

# Intelligent Code Example Search Engine: Automatic Code Example Recommendation

---

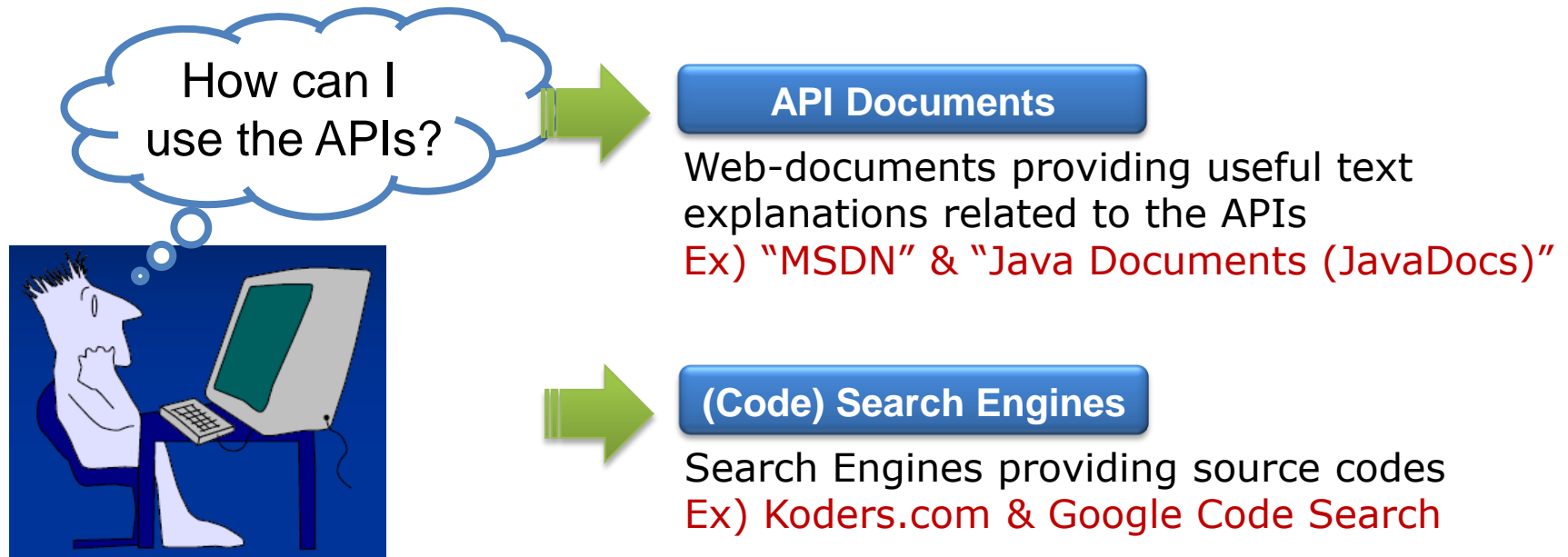
Jinhan Kim, IDS Lab, POSTECH

**POSTECH**

# Motivation

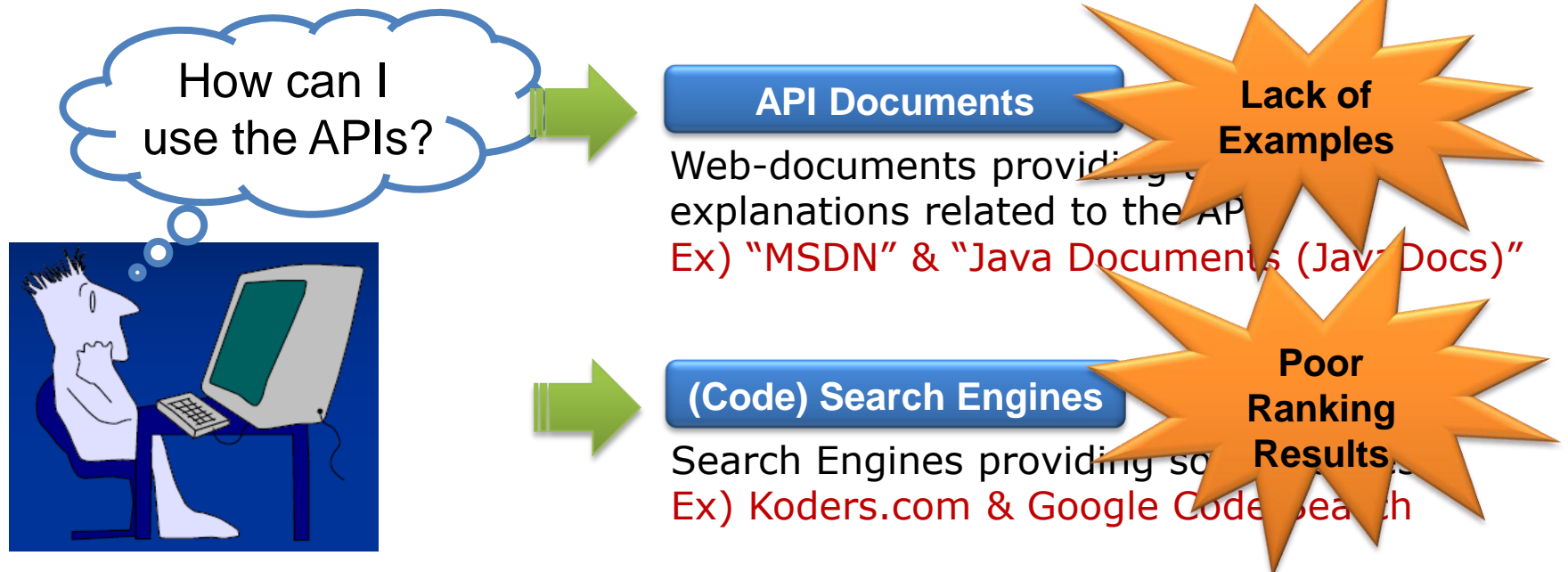
---

- Application Programming Interfaces (APIs)
  - Pre-defined useful functions  
