

Photo by Julia Wimmerlin
"Hong Kong by night"

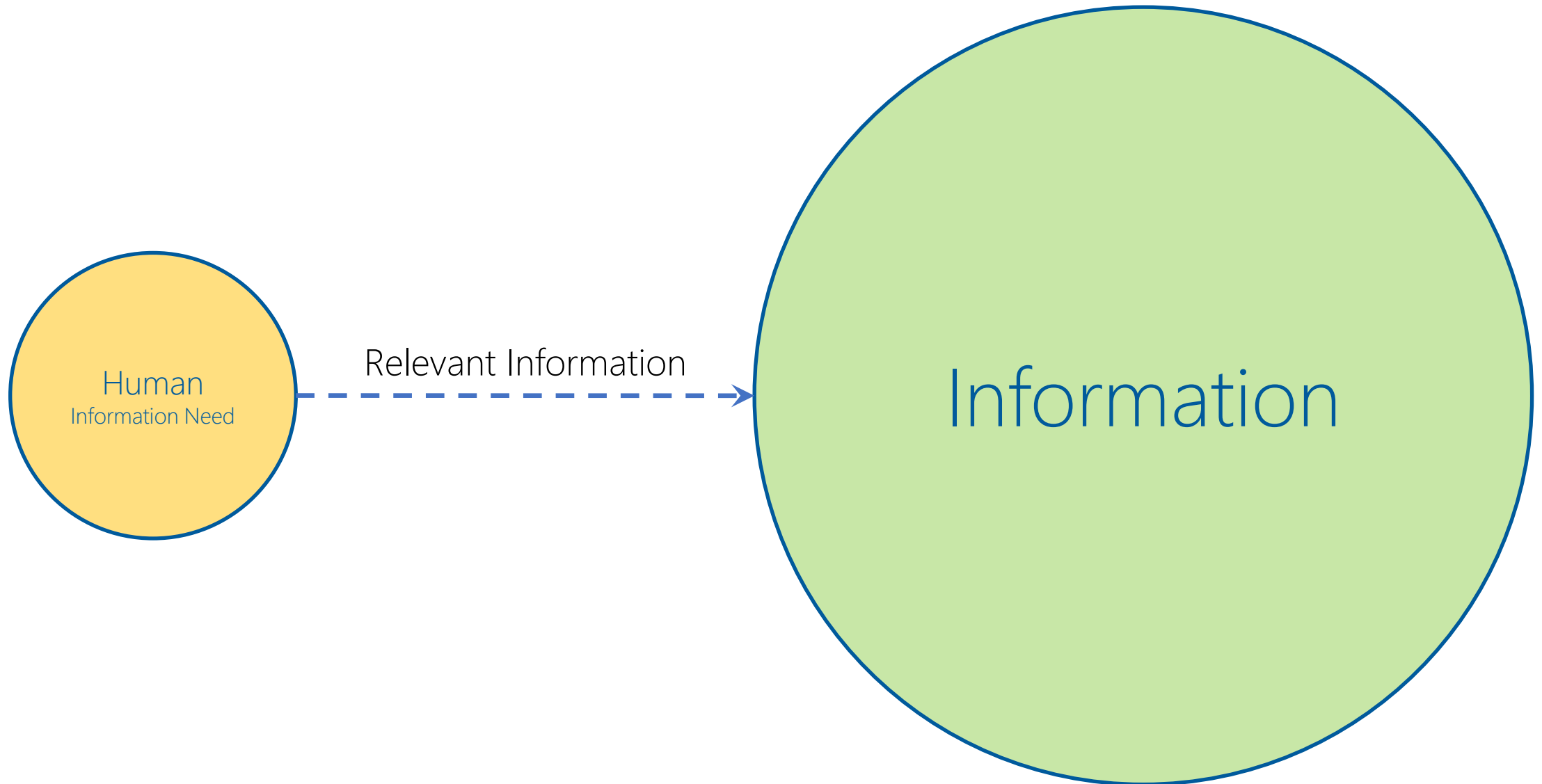


Information Retrieval

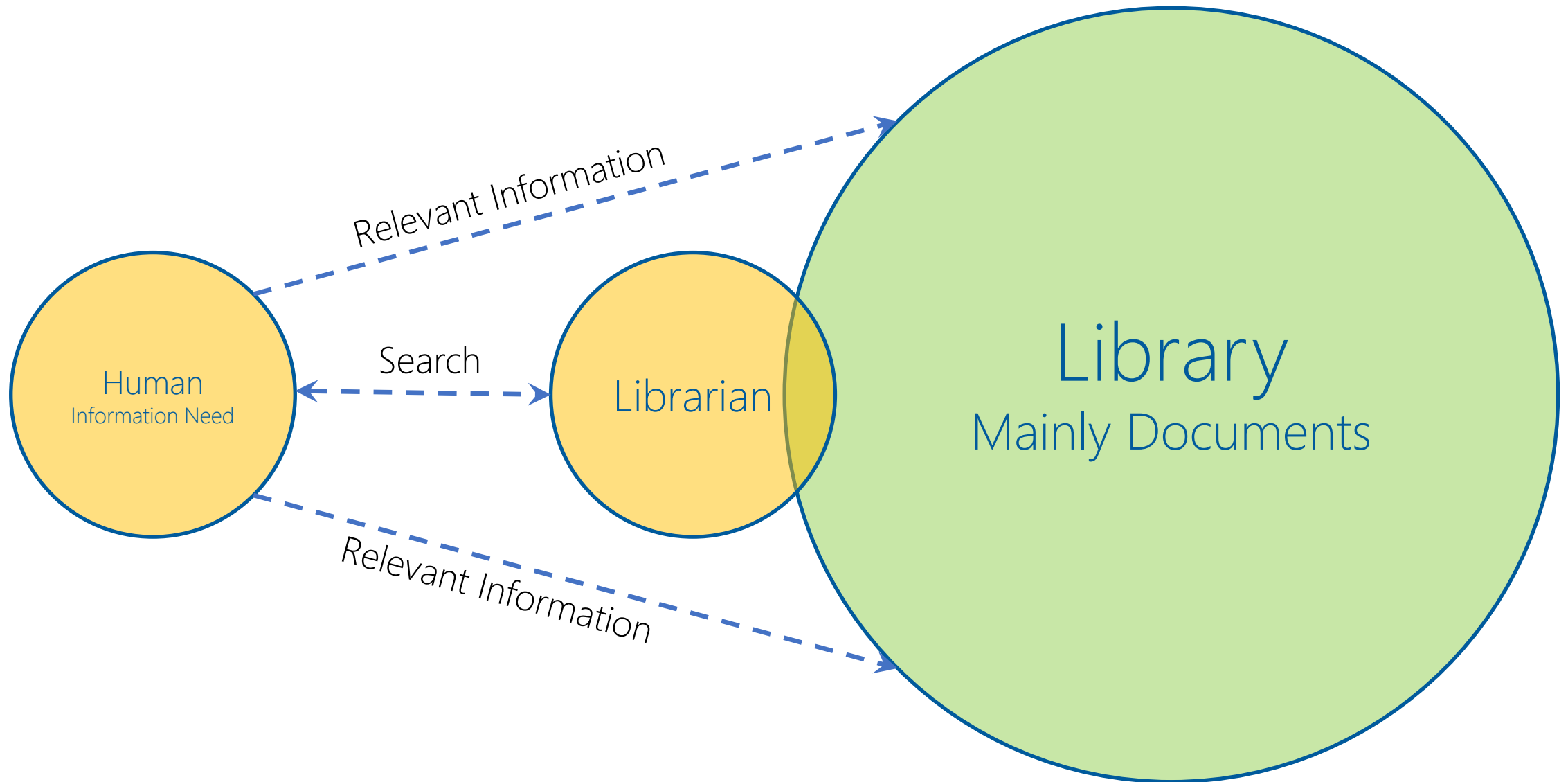
Information Retrieval



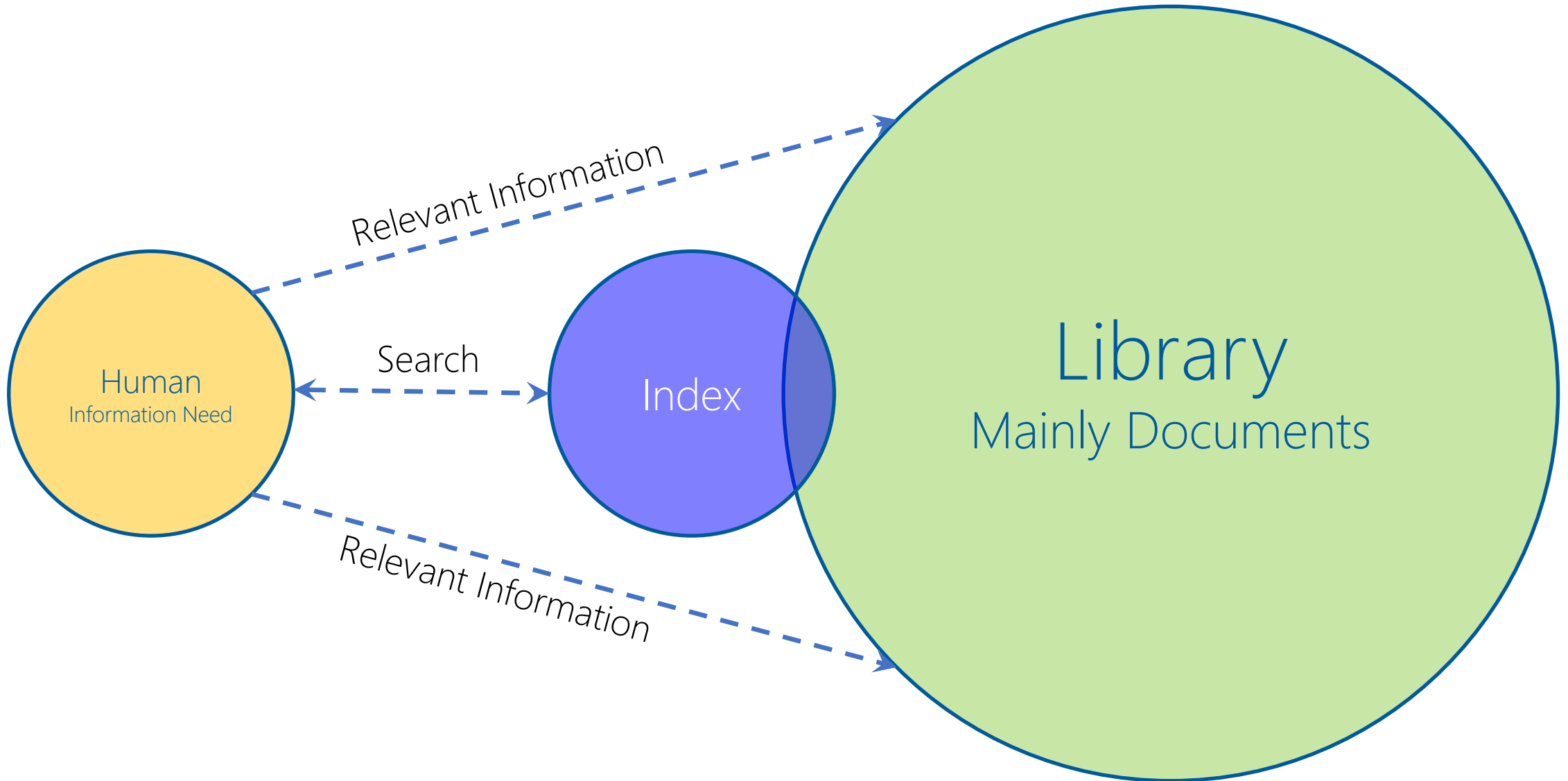
Information Retrieval



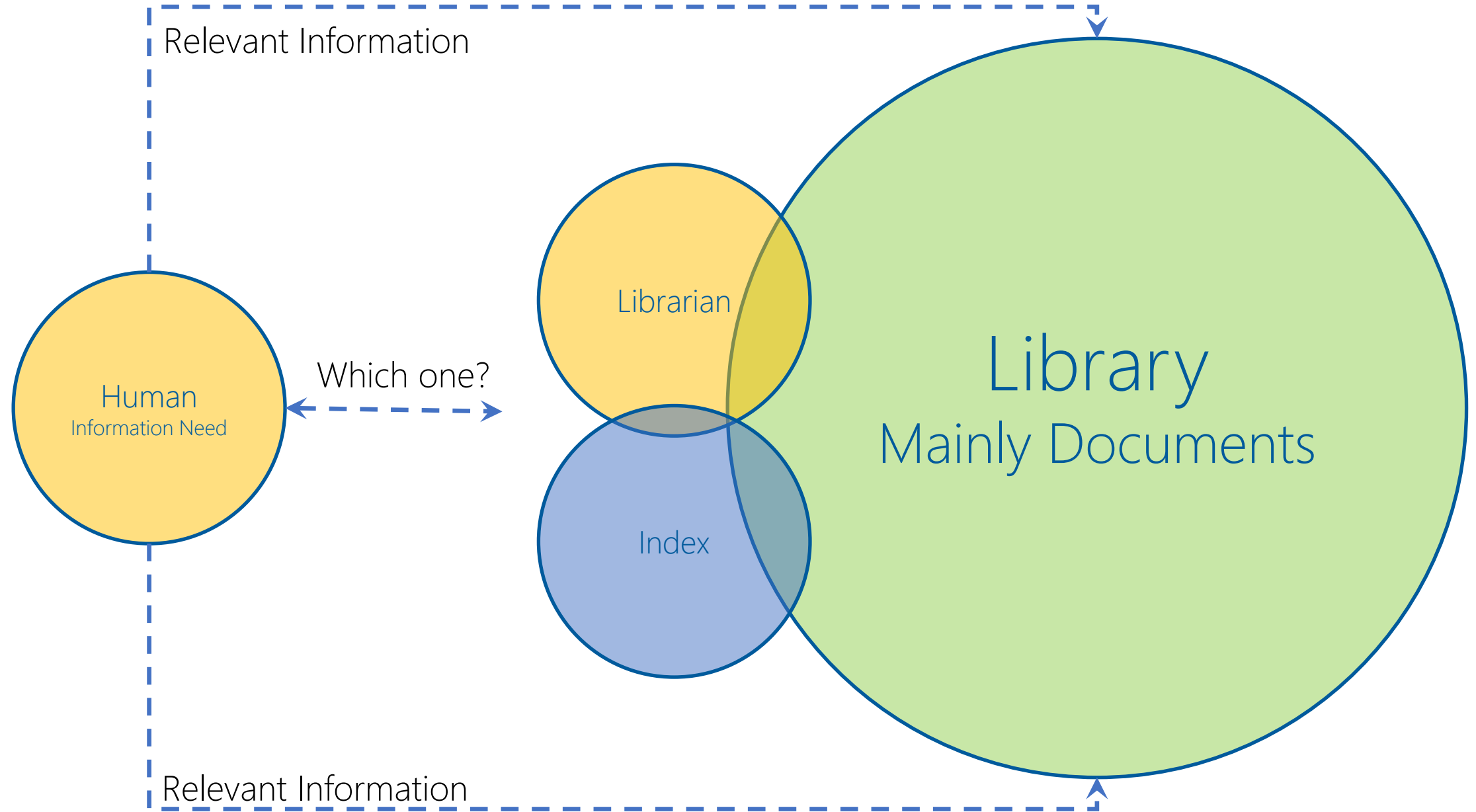
Information Retrieval: Library



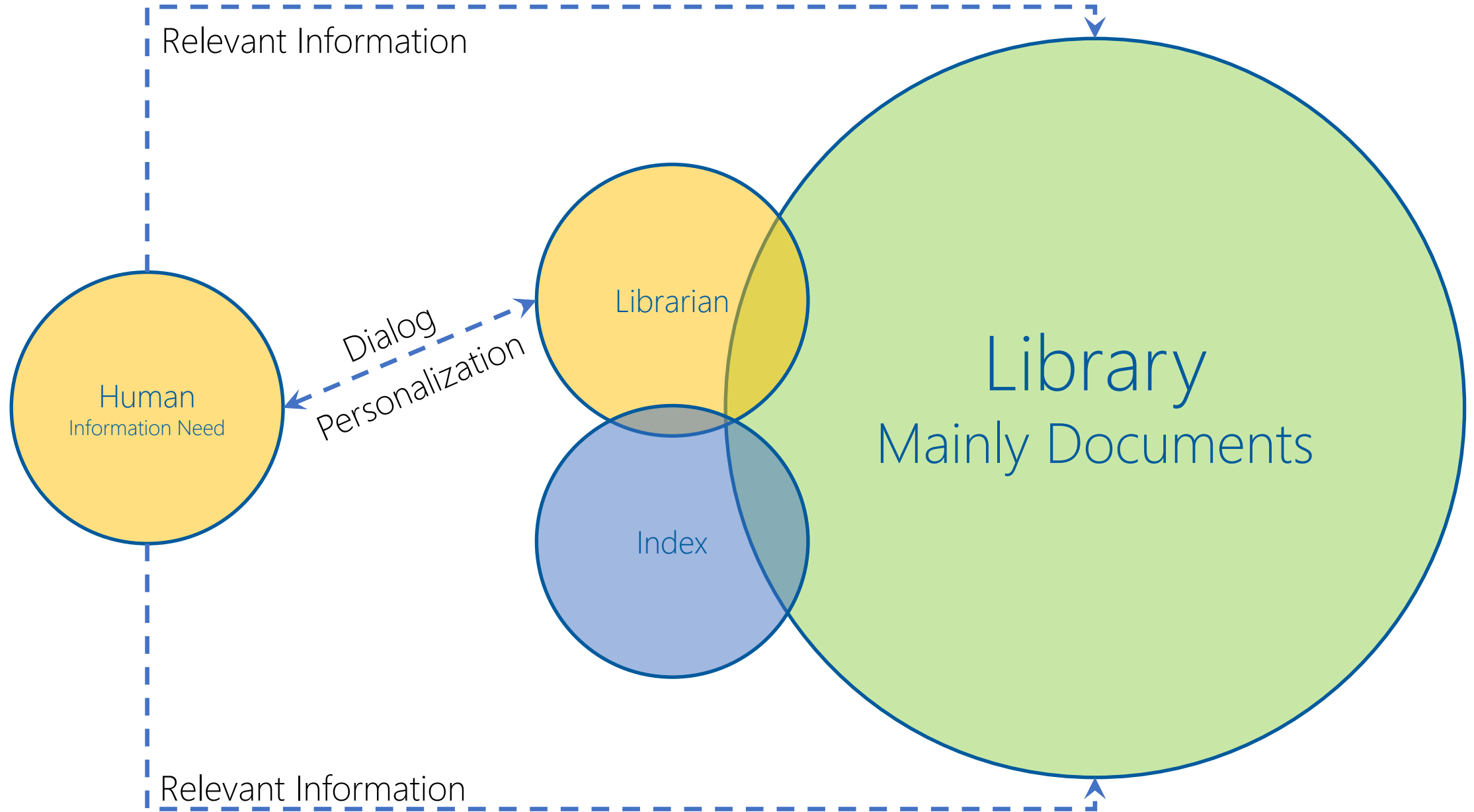
Information Retrieval: Library



Information Retrieval: Library

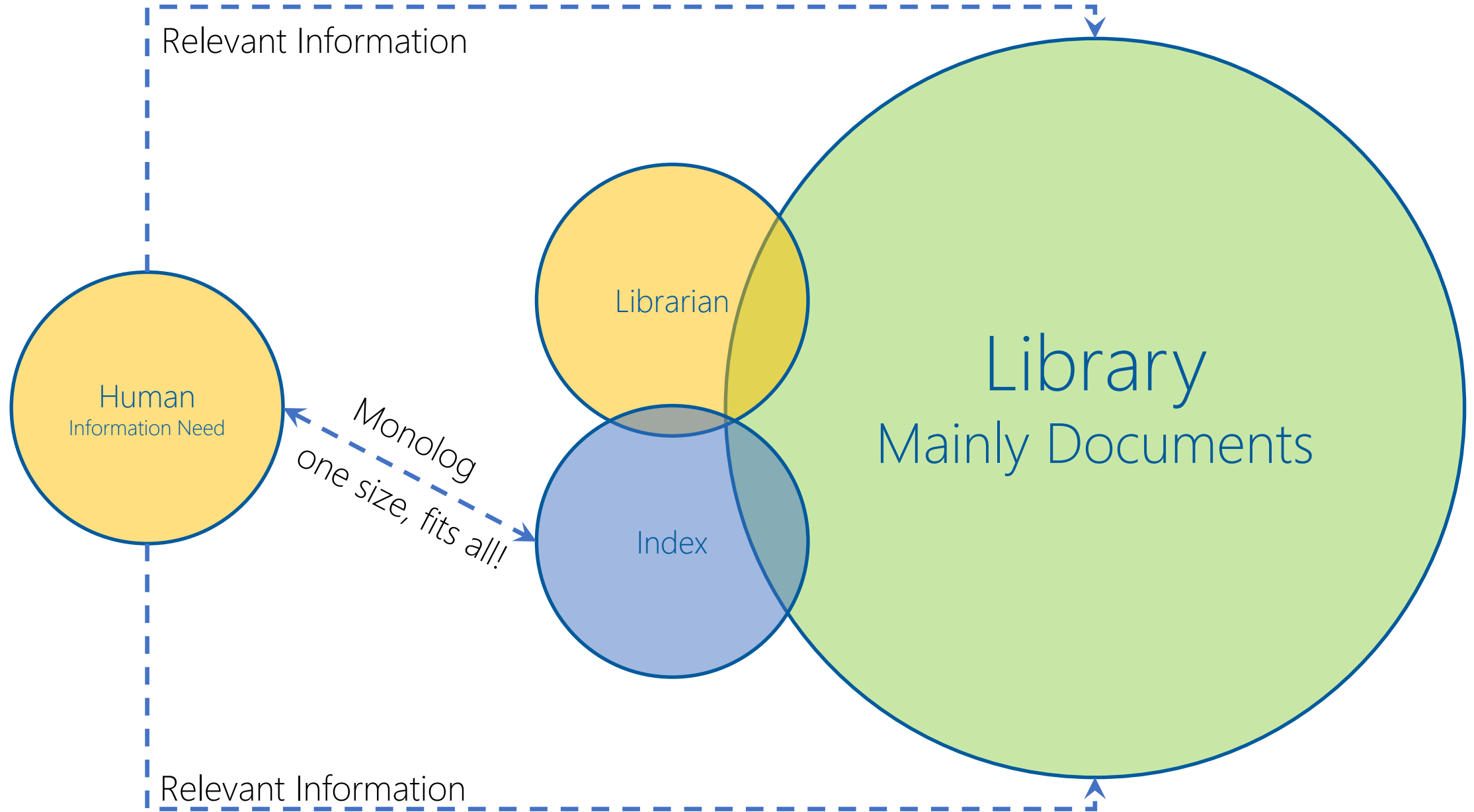


Information Retrieval: Library



Someone walks into a large library, tells the librarian that he is interested in China, and asks for some books. **What sort of books does the librarian recommend?** **That depends.** Is the person a small child who just saw a TV show about China and wants to see more pictures of such an exotic place? Is the person a high school student doing a term paper? Or maybe a prospective tourist? Or a scholar interested in Eastern thought? Can the person read Chinese? The librarian needs to know these things before he can point the reader to the right books. Some of what he needs to know he'll know before he even thinks about it, such as the approximate age of the person. Some things he'll assume until he has evidence to the contrary, such as that the person does not read Chinese. To find out other things, he'll ask a few specific questions. Only after he has a rough model of the person he's talking to can he answer the question.

