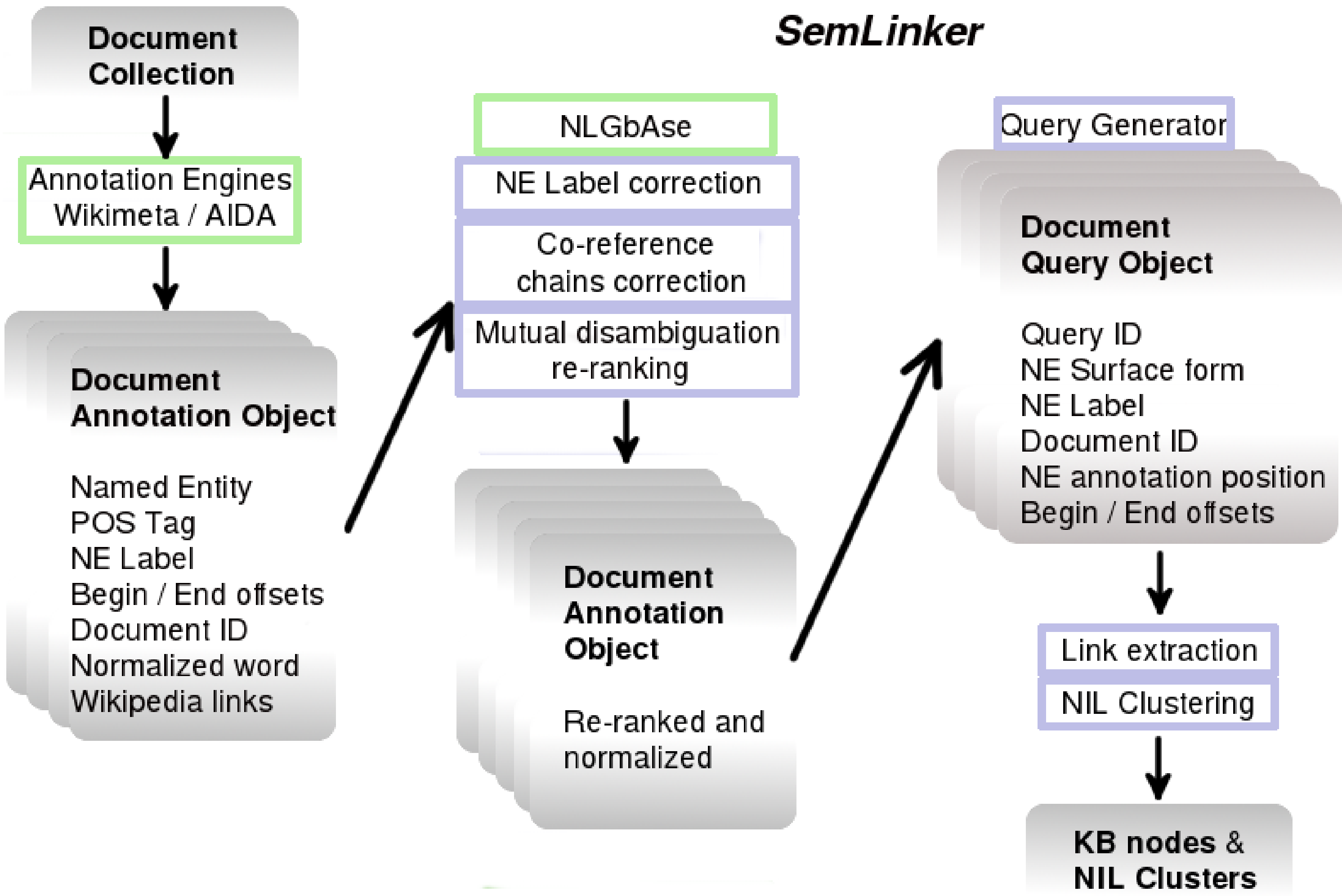


## The SemLinker System

**SemLinker** utilizes **two generic annotation engines** to perform entity discovery and linking. The linking is improved by a **re-ranking process** based on **mutual relations** between all entities found in the document. **SemLinker** merges the results from different annotators according to various heuristics.

## System Architecture



## System Components

1. Entity Extraction
2. NE Label Correction
3. Co-reference Correction
4. Query Generation
5. Mutual Disambiguation
6. Link Extraction
7. Query Merger
8. Clustering

## System Resources

- Annotation engines: **Wikimeta** and **AIDA**
- Corpus-based resources: **NLGbAse** and **Wikipedia dump**

## Annotation Object

Position	Word	POS	NE <sub>Label</sub>	Offset	DocID	Word <sub>Norm</sub>	Semantic Link
Wikimeta							
202	Chelsy	NAM	PERS.HUM	1235-1240	eng-NG-31-109383-9018012	chelsy	<a href="http://en.wikipedia.org/wiki/Chelsy_Davy">http://en.wikipedia.org/wiki/Chelsy_Davy</a>
203	Davy	FNAM	PERS.HUM	1242-1245	eng-NG-31-109383-9018012	davy	
204	in	IN	UNK	1247-1248	eng-NG-31-109383-9018012	in	
205	London	NAM	LOC.ADMI	1250-1255	eng-NG-31-109383-9018012	london	<a href="http://en.wikipedia.org/wiki/London">http://en.wikipedia.org/wiki/London</a>
AIDA							
18	Georgia	NNP	LOC	294-300	eng-NG-31-109034-9006695	georgia	<a href="http://en.wikipedia.org/wiki/Georgia%28U.S.state%29">http://en.wikipedia.org/wiki/Georgia%28U.S.state%29</a>
19	in	IN	UNK	302-303	eng-NG-31-109034-9006695	in	
20	the	DT	UNK	305-307	eng-NG-31-109034-9006695	the	
21	Caucasus	NNP	LOC	309-316	eng-NG-31-109034-9006695	caucasus	<a href="http://en.wikipedia.org/wiki/Caucasus">http://en.wikipedia.org/wiki/Caucasus</a>

## Query Generation

Query ID	NE <sub>surface_form</sub>	NE <sub>Label</sub>	DocID	NE <sub>position</sub>	Offset
Wikimeta					
EDL14_ENG_CSFG_0003349	Chelsy Davy	PER	eng-NG-31-109383-9018012	202	1235-1245
EDL14_ENG_CSFG_0003350	London	GPE	eng-NG-31-109383-9018012	205	1250-1255
AIDA					
EDL14_ENG_CSFG_0001291	Georgia	GPE	eng-NG-31-109034-9006695	18	294-300
EDL14_ENG_CSFG_0001292	Caucasus	GPE	eng-NG-31-109034-9006695	21	309-316

<query id="EDL14\_ENG\_CSFG\_0003349">  
 <name>Chelsy Davy</name>  
 <docid>eng-NG-31-109383-9018012</docid>  
 <beg>1235</beg>  
 <end>1245</end>  
</query>

<query id="EDL14\_ENG\_CSFG\_0003350">  
 <name>London</name>  
 <docid>eng-NG-31-109383-9018012</docid>  
 <beg>1250</beg>  
 <end>1255</end>  
</query>

<query id="EDL14\_ENG\_CSFG\_0001291">  
 <name>Georgia</name>  
 <docid>eng-NG-31-109034-9006695</docid>  
 <beg>294</beg>  
 <end>300</end>  
</query>

<query id="EDL14\_ENG\_CSFG\_0001292">  
 <name>Caucasus</name>  
 <docid>eng-NG-31-109034-9006695</docid>  
 <beg>309</beg>  
 <end>316</end>  
</query>