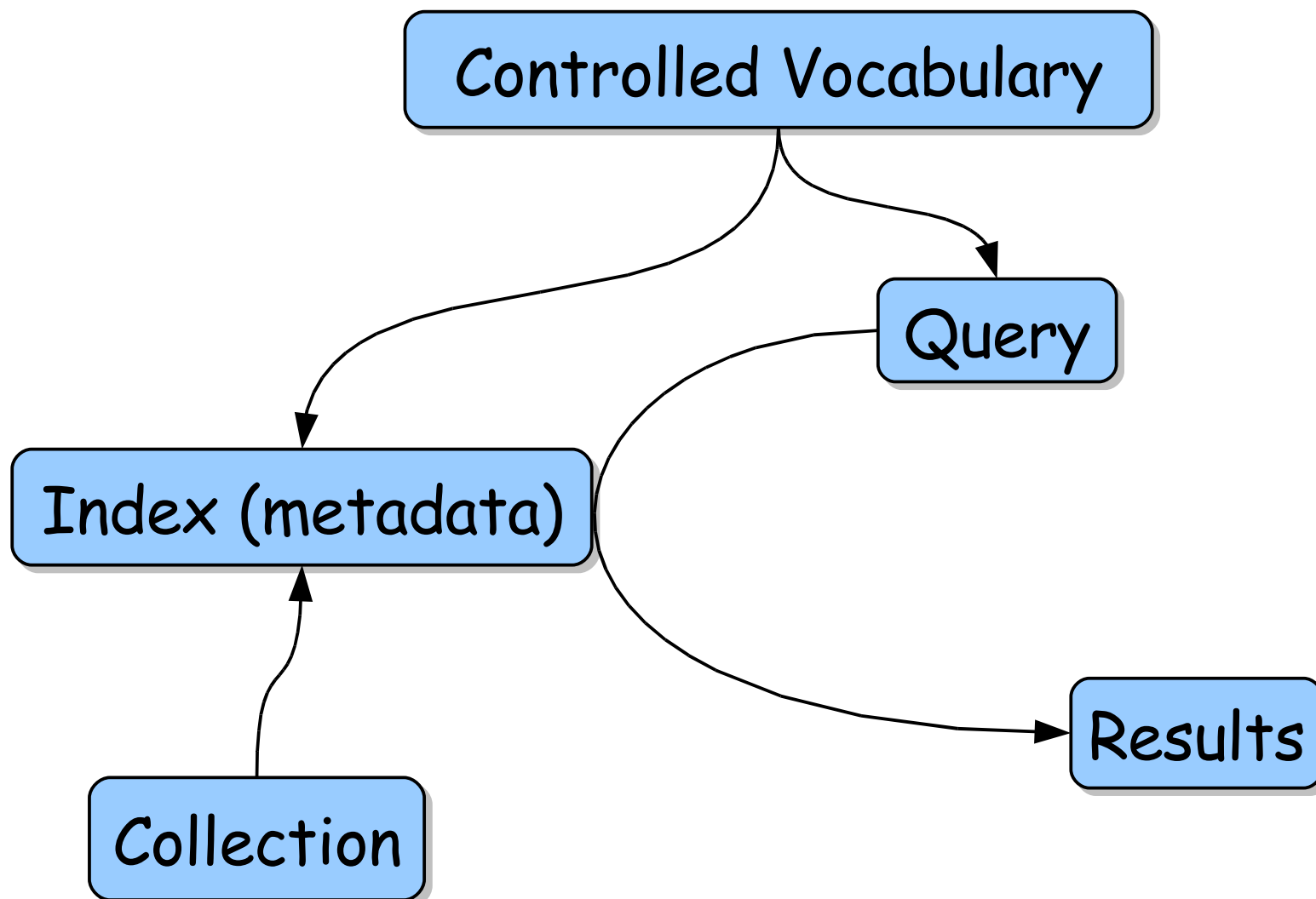
# Statement of Purpose

- Enable the use of controlled structured vocabularies for retrieval ...
- ... in decentralized information systems ...
- ... by providing an extensible language for sharing data between software systems.

