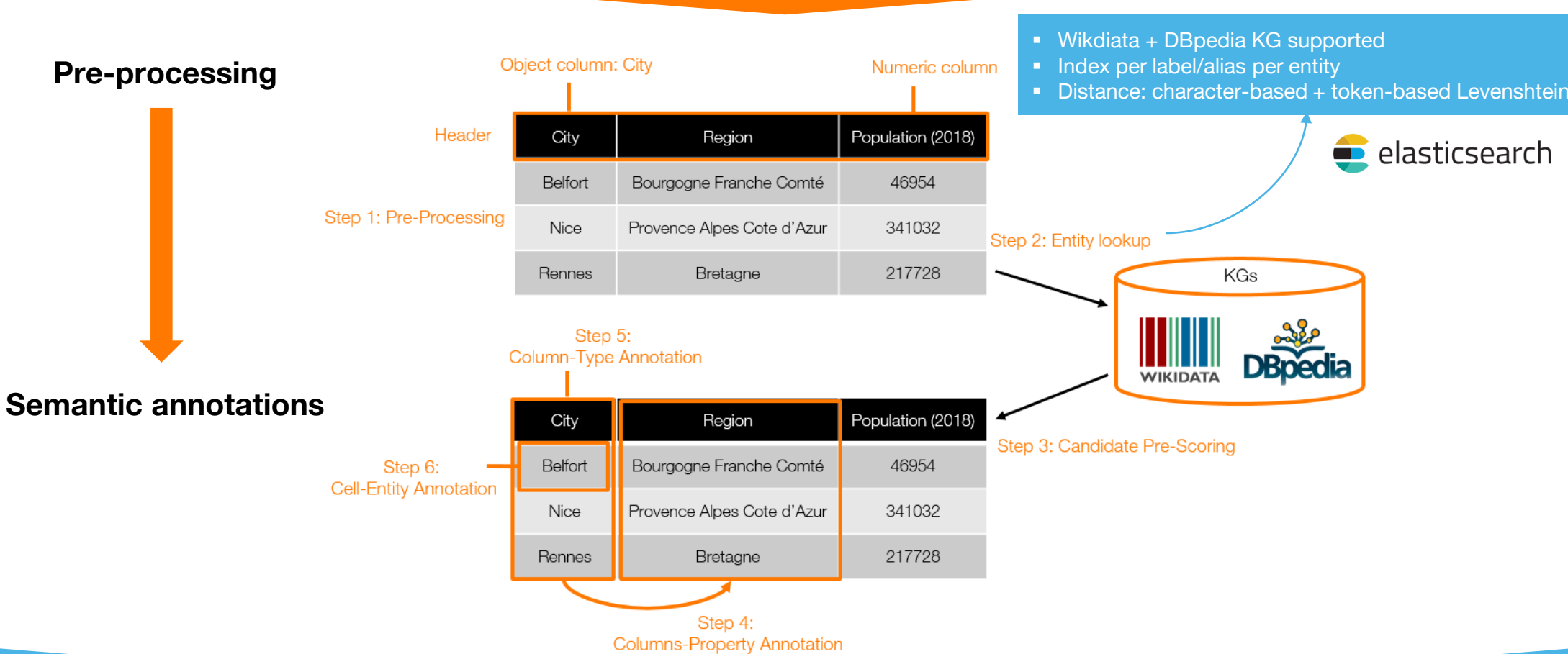
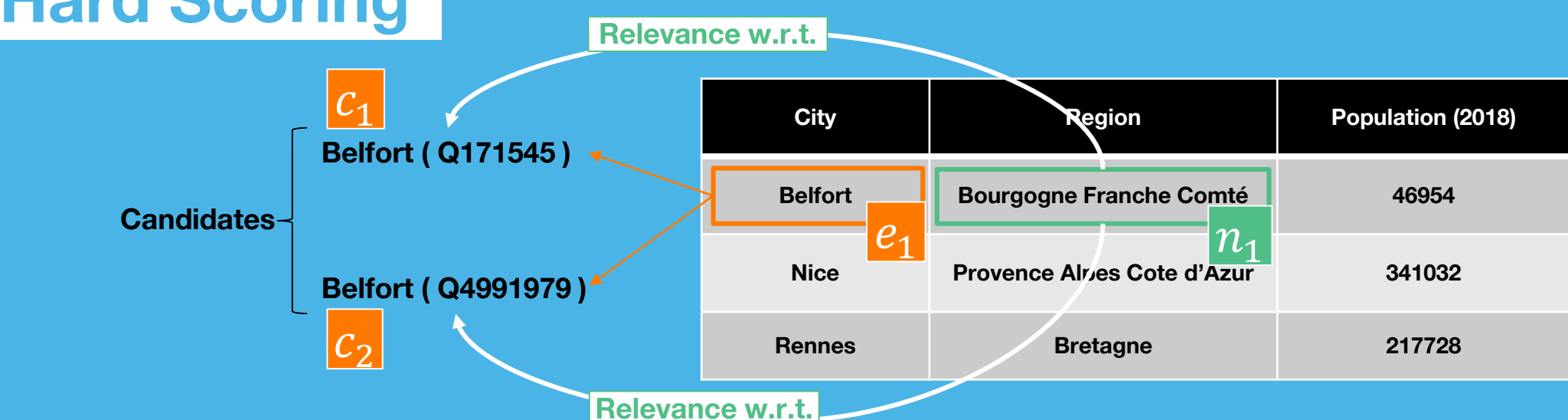


