

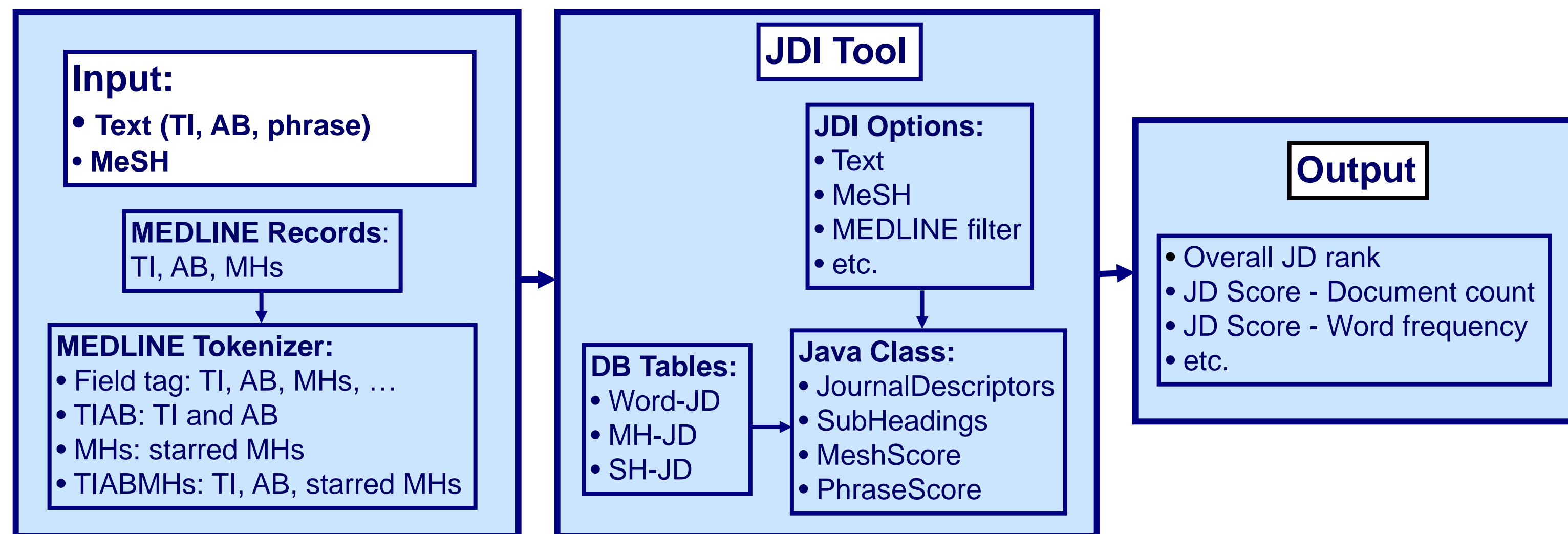
Text Categorization

<http://specialist.nlm.nih.gov/tc>



Journal Descriptor Indexing

122 JDs from NLM's List of Serials Indexed for Online Users (LSI)
Example: Journal of Pediatric Surgery - Pediatrics; Surgery