Ex) `System.out.print("Hello world")` in Java: print "Hello world"
  - We can use APIs **without knowing the detailed background information**
- Developers usually use APIs when developing software



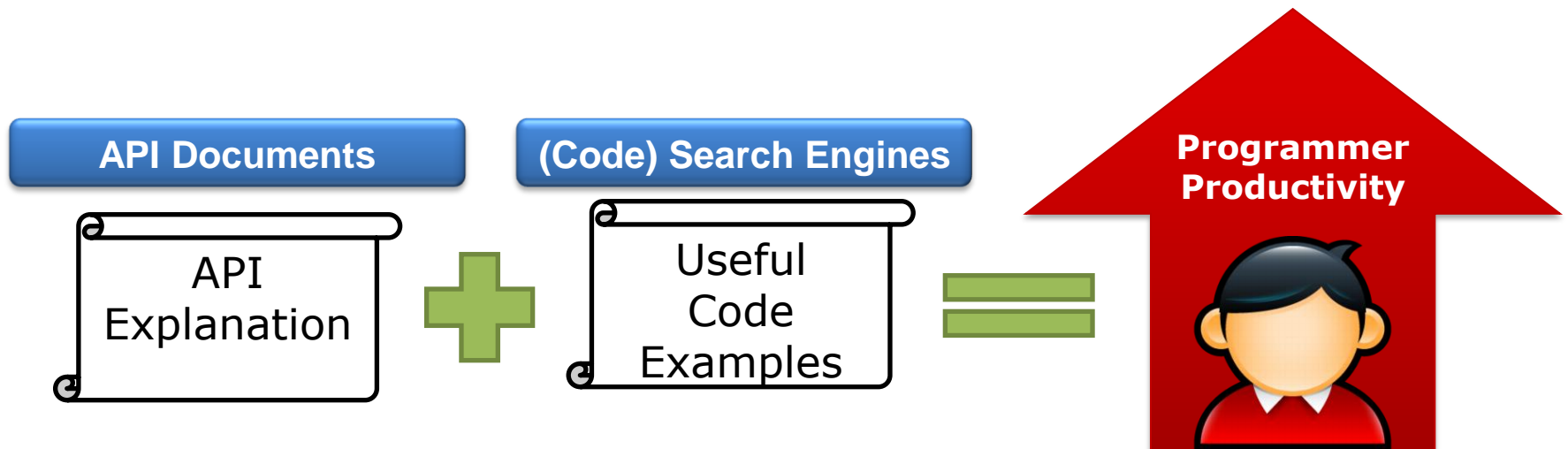
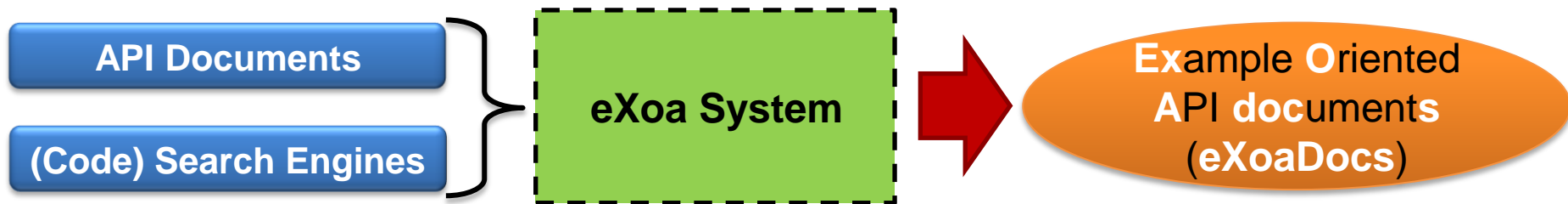
# Motivation