DAGOBDAH: Table and Graph Contexts For Efficient Semantic Annotation Of Tabular Data

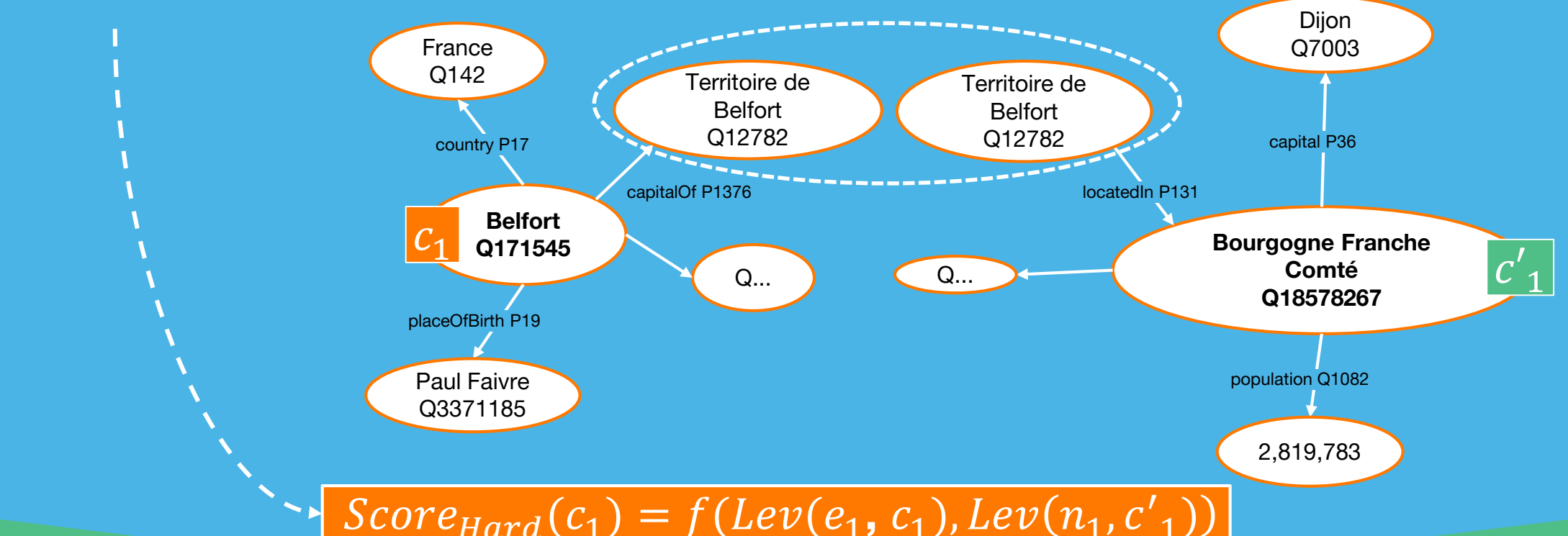
Annotation Workflow



Hard Scoring

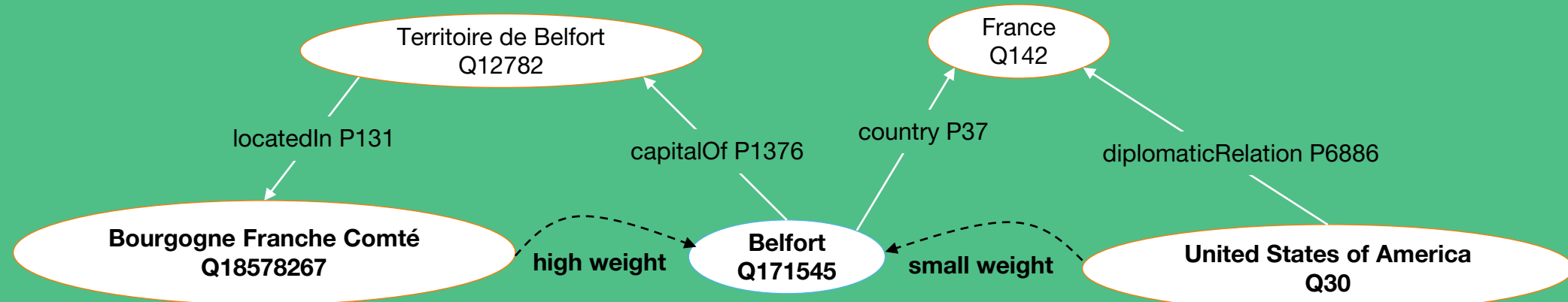


Context score of a candidate entity means quantifying the relevance of neighboring cells



Soft Scoring

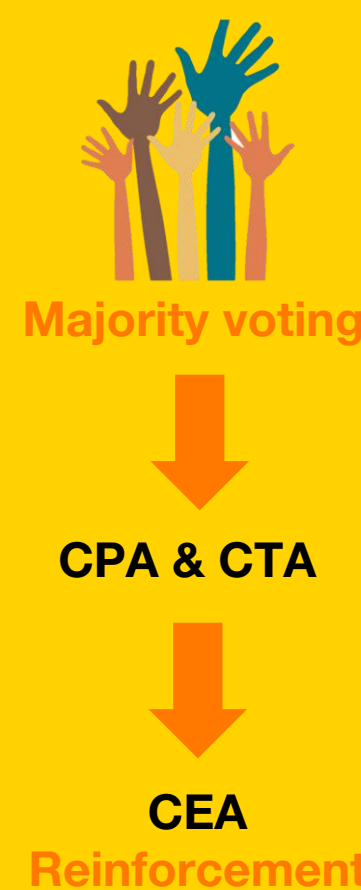
Each neighboring cell does not have an equal contribution