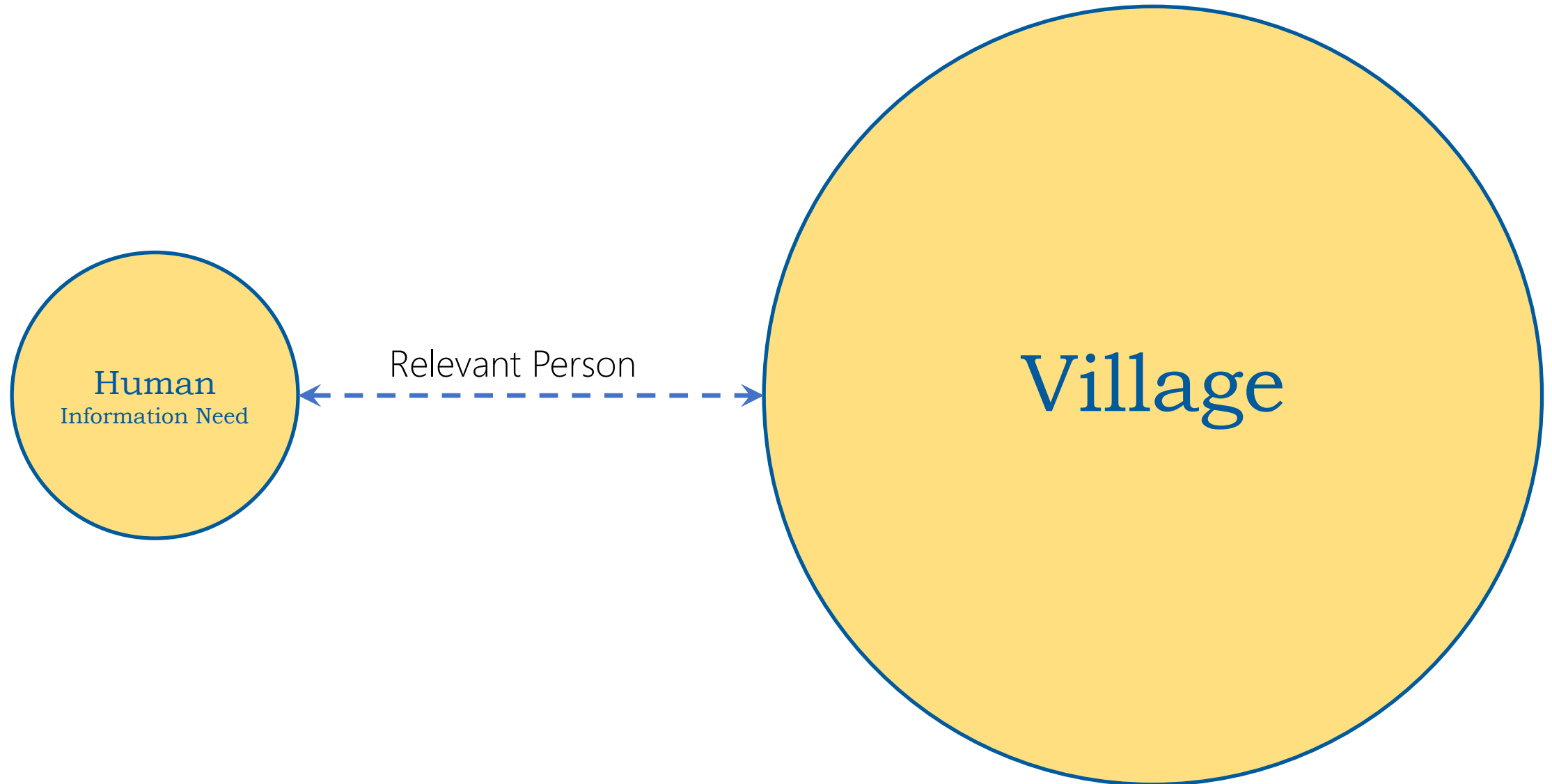
Information Retrieval: Library



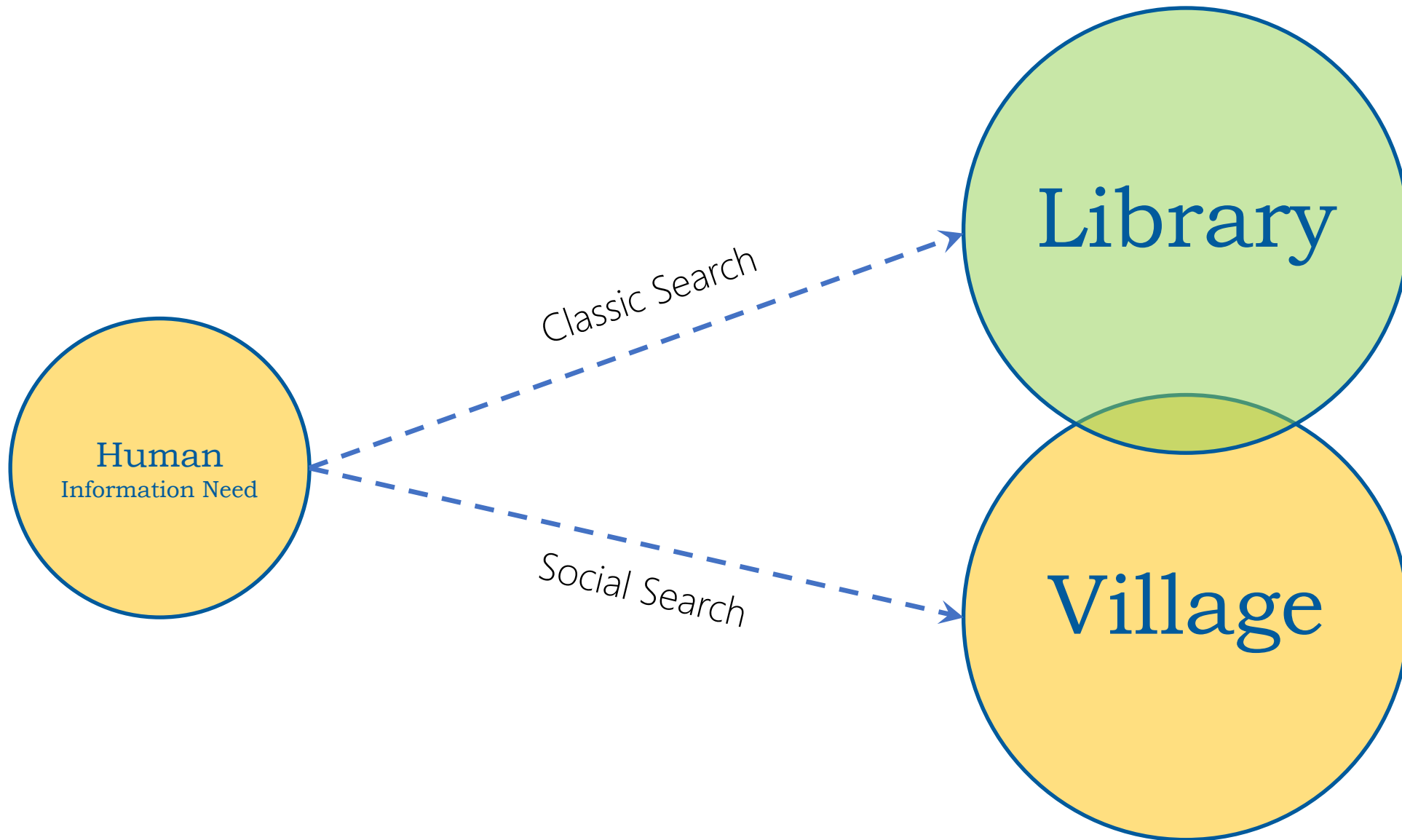
Information Retrieval: Librarian vs. Index

- Language Gap
 - Information Need vs. Information Content
 - “lost bullet” vs. “stray bullet”
- User Model
 - Neural Network
 - Biologist vs. Computer Scientist
- Temporal Language Model
 - Gay
 - Cheerful vs. Homosexual man