- Application Programming Interfaces (APIs)
  - Pre-defined useful functions  
Ex) `System.out.print("Hello world")` in Java: print "Hello world"
  - We can use APIs **without knowing the detailed background information**
- Developers usually use APIs when developing software

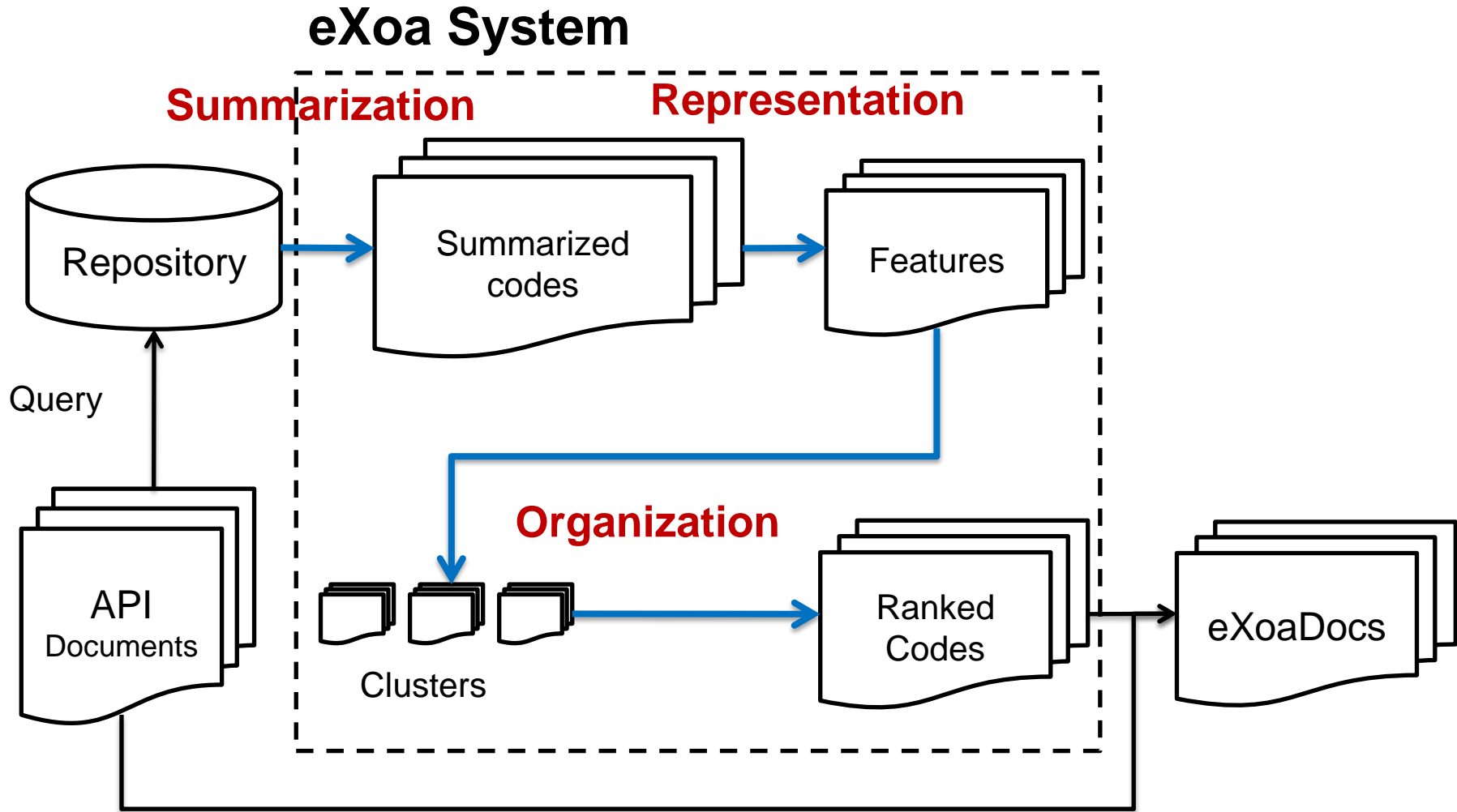


# Goal

- To address this problem...
  - Propose fully automated techniques that extract API usage examples from Code Search Engines
  - Embed them into API documents