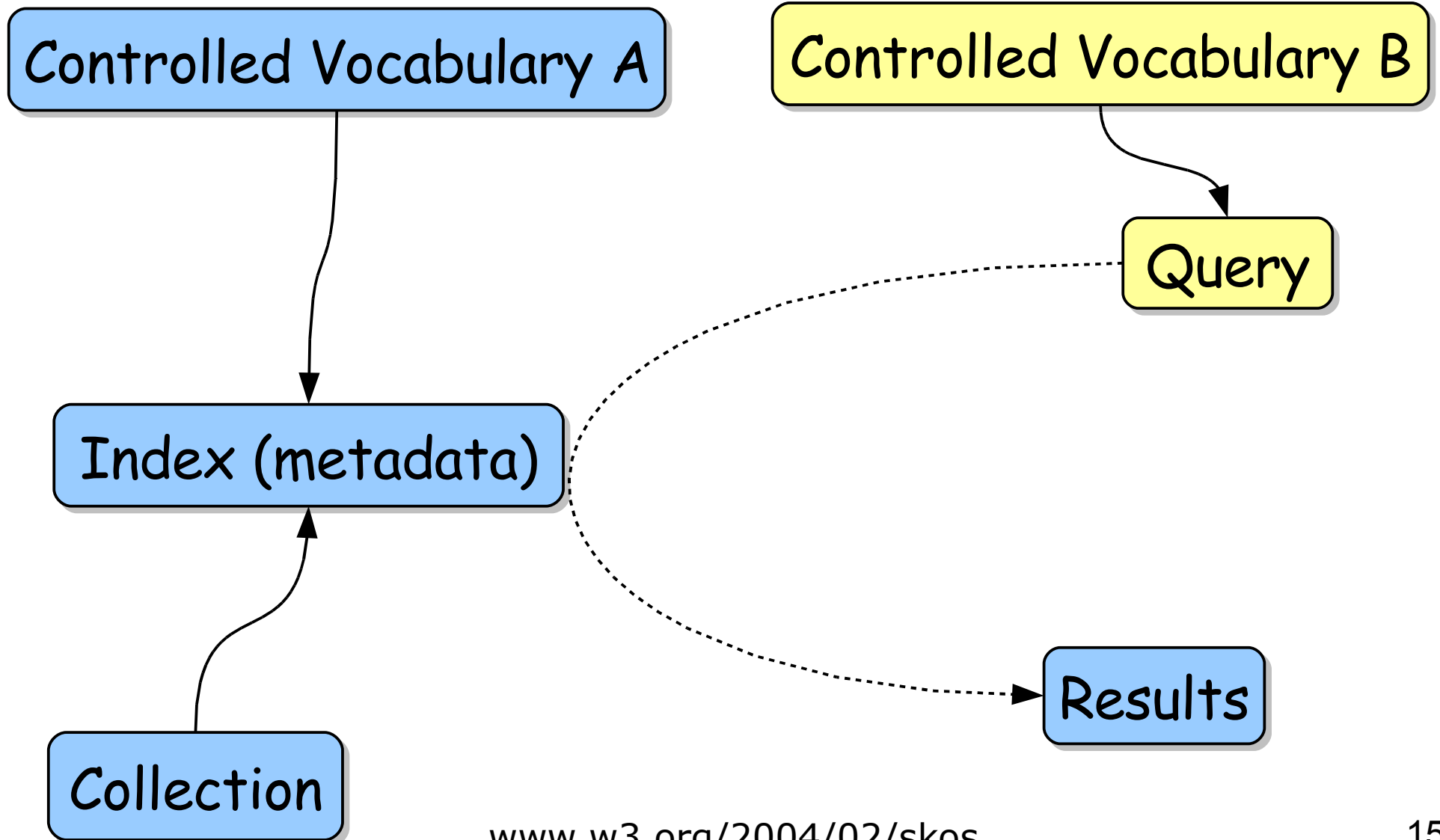
# Retrieval Scenarios

- Two general patterns ...
  - A) A controlled vocabulary is used to “index” and to “retrieve” some objects,
  - B) Different controlled vocabularies are used to “index” and to “retrieve” some objects.

# Retrieval Scenario (Pattern A)



# Retrieval Scenario (Pattern B)



# Use Cases

- Begin requirements analysis by describing **use cases**.
- **All use cases follow either pattern A or B.**
- If use cases are **satisfiable** then SKOS is “**good enough**”.
- N.B. Could still be used for other purposes, but **not a requirement**.