Information Retrieval: Village

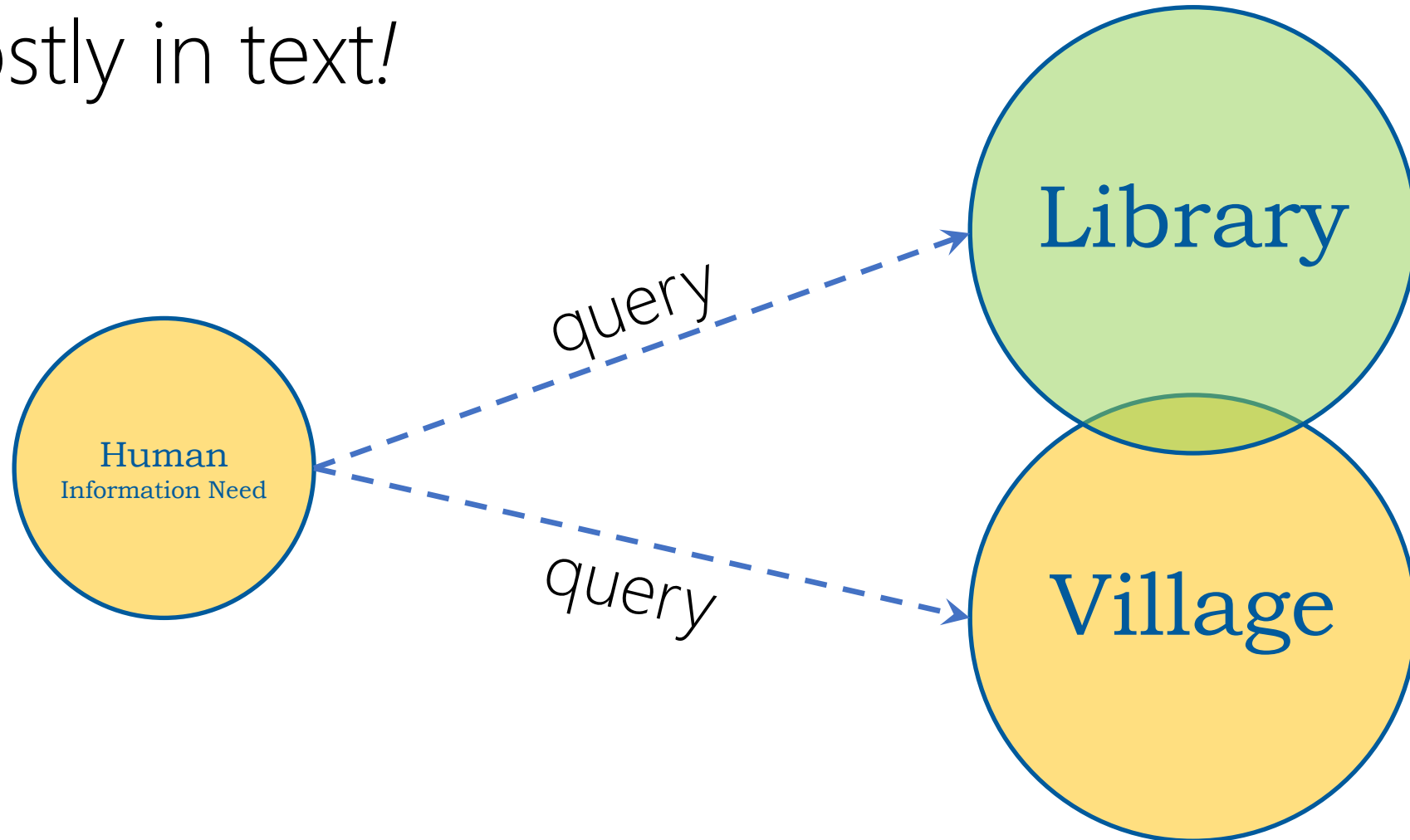


Information Retrieval

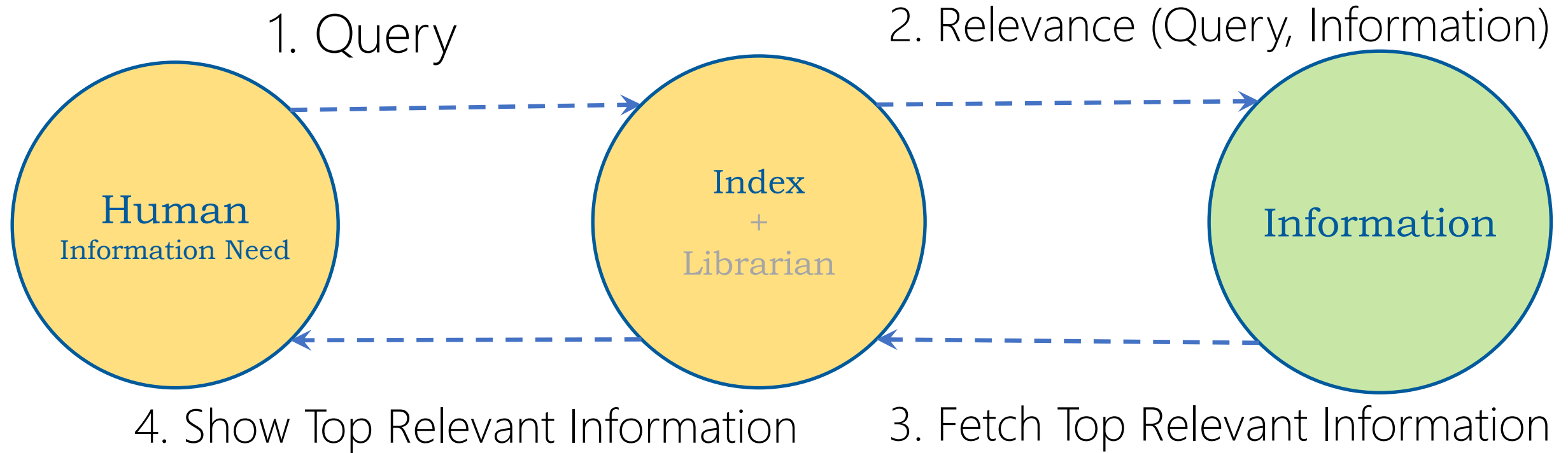


Information Retrieval: Query

- Query
 - *formal* statements of information need
 - mostly in text!

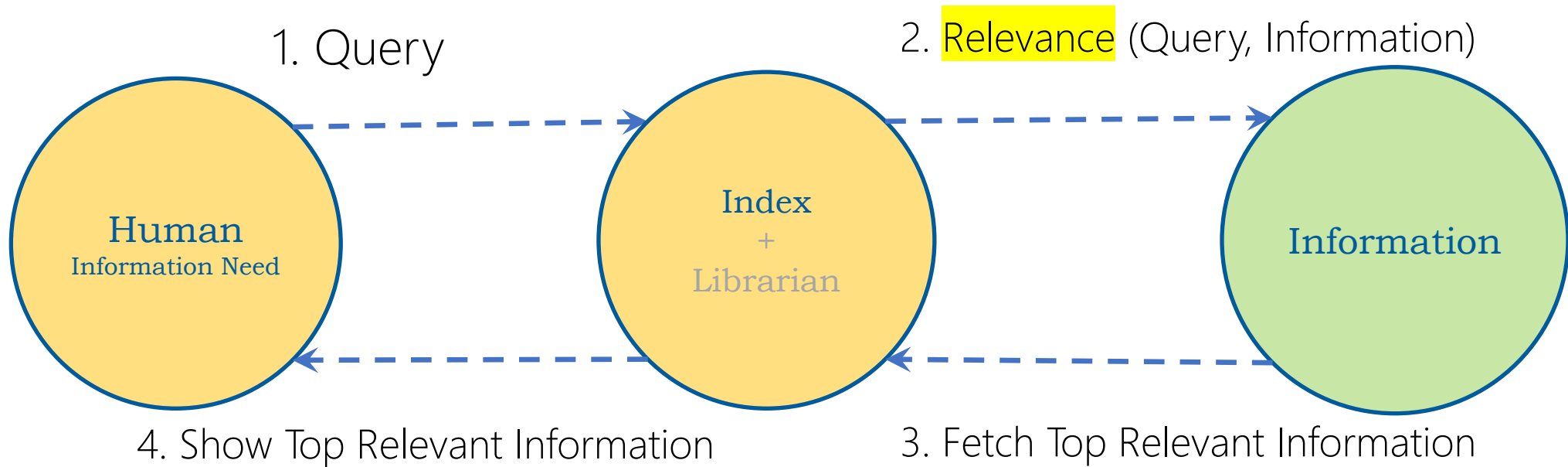


Information Retrieval: Today



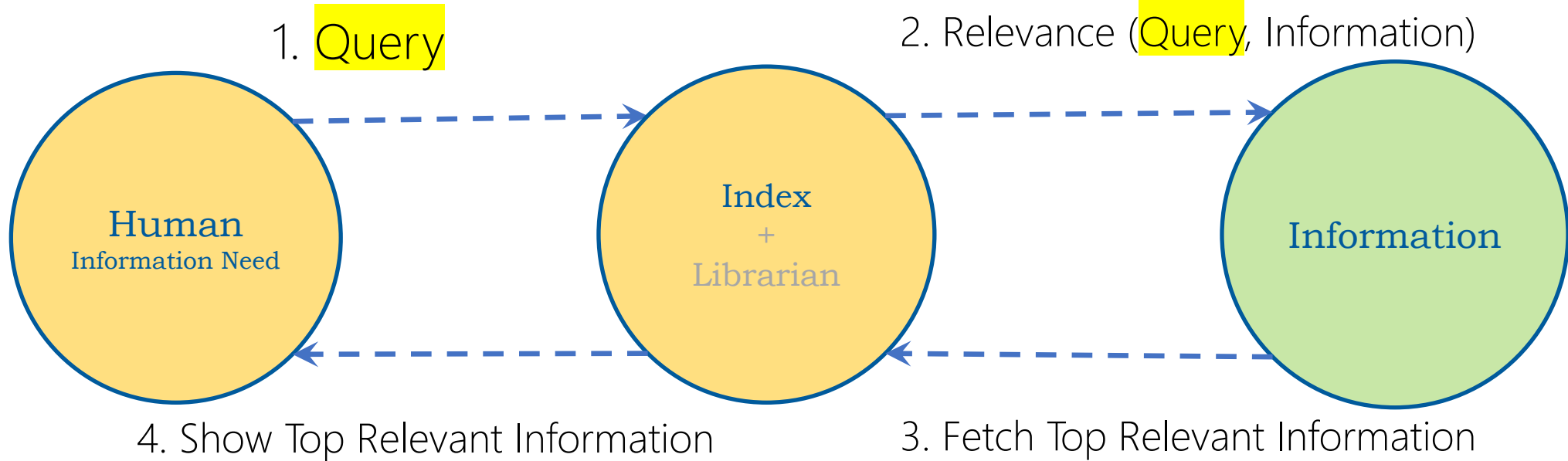
Information Retrieval: Process

- Relevance
 - Boolean (AND, OR, ...): include the query items
 - Vector Space: Similarity(Query Vector Rep., Info. Vector Rep.)
 - ...



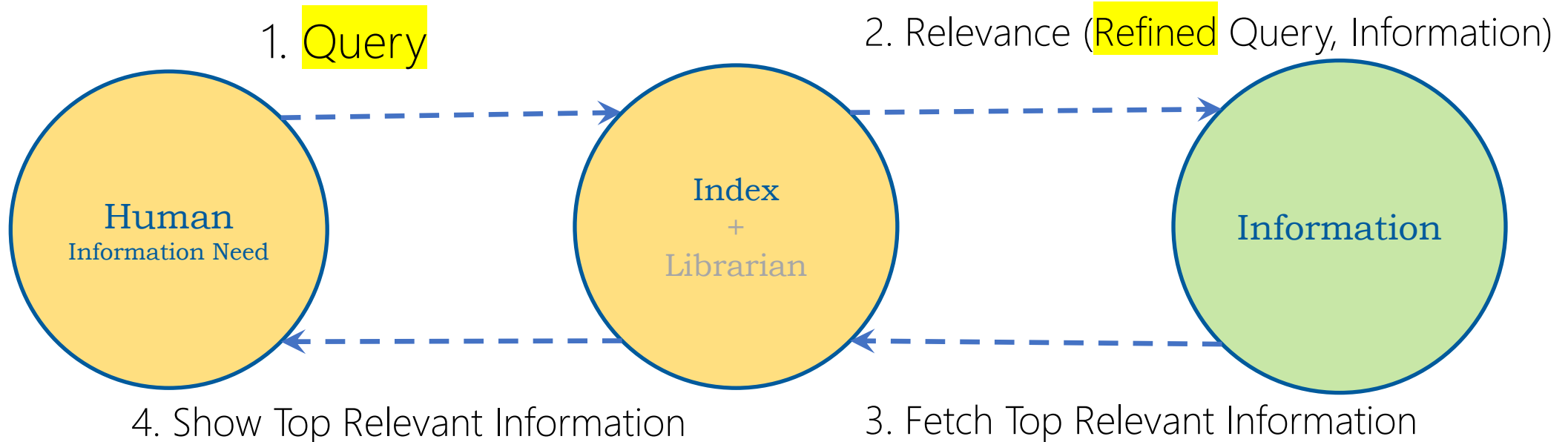
Information Retrieval: Query

- Relevance
 - Boolean (AND, OR, ...): include the query items
 - Vector Space: Similarity(Query Vector Rep., Info. Vector Rep.)
 - ...