generality(Q12782) << generality(Q142) → Q18578267 more informative

$$Score(c_1) = g(Score_{Soft}, Score_{Hard})$$

Annotation



SemTab @ ISWC 2021 Results

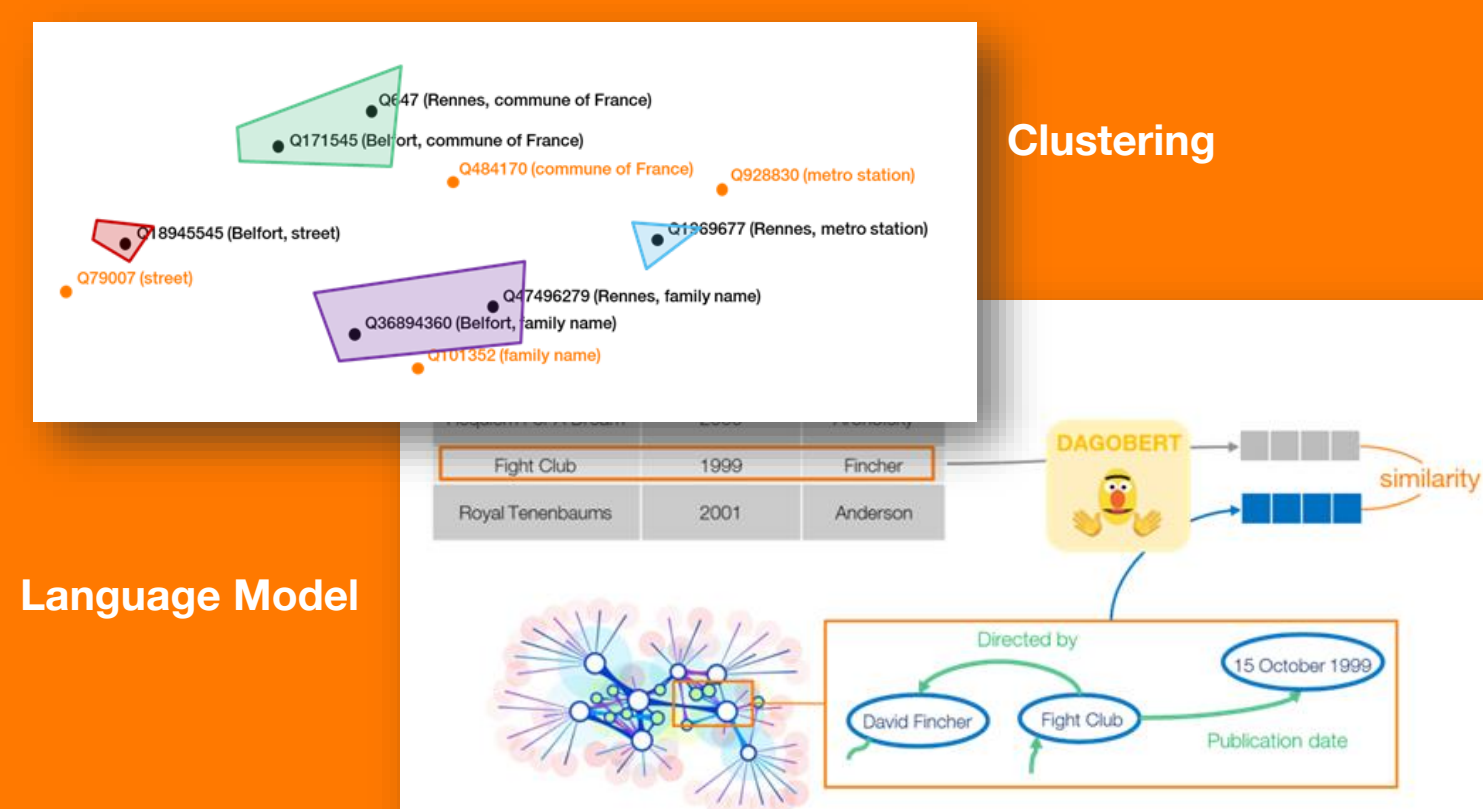
Task	Setting	CTA		CEA		CPA	
		F1	Precision	F1	Precision	F1	Precision
Round 1-WDTables	2*	0.832	0.832	0.923	0.923	x	x
Round 1-DBPTables	2*	0.422	0.424	0.945	0.946	x	x
Round 2-BioTables	3	0.916	0.916	0.970	0.970	0.899	0.899
Round 2-HardTables	3	0.976	0.976	0.975	0.976	0.996	0.996