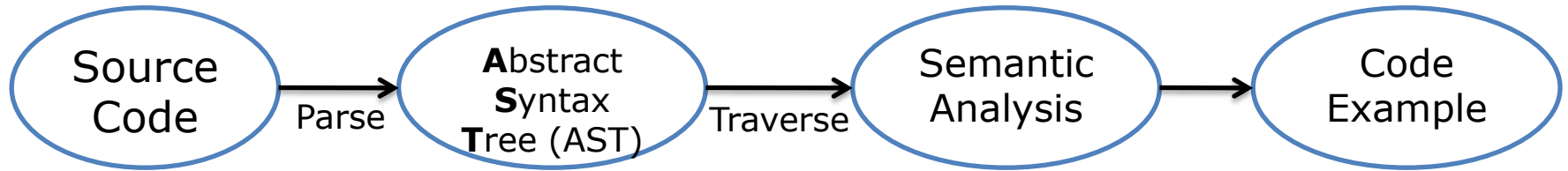
# Overall System Architecture



An Overview of System Architecture

# Modules

## □ Summarization



## □ Representation

- Approximate semantic features



## □ Organization

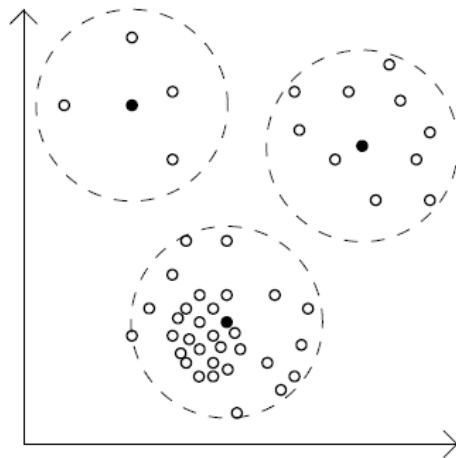
- Divide summarized code examples into clusters
- Compute score of each code example
  - Score Functions
    - Representativeness( $R_p$ ): Similarity to the center of cluster
    - Conciseness( $C_c$ ): line of code
    - Correctness( $C_r$ ): class and argument type checking
  - Aggregated Score Function

$$F = \text{Agg}(R_p + C_c + C_r)$$

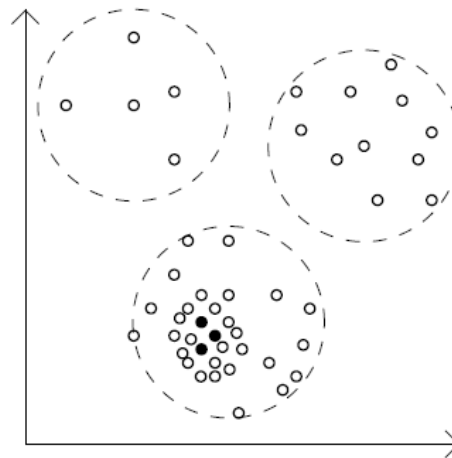
# Extend Organization Approaches

---

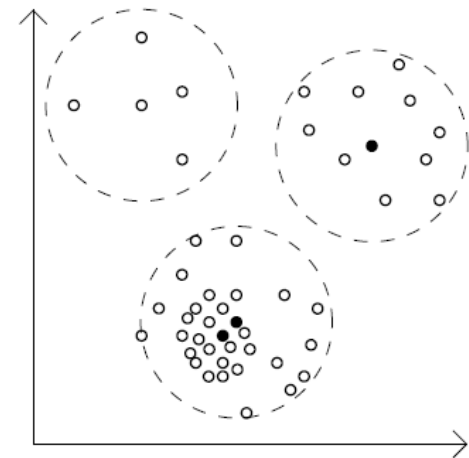
- ❑ Clustering-based Approach (Previous approach)
  - Cover diverse usage type but may include outlier types
- ❑ Ranking-based Approach
  - Consider different probabilities of code intents but may return skewed results
- ❑ Hybrid Approach