Information Retrieval: Query

- Query → Refined Query
 - User Model: a query by a biologist
 - Temporality: a query at year 2021
 - Language Gap: by “lost bullet” user means “stray bullet”
 - Autocorrection: by “precisn” user means “precision”
 - ...



Query Refinement

Information Retrieval: Query Refinement

- Query Refinement
 - Expansion: adding new words to original query
 - Reduction: removing words of original query
 - Replacement: replacing words of original query
 - Mixed → Pipeline of Changes

Learn to Refine Query

Supervised Query Refinement

Supervised Query Refinement

Given a training set of pairs of (q, q')

- $q' = \text{Refined}(q)$
- q' retrieves more relevant information
- q' retrieves more relevant information in order

We can train a translation from q to q' :

- seq2seq model
- seq2seq model with attention

Supervised Query Refinement

Given a training set of pairs of (q, q')

- $q' = \text{Refined}(q)$
- q' retrieves more relevant information
- q' retrieves more relevant information in order

Information Need: "a bullet that hit a person unintentionally"

q: "lost bullet"	q': "stray bullet"	Relevant Information in order of relevance	
D4	D1		D1
D2	D3		D2
D9	D4		D3
D7	D8		D4

Supervised Query Refinement

Metrics

- Set Retrieval: classification of relevant vs. non-relevant
 - Precision, Recall, F-measure, Success (Hit) ratio
- Ranked Retrieval
 - RR, AP, nDCG

Relevance Judgment



q: "lost bullet"	q': "stray bullet"	Relevant Information in order of relevance
D4	D1	D1
D2	D3	D2
D9	D4	D3
D7	D8	D4

Supervised Query Refinement

Metrics

- Set Relevance
- Precision
- Ranked
- RR, A

Evaluation Metrics

Jaime Arguello
INLS 509: Information Retrieval
jarguell@email.unc.edu

March 25, 2013

https://ils.unc.edu/courses/2013_spring/inls509_001/lectures/10-EvaluationMetrics.pdf

q: "lost bullet"

Order of relevance

D4

D2

D9

D7

D4

D8

D3

D4

Supervised Query Refinement

Given a training set of pairs of (q, q')

From where?

- Manual creation of dataset.
- Automatic creation of dataset!
 - Users Search Sessions & Click-through
 - ReQue

Supervised Query Refinement

Automatic creation of dataset!

- Users Search Sessions & Click-through

Scenario:

- Within a short time period
- User tries to find relevant information until she satisfies

Supervised Query Refinement

Automatic creation of dataset!

- Users Search Sessions & Click-through

Scenario:

- Within a short time period
 - User tries to find relevant information until she satisfies
- 1- Start with a naïve query about her information need

Supervised Query Refinement

Automatic creation of dataset!

- Users Search Sessions & Click-through

Scenario:

- Within a short time period
- User tries to find relevant information until she satisfies
 - 1- Start with a naïve query about her information need
 - 2- If she sees relevant information, she clicks on it

The truth about relevance and its order are captured by the user!

Supervised Query Refinement

Automatic creation of dataset!

- Users Search Sessions & Click-through

Scenario:

- Within a short time period
- User tries to find relevant information until she satisfies
 - 1- Start with a naïve query about her information need
 - 2- If she sees relevant information, she clicks on it
 - 3- She refine the query to find more relevant information

Supervised Query Refinement

Automatic creation of dataset!

- Users Search Sessions & Click-through

Conversational Search

Supervised Query Refinement



bullet



[All](#)

[Images](#)

[Videos](#)

[Shopping](#)

[News](#)

[More](#)

[Settings](#)

[Tools](#)

About 447,000,000 results (0.92 seconds)

<https://www.pcna.com> > bullet > en-us ▾

Bullet | PCNA

New look, new website, new experience. All to deliver on our promise to you. [Learn More](#). [Shop By Category](#).

Bullet

New look, new website, new experience. All to deliver on our ...

[More results from pcna.com »](#)

People also ask

What bullet means?



How do you type bullet points?



What is a bullet in computer?



Why is it called a bullet point?



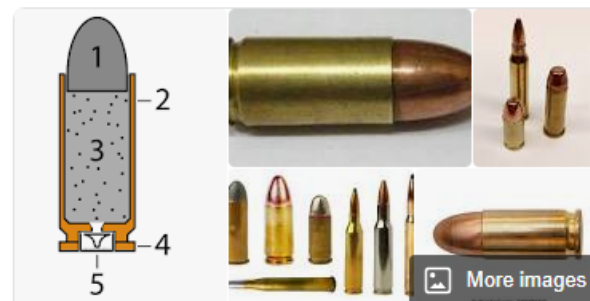
[Feedback](#)

<https://www.pcna.com> > bullet > en-ca > department > sho...

Shop All - Bullet - PCNA

[Bags \(340\)](#) · [Totes](#) · [Drawstring Bags](#) · [Backpacks](#) · [Coolers](#) · [Lunch Coolers](#) · [Event Coolers](#) ·

[Briefcases & Messengers](#) · [Duffels](#) · [Sport](#) · [Travel](#) · [Travel](#) · [Fanny](#)



[More images](#)

Bullet



A bullet is a kinetic projectile and a component of firearm ammunition that is expelled from a gun barrel during shooting. The term is from Middle French and originated as the diminutive of the word *boule*, which means "small ball". [Wikipedia](#)

People also search for

[View 10+ more](#)



Pistol



Rifle




Revolver



Shotgun

[Feedback](#)

Supervised Query Refinement



Google

bullet in street

All Images Videos News Shopping More Settings Tools

About 199,000,000 results (0.59 seconds)

Rating Hours Visit history

Bullet Digital Post
No reviews · Video production service
219 Dufferin St · (416) 531-0222
Closed · Opens 8:30 a.m. [Directions](#)

With a Bullet
No reviews · Company
2207 Queen St E #2 · (416) 646-2817 [Website](#) [Directions](#)

Stinson Design
5.0 ★★★★★ (1) · Design agency
401 Richmond St W #243 · +1 888-960-9851
Closed · Opens 9 a.m. [Website](#) [Directions](#)

[→](#) [View all](#)

Supervised Query Refinement

Google

bullet in street unintentionally

All Images Videos News Shopping More Settings Tools

About 3,530,000 results (0.51 seconds)

Showing results for **bullet in street unintentionally**
Search instead for **bullet in street unintentionally**

Images for bullet in street unintentionally

Report images

→ View all

<https://www.pinterest.com/david6212/2018-oops-i-acc...>
52 2018: Oops! I accidentally shot into my neighbor's property ...
Every year, the Richmond Police Department issues a **public** plea: Please don't fire guns in the air ... **Bullet unintentionally** fired into neighbor's home in Ventura.

<https://www.1011now.com/2020/07/31/man-accidenta...>
Man accidentally fires gun, sending bullet through neighbor's ...
Jul. 31, 2020 — 46th **Street**, just south of Van Dorn **Street**, on a report of a gunshot. Lincoln Police talked to the 70-year-old homeowner who said that a **bullet** had ...

<https://hazlitt.net/longreads/what-bullet-can-do> ▼
What A Bullet Can Do | Hazlitt
Mar. 7, 2018 — Today, the **public** health community, including the Centers for Disease Control (CDC), uses the term "**unintentional** shooting" instead of ...

Supervised Query Refinement

Google

accident bullet in street unintentionally

Q All Images Videos News Shopping More Settings Tools

About 3,120,000 results (0.60 seconds)

<https://www.1011now.com> › 2020/07/31 › man-accidenta...

Man accidentally fires gun, sending bullet through neighbor's ...

Jul. 31, 2020 — 46th **Street**, just south of Van Dorn **Street**, on a report of a gunshot. Lincoln Police talked to the 70-year-old homeowner who said that a **bullet** had ...

<https://www.nytimes.com> › 2013/09/29 › children-and-gu...

Children and Guns: The Hidden Toll - The New York Times

Sep. 28, 2013 — Children shot **accidentally** — usually by other children — are ... Believing he had removed all the **bullets**, he pointed the pistol at his sister and squeezed the trigger. ... But **accidents**, more than the other firearm-related deaths, come with ... states do not consider death certificate data a matter of **public** record.

<https://surgicalneurologyint.com> › surgicalint-articles ▼

Stray bullet: An accidental killer during riot control - Surgical ...

Background: The use of force to control **public** uprisings, riots, unruly mobs is an ... **Unintentional bullet** injuries have been reported in many case series. ... A number of victims believed that the weapon was not loaded at the time of the **accident**.

Stray bullet: An accidental killer during riot control - Surgical ...

Supervised Query Refinement

The diagram illustrates the process of supervised query refinement. It shows a sequence of search results for the query "stray bullet" on Google. Three blue arrows point from the left side of the image to specific search results, indicating a refinement process. The first arrow points to a video result titled "Caught on camera: Man struck by stray bullet". The second arrow points to a Wikipedia result titled "Stray bullet - Wikipedia". The third arrow points to another Wikipedia result titled "Celebratory gunfire - Wikipedia".

Google

stray bullet

All Images News Videos Shopping

About 16,300,000 results (0.62 seconds)

Videos

Caught on camera: Man struck by stray bullet

YouTube · WPLG Local 10
Dec. 19, 2018

View all

[https://en.wikipedia.org › wiki › Stray_bullet](https://en.wikipedia.org/wiki/Stray_bullet)

Stray bullet - Wikipedia

A **stray bullet** is a **bullet** that, after being fired from a gun, hits an unintended target. A shooting accident may occur due to missing a target when hunting or sports shooting, or due to accidental/negligent discharges, or during crossfire or celebratory gunfire.

[https://en.wikipedia.org › wiki › Celebratory_gunfire](https://en.wikipedia.org/wiki/Celebratory_gunfire)

Celebratory gunfire - Wikipedia

Celebratory gunfire is the shooting of a firearm into the air in celebration. The practice sometimes results in random death and injury from **stray bullets**.

Supervised Query Refinement

Automatic creation of dataset!

- Users Search Sessions & Click-through

The first query $\rightarrow q$

The last query $\rightarrow q'$

We assume the retrieve information for q' is better than q . Why? Because the user quit the search!

Supervised Query Refinement

Automatic creation of dataset!

- Users Search Sessions & Click-through
- AOL
- MSMARCO



Supervised Query Refinement

Session 9A: Queries

CIKM'17, November 6-10, 2017, Singapore

Learning to Attend, Copy, and Generate for Session-Based Query Suggestion

Mostafa Dehghani*
University of Amsterdam
dehghani@uva.nl

Sascha Rothe
Google Research
rothe@google.com

Enrique Alfonseca
Google Research
ealfonseca@google.com

Pascal Fleury
Google Research
fleury@google.com

Session 4B: Queries

SIGIR '19, July 21–25, 2019, Paris, France

Context Attentive Document Ranking and Query Suggestion

Wasi Uddin Ahmad
University of California, Los Angeles
Los Angeles, CA
wasiahmad@ucla.edu

Kai-Wei Chang
University of California, Los Angeles
Los Angeles, CA
kwchang.cs@ucla.edu

Hongning Wang
University of Virginia
Charlottesville, VA
hw5x@virginia.edu

Supervised Query Refinement

Automatic creation of dataset!

- Users Search Sessions & Click-through

The first query $\rightarrow q$

The last query $\rightarrow q'$

We assume the retrieve information for q' is better than q . Why? Because the user quit the search!

was she satisfied when quit?!

