					Language coverage		Approach	Source type	Metadata	Metadata structures	Required information	Limitations	Data gathering
					Includes	Excludes							
Stocker, M., Seaborne, A., Bernstein, A., Kiefer, C., and Reynolds, D. 2008. SPARQL basic graph pattern optimization using selectivity estimation. In Proceeding of the 17th international Conference on World Wide Web (Beijing, China, April 21 - 25, 2008). WWW '08. ACM, New York, NY, 595-604.				1	BGP optimization	OPTIONAL, UNION, FILTER	static optimization before evaluation: - join ordering	single source oriented optimization	Use of heuristics: - without pre-computed statistics - with pre-computed statistics	Summary statistics: - single triple patterns - joined triple patterns use of histograms	Selectivity of patterns, and estimation of selectivity	Only main memory graph implementations	Random sampling from the full dataset
Harth, A., Hose, K., Karnstedt, M., Polleres, A., Sattler, K., and Umbrich, J. 2010. Data summaries for on-demand queries over linked data. In Proceedings of the 19th international Conference on World Wide Web (Raleigh, North Carolina, USA, April 26 - 30, 2010). WWW '10. ACM, New York, NY, 411-420				2	Conjunctive SPARQL BGPs		Source selection, join source selection, discard unnecessary sources Source ranking	Distributed data - linked open data	Prefetched data summaries	QTree, combination of histogram + RTree MBB Bucket [S.low, S.hi][P.low,P.hi][O. low,O.hi]	Count of data items	Focused on source selection, no SPARQL endpoints usage	Crawling to fill the QTree. Alternatively start w empty QTree and fill on every query
Schmidt, M., Meier, M., and Lausen, G. 2010. Foundations of SPARQL query optimization. In Proceedings of the 13th international Conference on Database theory (Lausanne, Switzerland, March 23 - 25, 2010). L. Segoufin, Ed. ICDT '10. ACM, New York, NY, 4-3				3	OPTIONAL UNION FILTER		complexity analysis SPARQL rewriting rules with equivalences, algebraic optimization and semantic optimization exploiting constraints	Single source, rewriting can be used in dsitributed approaches	no	No	No	only rewriting rules heuistics, statistics, access paths, cost models out of scope	No
Olaf Görlitz and Steffen Staab. Federated Data Management and Query Optimization for Linked Open Data. Book chapter. In: Athena Vakali & Lakhmi Jain(eds.) Web Data Management Trails, Springer, to appear.				4	BGPs	OPTIONAL, UNION	Join order optimization using dynamic programming. Use of semi-join operators.	Federated SPARQL sources	Dynamic programming using a Cost Model Definition of a set of cost function for operators, based on cardinality and selectivity	Not specified	Estimated cardinality and selectivity of patterns	Theoretical work, no implementation or evaluation	Several approaches explained, not settled to any.
Quilitz, B. and Leser, U. 2008. Querying distributed RDF data sources with SPARQL. In Proceedings of the 5th European Semantic Web Conference on the Semantic Web: Research and Applications (Tenerife, Canary Islands, Spain, June 01 - 05, 2008). S. Bechhofer, M. Hauswirth, J. Hoffmann, and M. Koubarakis, Eds. Lecture Notes In Computer Science. Springer-Verlag, Berlin, Heidelberg, 524-538.				5	FBGPs	GRAPH, UNION, OPTIONAL	Basic logical rewriting optimization Physical optimization, join reordering using Cost model, uses the statistics	Distributed SPARQL	Statistics provided in Source descriptions	No histograms, statistics provided as metadata for each source	total number of triples, selectivity of patterns for a predicate, number of triples with a predicate	Only basic optimization techniques. could be extended.	Statistics provided beforehand
Obermeier, P., Nixon, L.: A Cost Model for Querying Distributed RDF-Repositories with SPARQL. In: Proceedings of theWorkshop on Advancing Reasoning on theWeb: Scalability and Commonsense. CEUR Workshop Proceedings, vol. 350				6	Seemingly includes OPTIONAL and UNION		Estimates phyical costs of operations	Distributed query processing	Statistics, cardinalities	Not specified	Cardinalities, estimation of CPU instructions and estimation of disk access	Still to implement in a query processor	Recommended statistics beforehand, used to compute cost functions
Stuckenschmidt, H., Vdovjak, R., Houben, G., and Broekstra, J. 2004. Index structures and algorithms for querying distributed RDF repositories. In Proceedings of the 13th international Conference on World Wide Web (New York, NY, USA, May 17 - 20, 2004). WWW '04. ACM, New York, NY, 631-639				7	SeRQL	Previous to SPARQL standardization	Cost model, uses cardinality Heuristics for Join ordering	Distributed RDF storage	Cardinality of relations	Source index hierarchies	Cardinalities	Purely theoretical Not SPARQL really Source selection oriented	Not specified, assumed to be given
Neumann, T. and Weikum, G. 2008. RDF-3X: a RISC-style engine for RDF. Proc. VLDB Endow. 1, 1 (Aug. 2008), 647-659.				8	Includes OPTIONAL and UNION, though hints at other specific techniques	No	Indexes over 6 permutation of triples Indexes compressed, claimed < size of data	Single source	Counters of frequent predicate sequences Paths are potential join Selectivity estimation	B+ sorted lexicographically Selectivity Histograms	Selectivity estimation is computed	Compression tradeoff, redundancy indexes	Frequent path mining algorithm