# Pattern A – Use Cases

- Describe the vocabulary (structure, management).
- Describe the index (structure, management).
- Describe the retrieval system (retrieval functionality, user interface).
- N.B. See “Retrieval and the Semantic Web” ... framework for comparative analysis of retrieval systems ... **[purl.org/net/retrieval](http://purl.org/net/retrieval)**

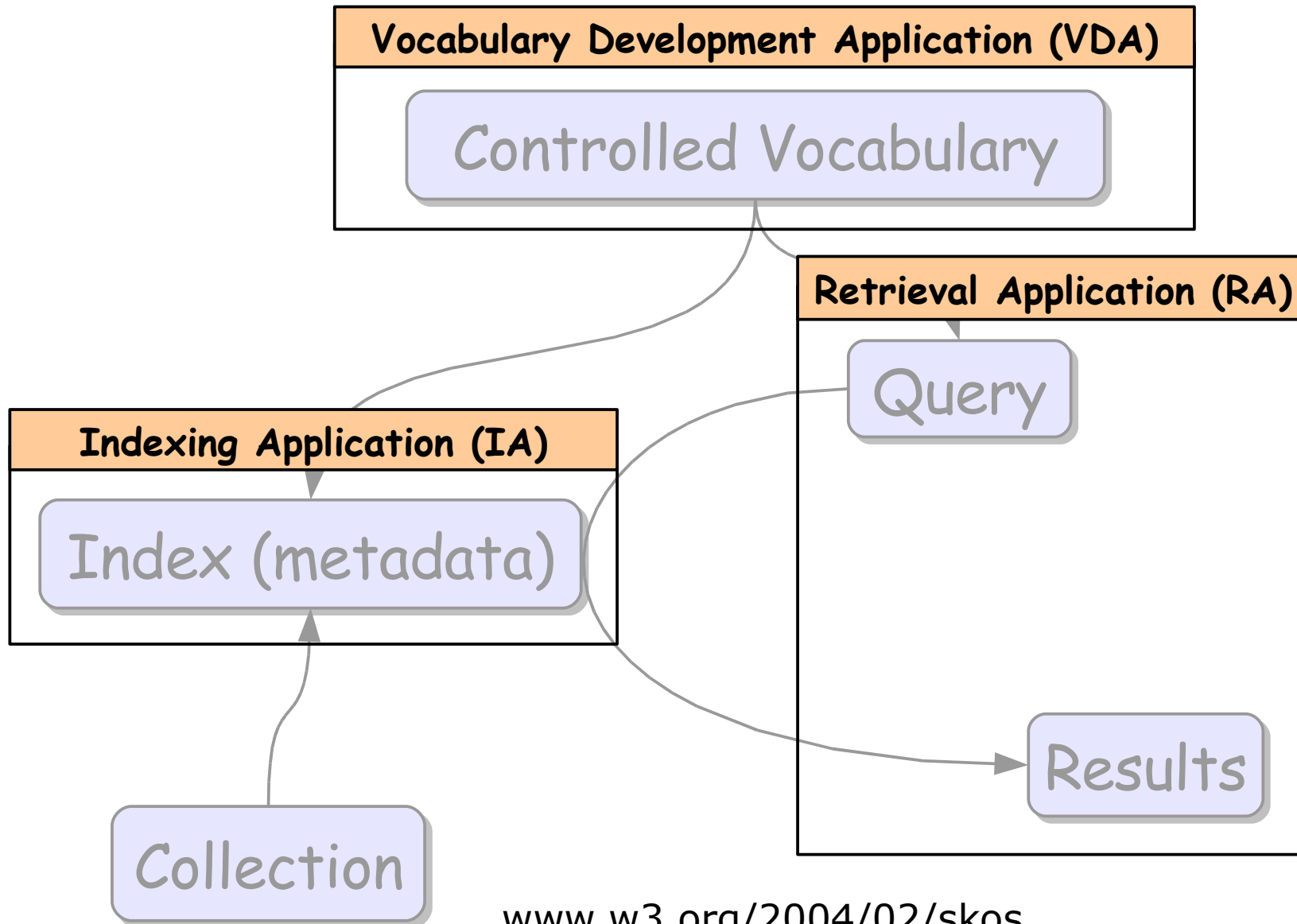
# Statement of Purpose

- Enable the use of controlled structured vocabularies for retrieval ...
  - ... in decentralized information systems ...
  - ... by providing an extensible language for sharing data between software systems.
- 
- **Which data?**
  - **Which software systems?**

# Software & Data

- We must understand what generic software components we are intending to support ...
- ... what functionalities they provide ...
- ... what data they require and ...
- ... how data are shared.