Supervised Query Refinement

Automatic creation of dataset!

- Users Search Sessions & Click-through

The first query $\rightarrow q$

The last query $\rightarrow q'$

We assume the retrieve information for q' is better than q . Why? Because the user quit the search!

Topic Drift!

Supervised Query Refinement: Topic Drift

Automatic creation of dataset!

- Users Search Sessions & Click-through

#	Session Queries
1	[is sicily part of italy]→[is sri lanka part of africa]→[what are the maldives]→[what is great barrier reef]→[what is tanzania]
2	[immutable, definition]→[define obliged]→[meaning of industrious]→[what do vocation mean]→[legal definition capricious]→[definition of. contempt]→[definition of famine]→[meaning of obstinate]
3	[what is google classroom]→[synonym for commotion]→[missionaries definition]→[intersect definition]→[types of intersecting lines]→[stimulus value definition]→[definition of system unit] →[destruction of lesions definition]→[touch definition]→[top load washer machine]
4	[definition tangible]→[define translucent]→[astringent define]→[definition of retribution]→[define defined contribution]
5	[are great white shark endangered]→[german shepherd/labrador]→[australian shepherd price] [longevity of boston terrier]→[cost for cairn terrier]→[is chihuahua a dog]

Supervised Query Refinement

Automatic creation of dataset!

Resource Track

CIKM '20, October 19–23, 2020, Virtual Event, Ireland

ReQue: A Configurable Workflow and Dataset Collection for Query Refinement

Mahtab Tamannaee
mtamannaee@ryerson.ca
Ryerson University, Canada

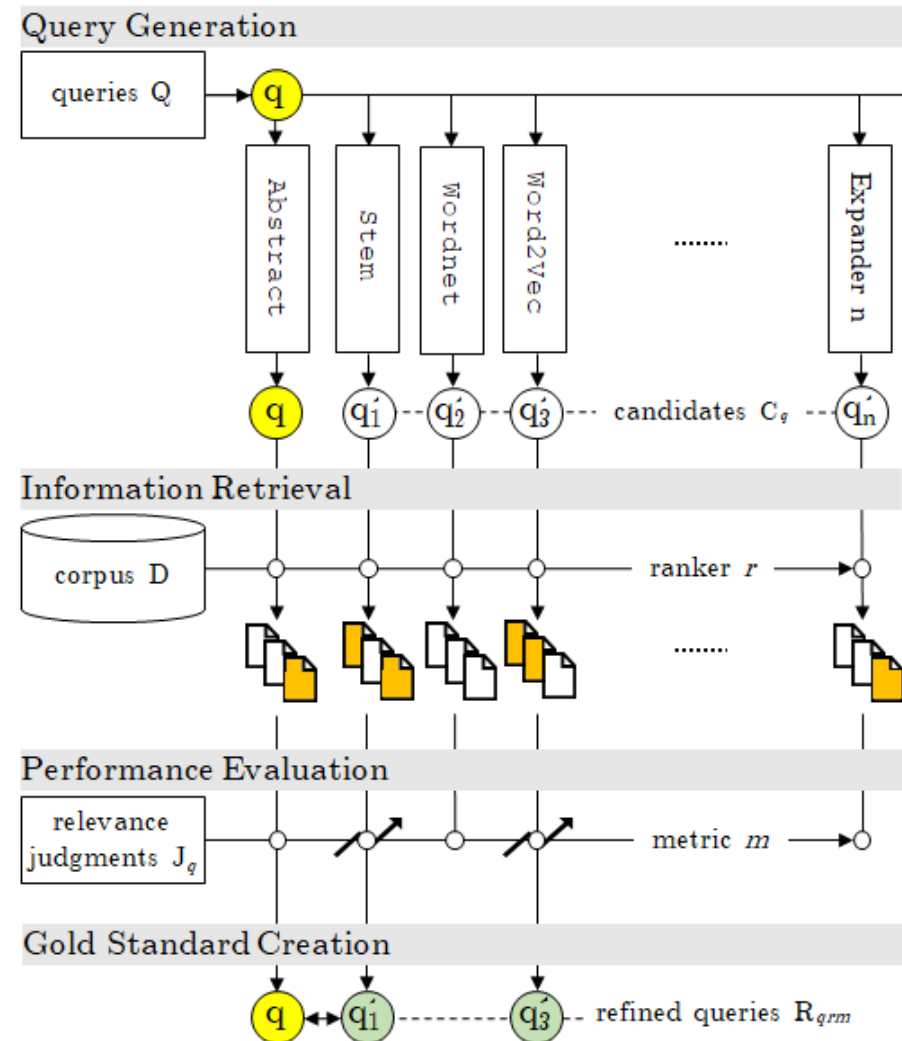
Jamil Samouh
jtsamouh@gmail.com
Ryerson University, Canada

Hossein Fani
hfani@uwindsor.ca
University of Windsor, Canada

Samad Paydar
paydar@ryerson.ca
Ryerson University, Canada

Fattane Zarrinkalam
fzarrinkalam@ryerson.ca
Ryerson University, Canada

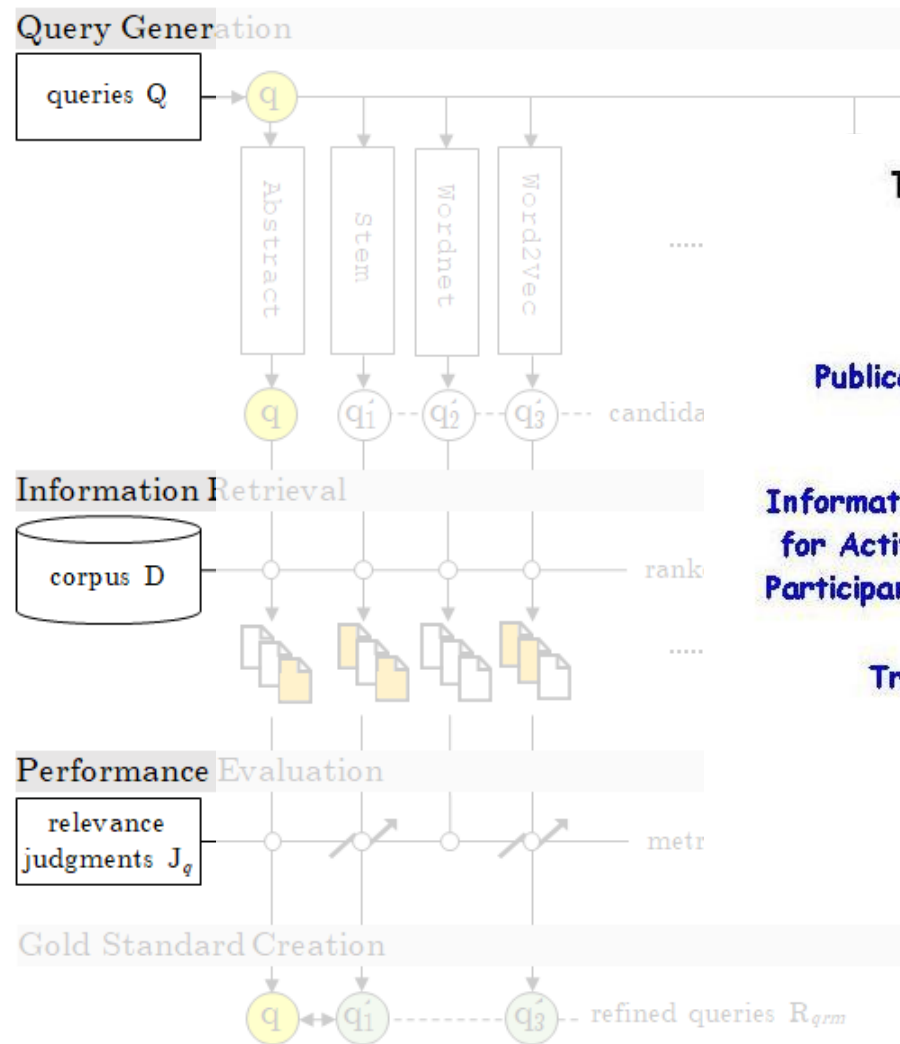
Ebrahim Bagheri
bagheri@ryerson.ca
Ryerson University, Canada



Supervised Query Refinement

Automatic creation of dataset!

- ReQue



Text REtrieval Conference (TREC)

...to encourage research in information retrieval from large text collections.



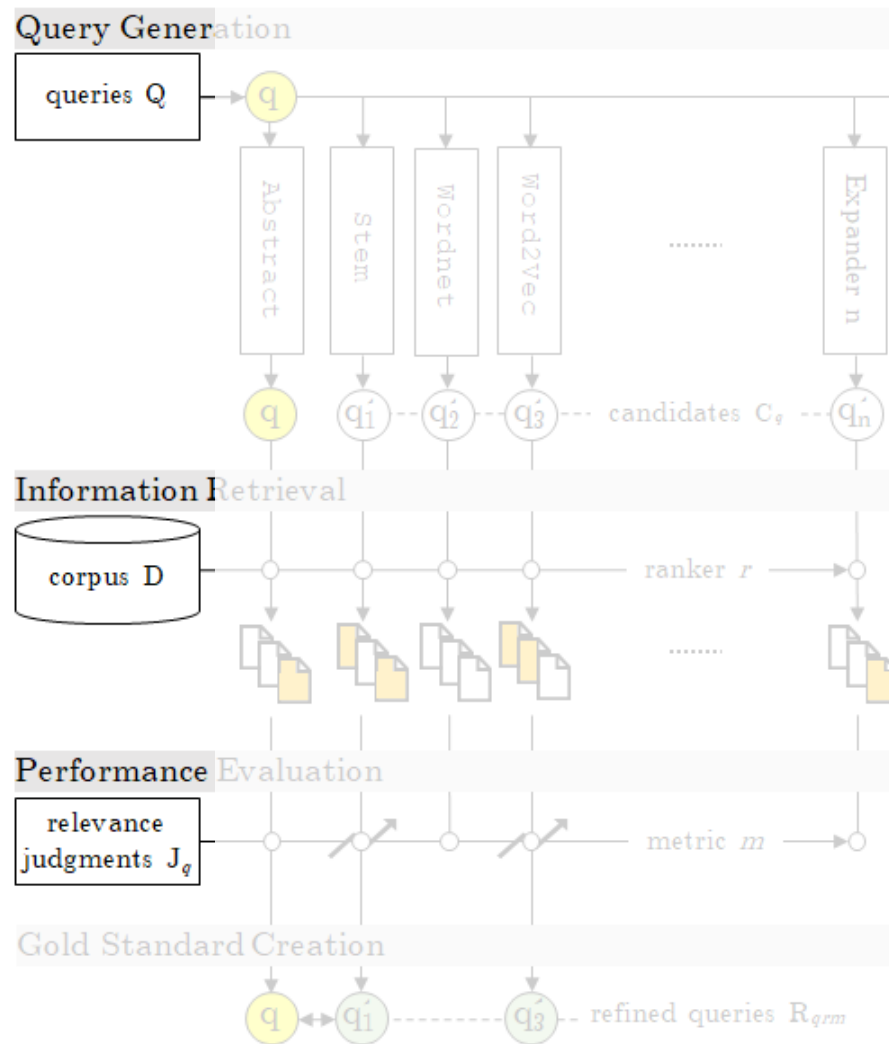
<https://trec.nist.gov/>

Supervised Query Refinement

Automatic creation of dataset!