Semantic Type Indexing

135 STs in the Semantic Network in NLM's UMLS. One UMLS concept is assigned to one or more STs.

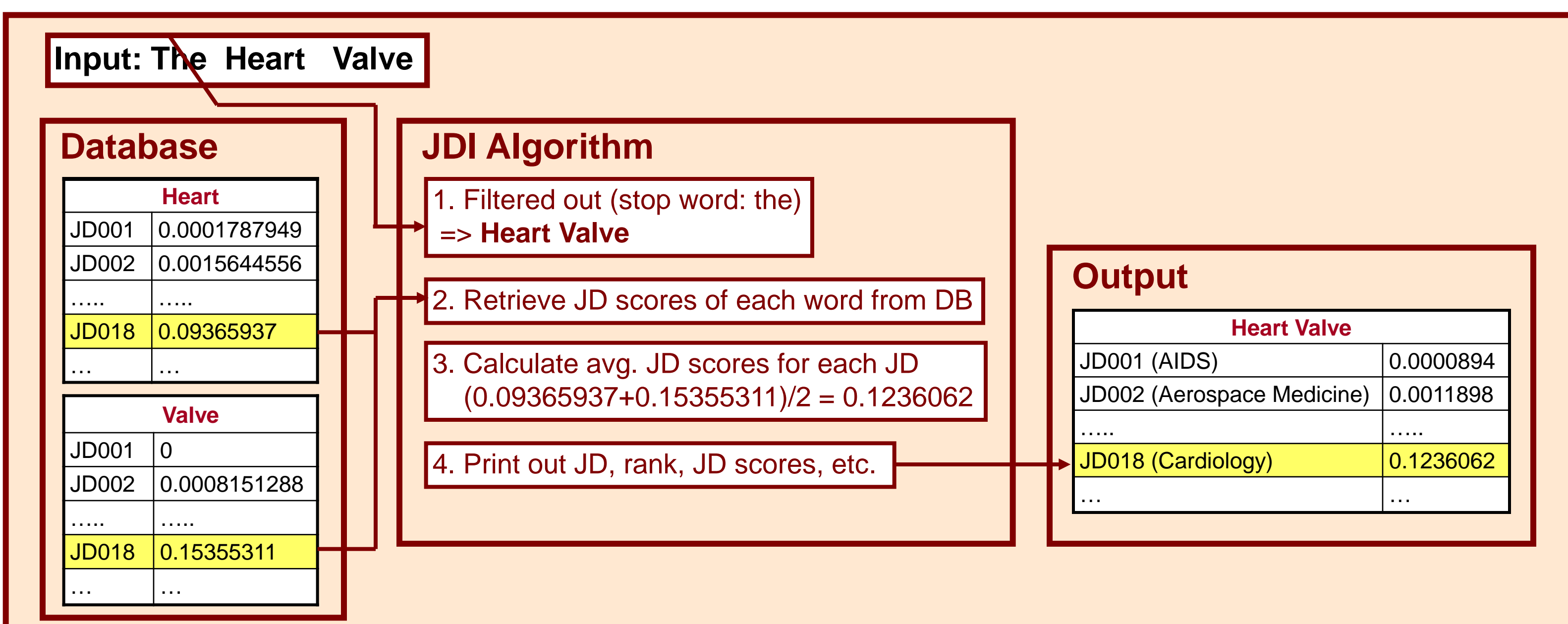
Input: Race, ethnicity, culture, and disparities in health care

Output:

--- ST scores and rank based on document count for word ---

popg
0|0.8373|popg|Population Group
1|0.7722|socb|Social Behavior
2|0.7591|aggp|Age Group
3|0.7385|idcn|Idea or Concept
4|0.7385|shro|Self-help or Relief Organization
5|0.7272|famg|Family Group
6|0.7232|orgt|Organization
7|0.7086|inbe|Individual Behavior
8|0.6965|gora|Governmental or Regulatory Activity
9|0.6905|edac|Educational Activity

JDI - Example



Application: Word Sense Disambiguation

Title: Race, ethnicity, culture, and disparities in health care
where "culture" has three UMLS concepts/Semantic Type mapping from MetaMap with same score:

- Anthropological Culture - Idea or Concept
- Cultural - Functional Concept
- Laboratory Culture - Laboratory Procedure

Apply STI to disambiguate "culture" in the above title.

Race, ethnicity, culture, and disparities in health care		
popg	0.8373	Population Group
socb	0.7722	Social Behavior
aggp	0.7591	Age Group
idcn	0.7385	Idea or Concept
...
ftcn	0.4647	Functional Concept
...
lbpr	0.3749	Laboratory Procedure
....

Select the ST with highest score, IDCN