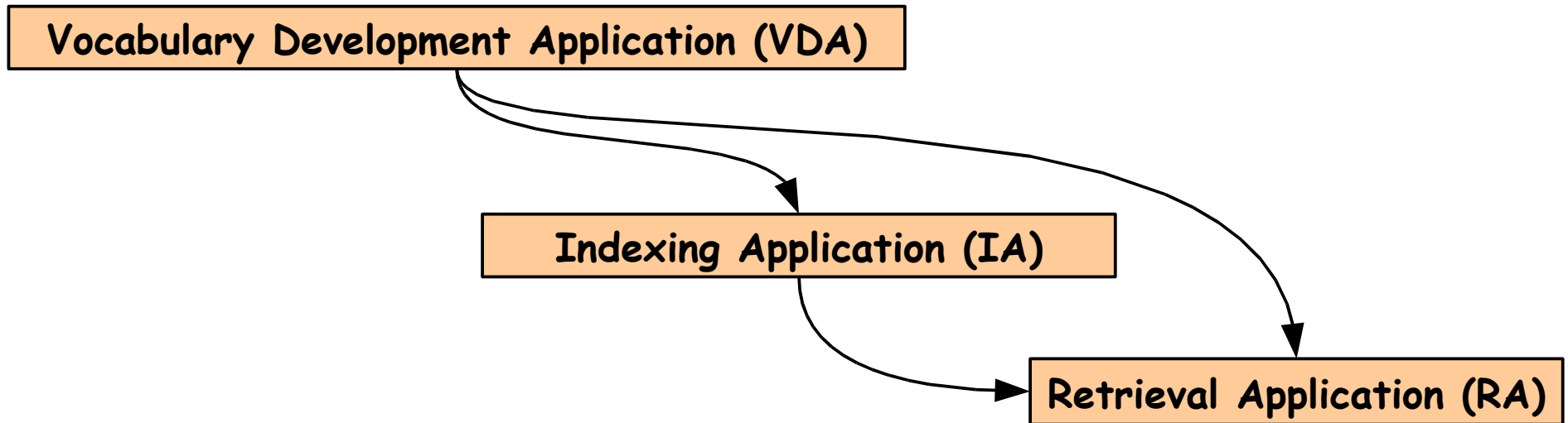
# Pattern A – Software Components



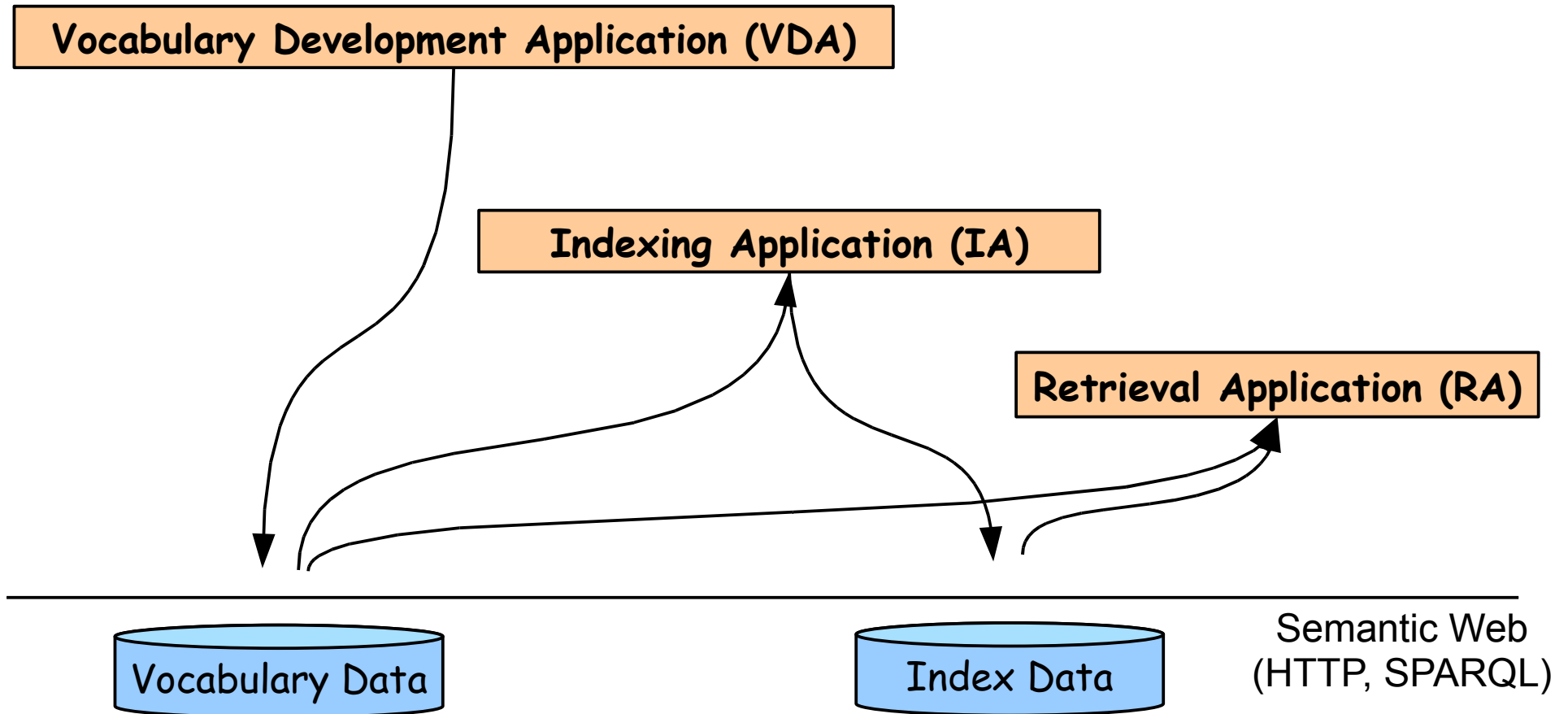
# Pattern A – Software Components

- Ideal Functional Specifications.
- N.B. Standardization means trying to hit a moving target.
- Hopefully, use cases can capture current functional requirements.

# Pattern A – Data Flow

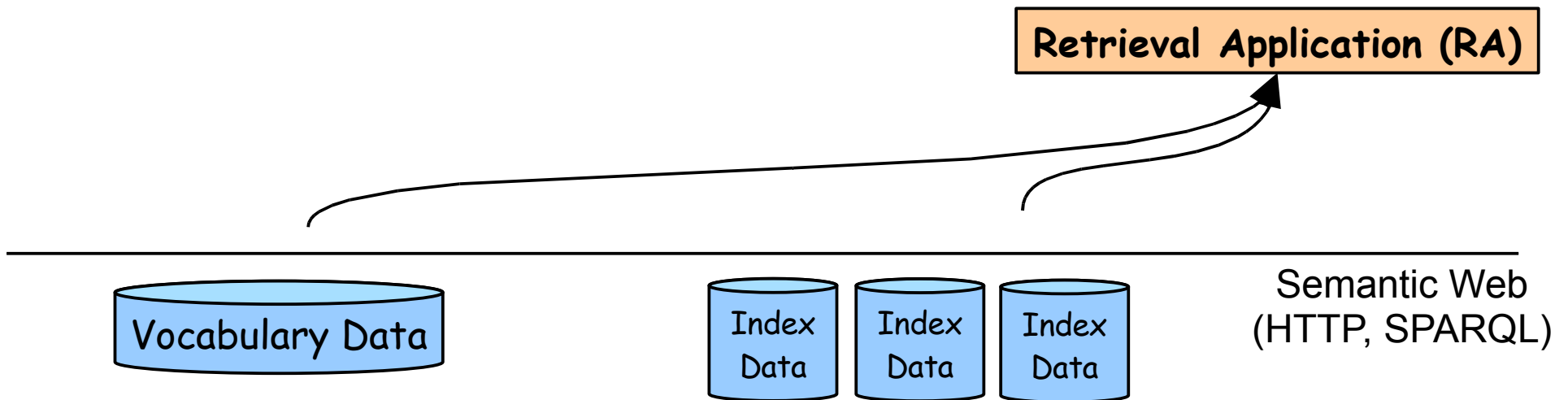


# Pattern A – Interaction



# Pattern A – Interaction ... Why?

- Merge data from multiple sources.



# Summary So Far ...

- SKOS to support retrieval using controlled vocabularies.
- Two general scenarios (A & B).
- Use cases.
- Software architecture ...
  - Functionality,
  - Interaction.

# Part 2

## Trends & Constraints (Gazing into the crystal ball ...)

# Cost!!!

- Vocabularies cost money!
- Indexing costs money!
- Mapping costs money!

# Competition

- Text retrieval.
- Wisdom of crowds ...
  - Pagerank,
  - Search behaviour,
  - Social tagging.

# Bottom Line

- Pressure on solutions based on controlled vocabularies to **cut costs** and **max perform.**
- Pressure creates trends in use of controlled vocabularies ...
  - Integrated solutions,
  - Interoperability,
  - Collaboration,
  - Maintenance.

# Implications for Mapping

- Simplest mapping that achieves required retrieval performance ... ?
- Role for mapping in describing change ... self-updating indexes!

# Final Word

- Semantic Web Deployment starts soon ...
- ... make your voice heard!
- Formal participation, contact your W3C A/C rep.
- Informal participation, subscribe to [public-esw-thes@w3.org](mailto:public-esw-thes@w3.org)
- SKOS and Dublin Core are a match made in heaven :)