- ReQue

Robust04
Clueweb
Gov



```
<num> Number: 301
<title> International Organized Crime

<desc> Description:
Identify organizations that participate in international criminal
activity, the activity, and, if possible, collaborating organization
and the countries involved.

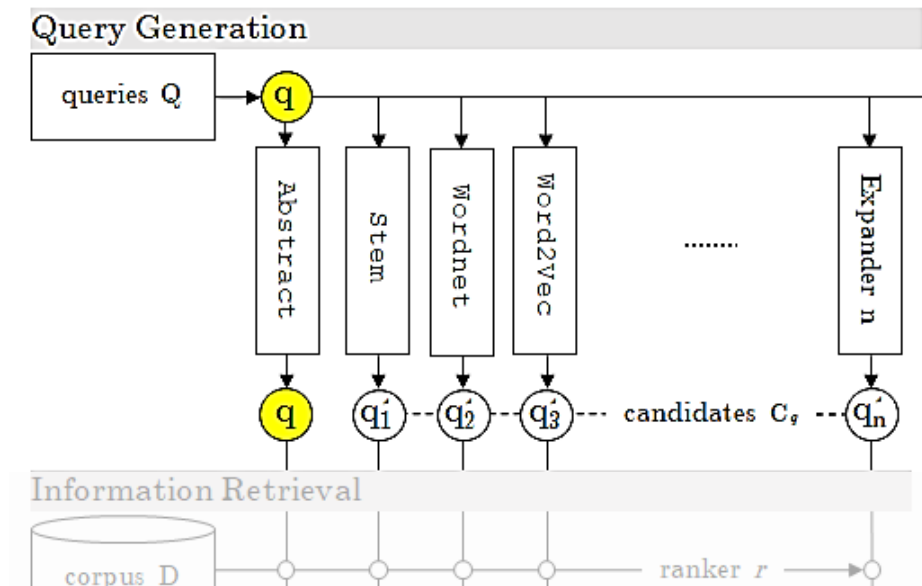
<narr> Narrative:
A relevant document must as a minimum identify the organization and
type of illegal activity (e.g., Columbian cartel exporting cocaine).
Vague references to international drug trade without identification
the organization(s) involved would not be relevant.
```

```
301 0 FBIS3-10082 1
301 0 FBIS3-10169 0
301 0 FBIS3-10243 1
301 0 FBIS3-10319 0
301 0 FBIS3-10397 1
301 0 FBIS3-10491 1
301 0 FBIS3-10555 0
301 0 FBIS3-10622 1
301 0 FBIS3-10634 0
301 0 FBIS3-10635 0
301 0 FBIS3-10721 1
301 0 FBIS3-10805 1
301 0 FBIS3-10910 1
301 0 FBIS3-10937 1
301 0 FBIS3-10938 0
301 0 FBIS3-11003 1
301 0 FBIS3-11028 0
301 0 FBIS3-11095 1
301 0 FBIS3-11099 1
301 0 FBIS3-11210 1
301 0 FBIS3-11318 0
301 0 FBIS3-11405 1
301 0 FBIS3-11502 1
```

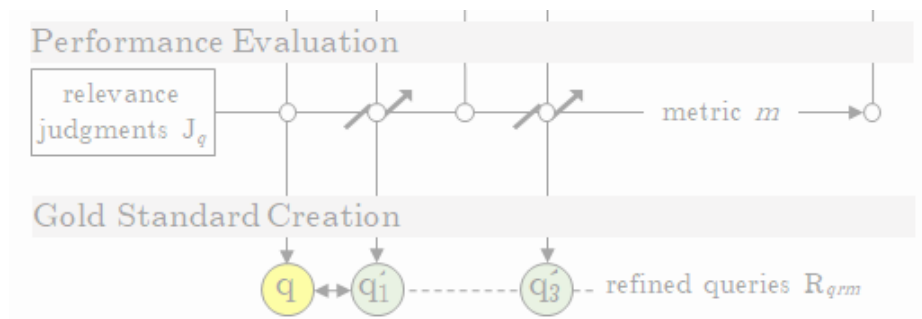
Supervised Query Refinement

Automatic creation of dataset!

- ReQue



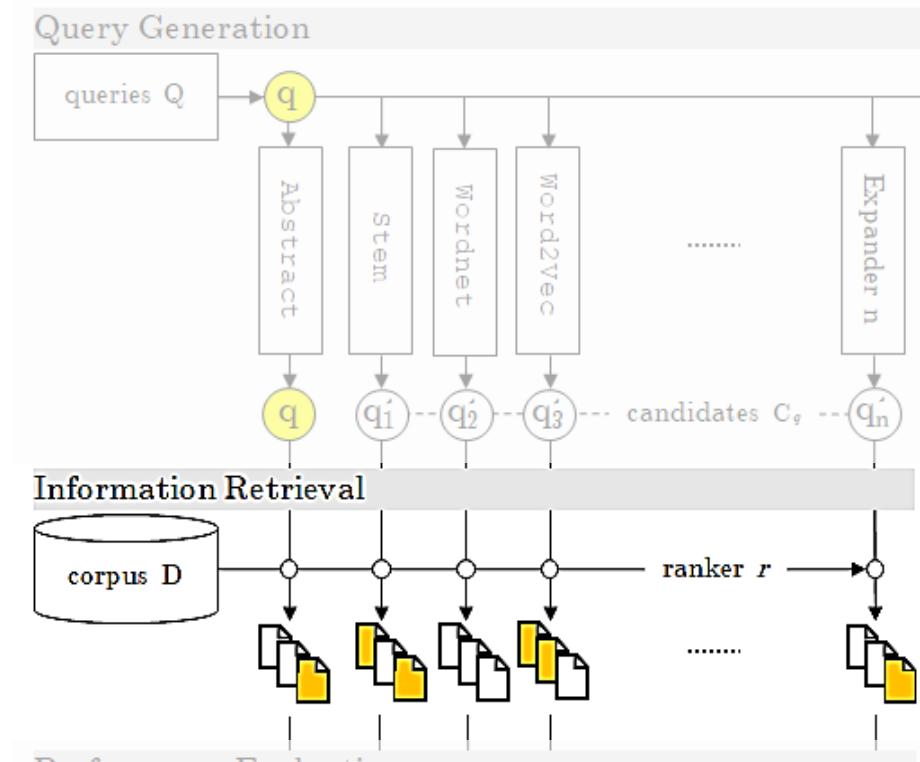
Create collection of variations to q



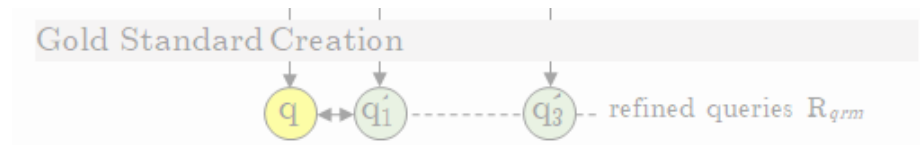
Supervised Query Refinement

Automatic creation of dataset!

- ReQue



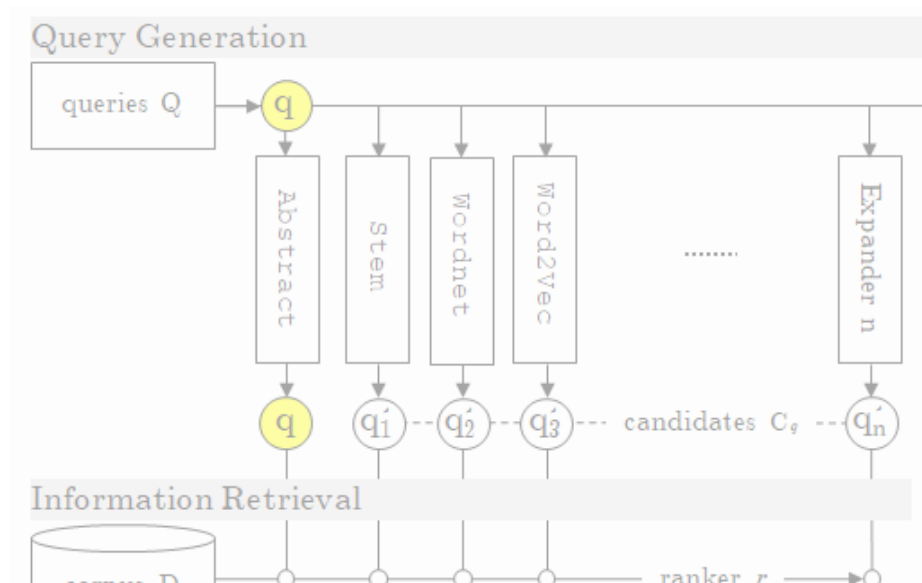
Try to retrieve the relevant information



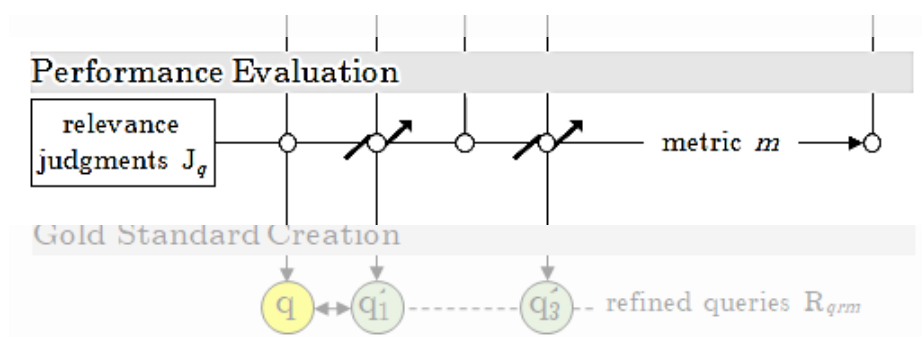
Supervised Query Refinement

Automatic creation of dataset!

- ReQue



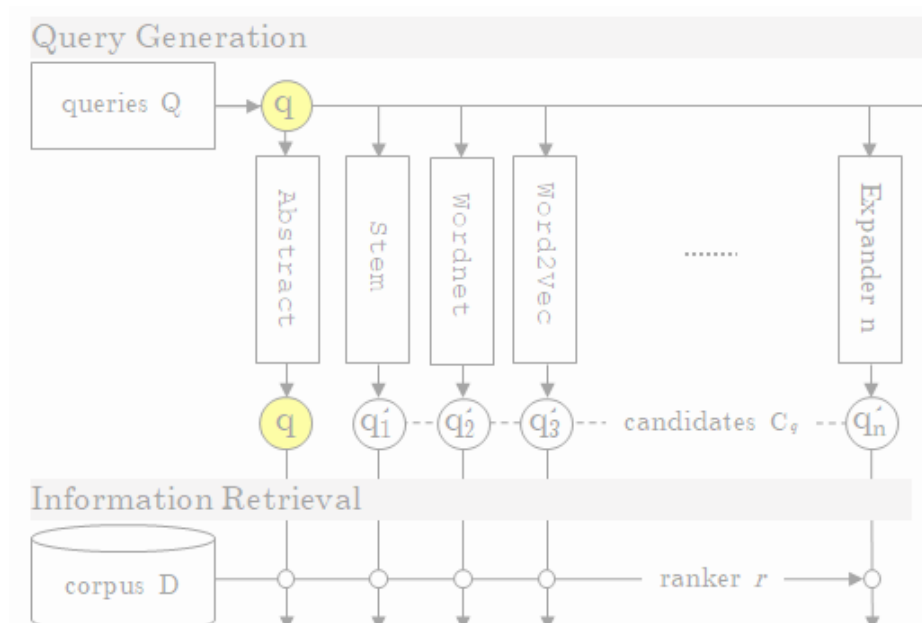
Evaluate the retrieved information



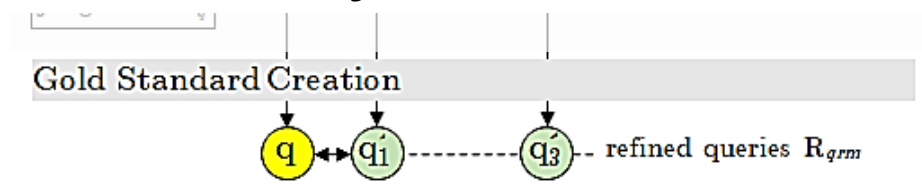
Supervised Query Refinement

Automatic creation of dataset!

