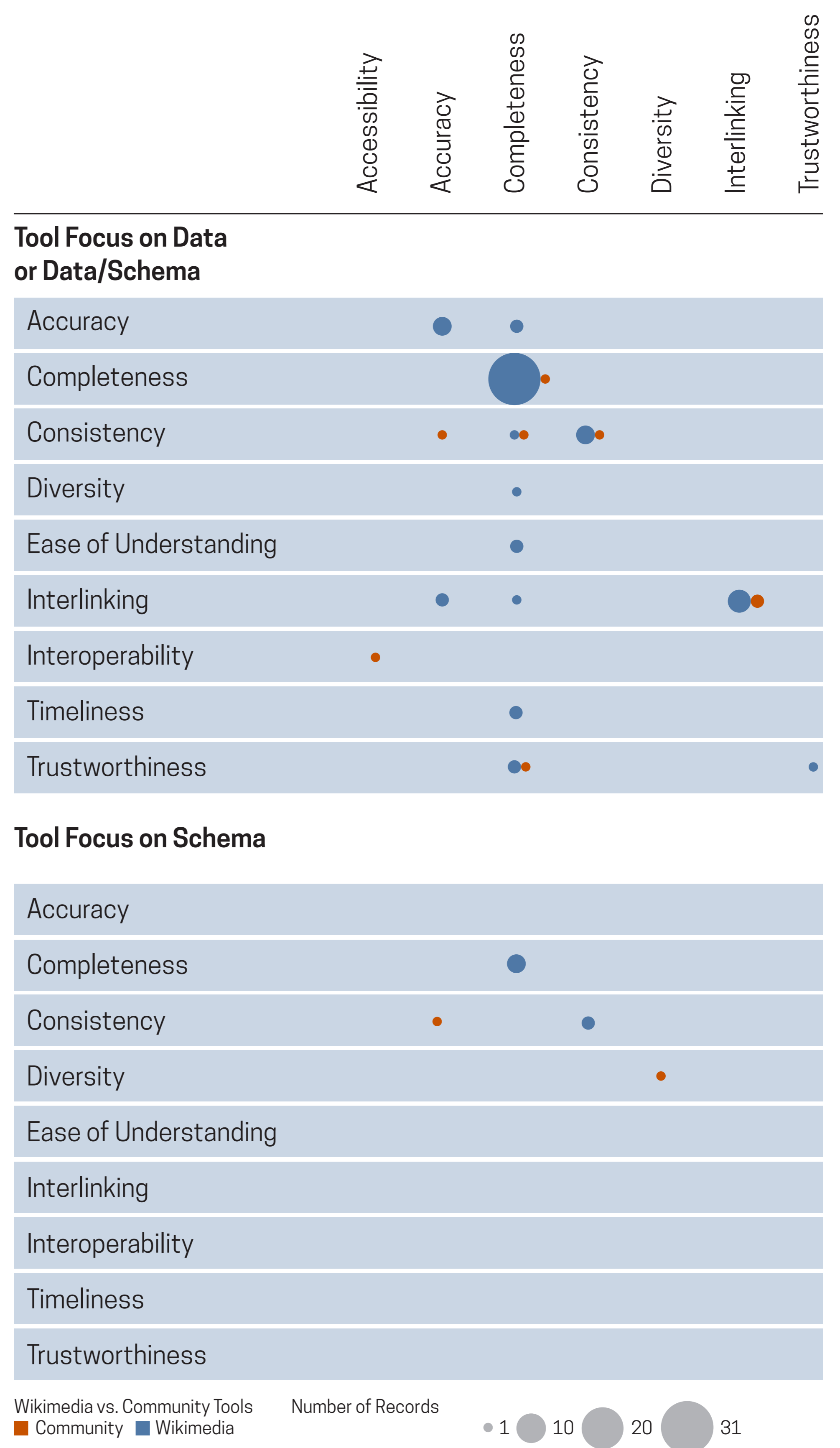# Data Quality in Wikidata Status Quo

Ensuring data quality is of utmost importance, as the goal of Wikidata is to "give more people more access to knowledge" and therefore, the data needs to be "fit for use by data consumers" (Wang & Strong, 1996). The Wikidata community has already developed many tools that help to deal with data quality issues on Wikidata. Moreover, the community started global discussions about relevant dimensions of data quality to establish a data quality framework for Wikidata. We reviewed this existing data quality framework and compared it, with current approaches. We identify dimensions that are not represented yet. Furthermore, when classifying existing tools provided by Wikimedia or developed by the community, we can see a focus on completeness.

		Wang & Strong (1996)	Zaveri et al. (2015)	Färber et. al. (2018)	Piscopo et al. (2016)	
Intrinsic	Believability		×			
	Objectivity		×			×
	Reputation/Trustworthiness		×		×	×
	Accuracy		×	×	×	×
	Consistency			×	×	×
	Conciseness			×		
Contextual	Value-added		×			
	Relevancy		×	×	×	
	Timeliness		×		×	×
	Completeness		×	×	×	×
	Appropriate amount of data		×	×		
Representational	Interpretability		×	×		
	Ease of understanding		×	×	×	×
	Representational consistency		×	×		
	Concise representation		×	×		
	Versaltility			×		×
	Interoperability				×	
Accessibility	Accessibility/Availability		×	×	×	
	Security		×			×
	Interlinking			×	×	
	Licensing			×	×	
	Perfomance			×		
Trust	Reputation			×		
	Believability			×		
	Verifiability			×		
	Objectivity			×		
Dynamicity	Currency			×		
	Volatility			×		
	Timeliness			×		

Tool Focus on Data or Data/Schema

Columns: Accessibility, Accuracy, Completeness, Consistency, Diversity, Interlinking, Trustworthiness

- Accuracy
- Completeness
- Consistency
- Diversity
- Ease of Understanding
- Interlinking
- Interoperability
- Timeliness
- Trustworthiness

Tool Focus on Schema

- Accuracy
- Completeness
- Consistency
- Diversity
- Ease of Understanding
- Interlinking
- Interoperability
- Timeliness
- Trustworthiness

Wikimedia vs. Community Tools
■ Community ■ Wikimedia

Number of Records
• 1 10 20 31

Färber, M., Bartscherer, F., Menne, C., Rettinger, A., 2018. Linked data quality of DBpedia, Freebase, OpenCyc, Wikidata, and YAGO. https://doi.org/10.3233/SW-170275

Wang, R.Y., Strong, D.M., 1996. Beyond Accuracy: What Data Quality Means to Data Consumers. https://doi.org/10.1080/07421222.1996.W11518099

Zaveri, A., Rula, A., Maurino, A., Pietrobon, R., Lehmann, J., Auer, S., 2015. Quality assessment for Linked Data: A Survey: A systematic literature review and conceptual framework. https://doi.org/10.3233/SW-150175

Piscopo, A., 2016. Requests for comment/Data quality framework for Wikidata. https://www.wikidata.org/wiki/Wikidata:Requests_for_comment/Data_quality_framework_for_Wikidata

Freie Universität Berlin · Universität Zürich UZH · Wikimedia DEUTSCHLAND