(a) eXoaCluster



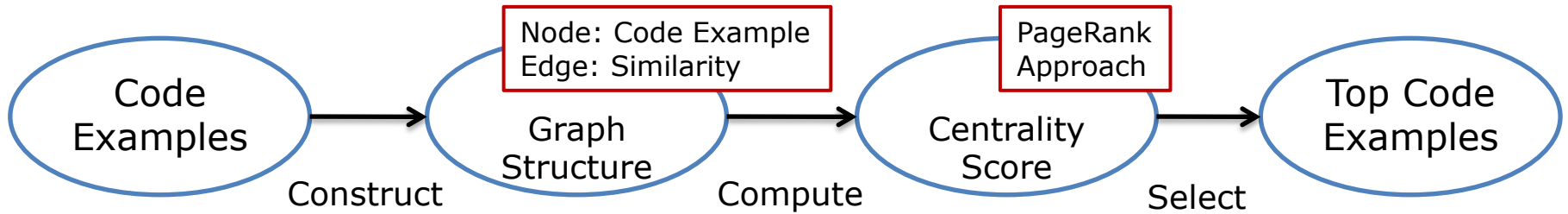
(b) eXoaRank



(c) eXoaHybrid

# Extend Ranking Approaches

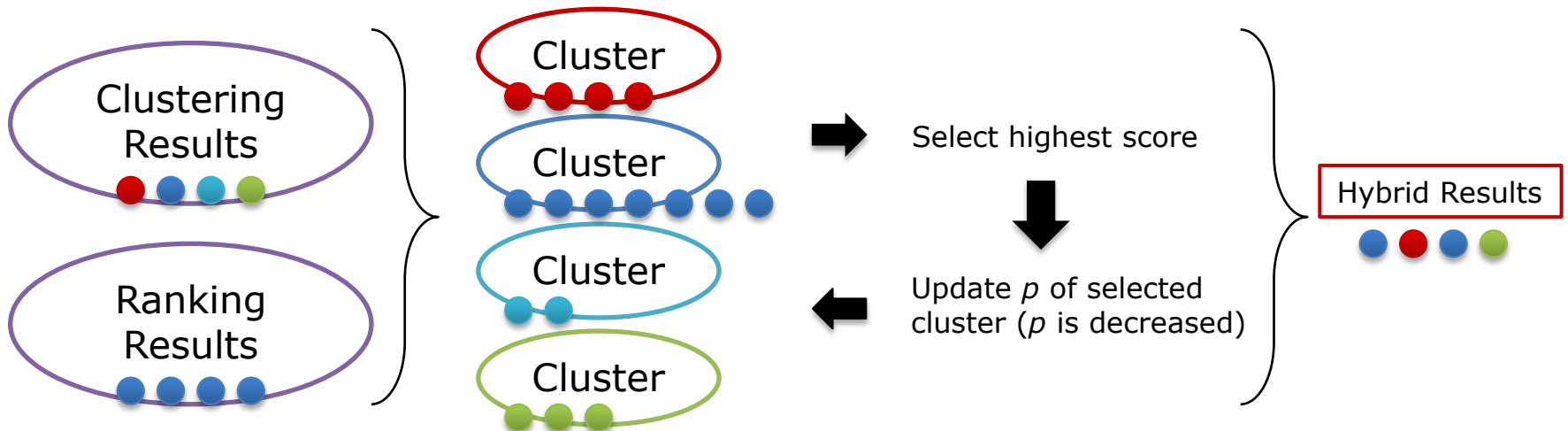
## Ranking-based Approach



## Hybrid Approach

- Balance between Clustering-based and Ranking-based

- Score:  $p$  of cluster \*  $p$  of code example





# Sample Result

(a) **Popularity**: A blue circle highlights the 'clone' method in the 'Method Summary' section, with a callout bubble containing the text 'Creates and returns a copy of this object.'

(b) **Code example**: A blue circle highlights the 'format(parsed)' code example, with a callout bubble containing the text 'format(parsed)'. Another blue circle highlights the 'format' method signature, with a callout bubble containing the text 'format'.

(c) **User feedback**: A blue circle highlights the navigation icons (back, forward, search, etc.) at the bottom of the page.

The screenshot shows the Java API documentation for `java.lang.String`. The 'Method Summary' section lists several methods, including `clone()`, `format(Object...)`, `formatToCharacterIterator(Object)`, `parseObject(String source)`, and `parseObject(String source, ParsePosition pos)`. The 'Method Detail' section provides the signature and description for `format(Object...)`, including its parameters, return value, and examples. The 'Code example' section shows the `format(parsed)` method call.

End

---

Question??