Hard Scoring + Soft Scoring → Best results

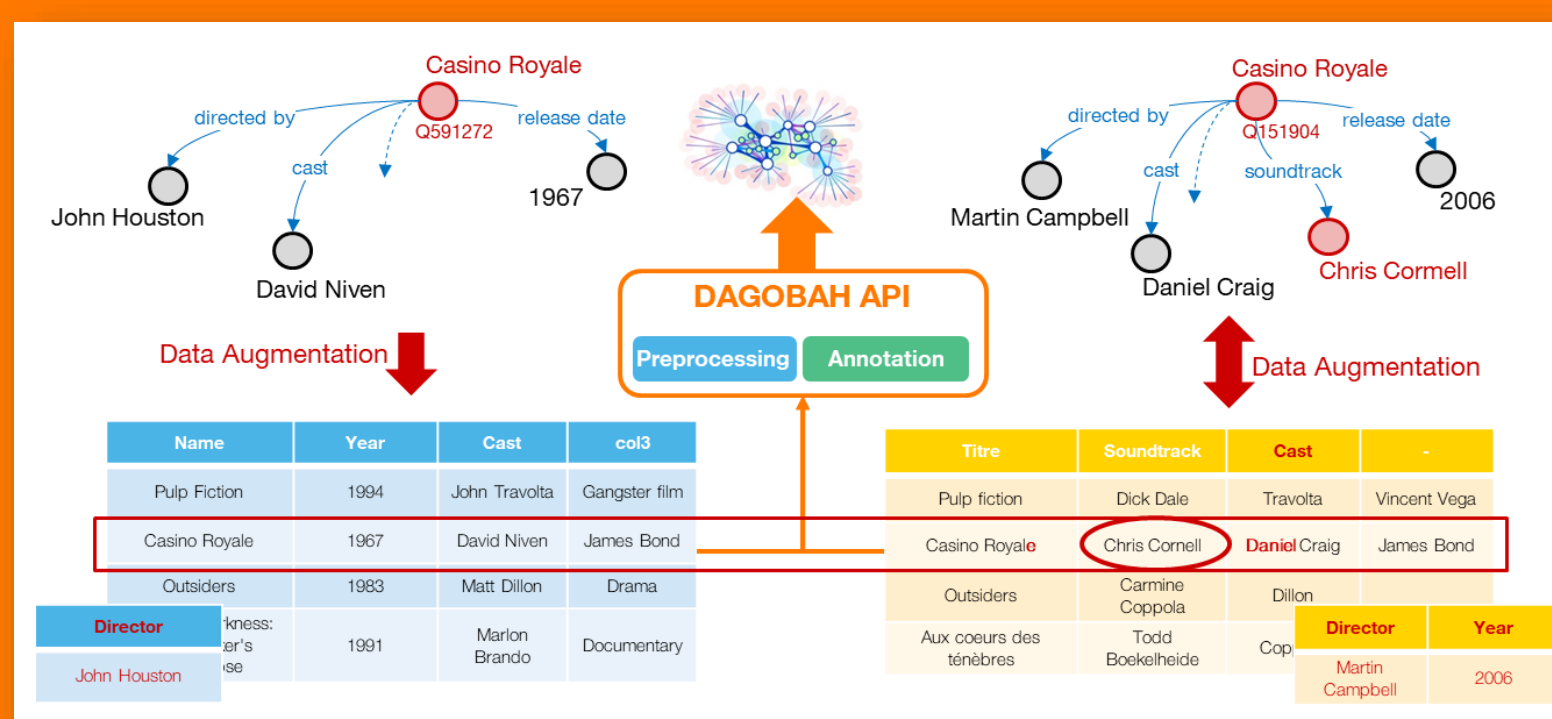
Team_DAGOBDAH 65.0
Top-ranked system (Rounds 1 + 2) (with few submissions)

Future Works

Annotation enhancement (dictionaries, embeddings)



Data augmentation



Hardcore corpus

Team	Est.	West	Total	Est.	West	Total
Monopoly / Los Angeles Lakers	0	30	31K	0	9	9
Bayern Munich	21	9	30	21	9	30
Spartan National / Philadelphia 76ers	21	9	30	21	9	30
2020	21.6	not available	not available	not available	not available	not available
2019	20.6	not available	not available	not available	not available	not available
2018	20.5	not available	not available	not available	not available	not available
2017	20.9	not available	not available	not available	not available	not available
2016	24.3	not available	not available	not available	not available	not available
2015	37.2507196	1.957191	2.1	4.35047	174.5014	178.8516
2014	40.7407102	1.8 - 1.9101	1.94 - 2.0	1.94 - 2.0	1.94 - 2.0	1.94 - 2.0
2013	40.3702104	1.65 - 1.8175	1.81 - 1.9	1.81 - 1.9	1.81 - 1.9	1.81 - 1.9
2012	39.4602104	1.610	1.70	1.70	1.70	1.70
2011	37.919	1.354	1.56	1.56	1.56	1.56
2010	41.000	1.1207	1.32	1.32	1.32	1.32