- ReQue



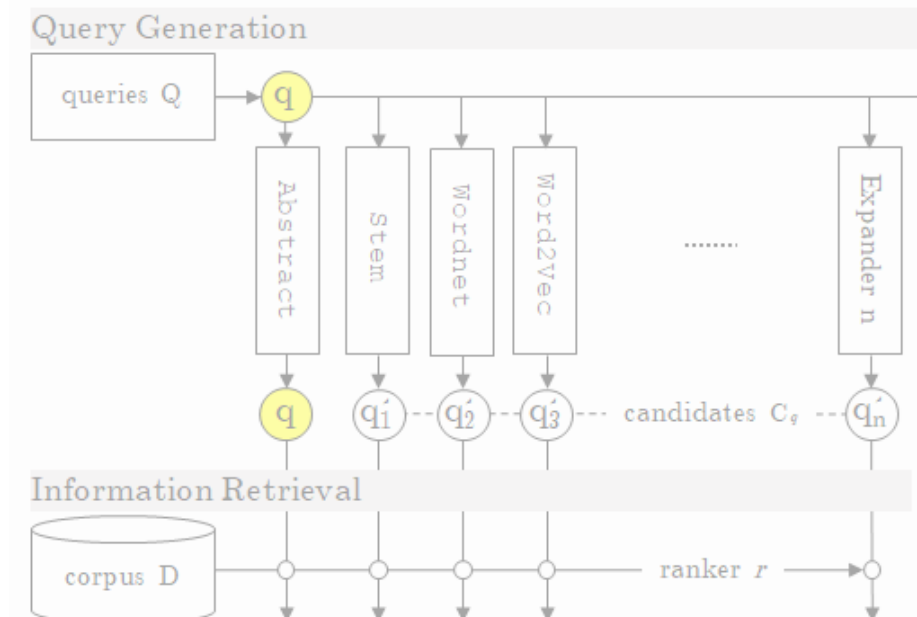
Those that increase the performance are the refined queries!
There may be more than 1.



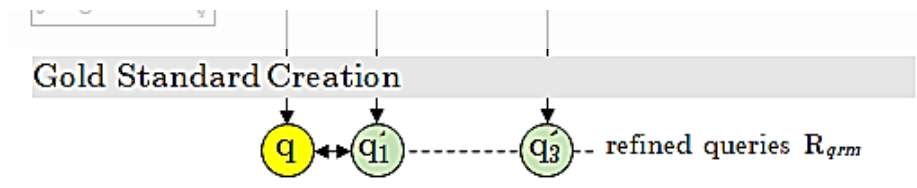
Supervised Query Refinement

Automatic creation of dataset!

- ReQue



The improvement for q' is guaranteed!
We should train the models on this dataset.



Unsupervised Query Refinement

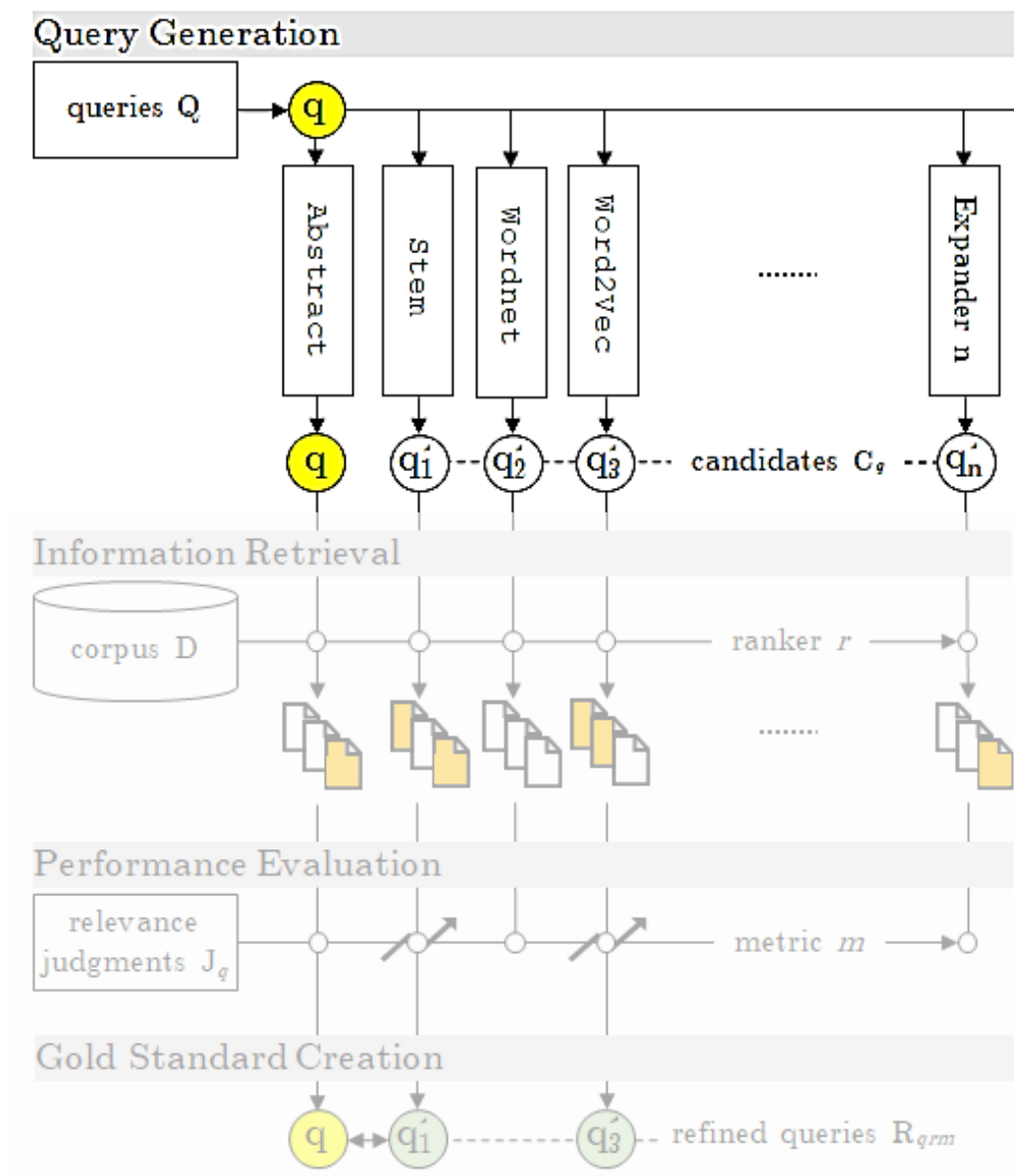
Creating collection of variations to q

- Brute-force change to q
- Guided change to q

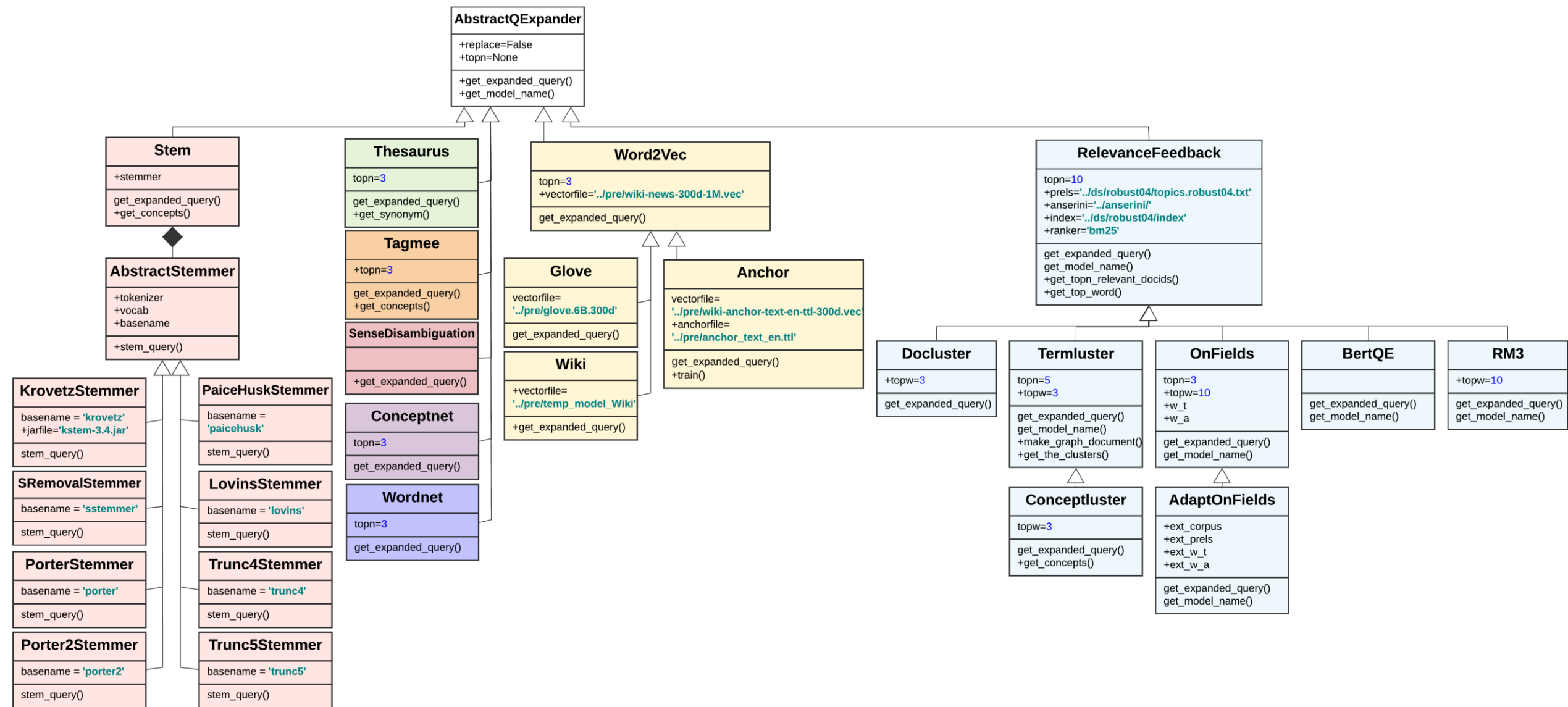
ReQue

Unsupervised Query Refinement

- Stem
- Thesaurus
- n-gram LM
- Neural LM
- ...



ReQue



Global Analysis

Local Analysis