

The Trials and Tribulations of Working with Structured Data - a Study on Information Seeking Behaviour

Laura M Koesten¹, Emilia M Kacprzak¹, Jenifer F A Tennison², Elena Simper¹

¹ University of Southampton, UK ² The Open Data Institute, UK

オンラインの構造化されたデータから情報を探す行動を対象に、職業や立場の違う20人の被験者のインタビューと分析を実施。

Human-Data Interactionのフレームワーク

TASK	SEARCH	EVALUATE	EXPLORE	USE
goal or process oriented	STRATEGIES:			
linking	• web search engines	• relevance ¹	• basic visual scan	
time series analysis	• data portals	• usability ²	• obvious errors	
summarising	• asking people	• quality ³	• summarizing statistics	
presenting	• FOI request		• headers	
exporting			• documentation	
			• metadata	
1: context, coverage, summary, original purpose, granularity, time frame	2: labeling, documentation, licence, access, machine readability, language used, format, ability to share, format		3: collection methods, provenance, consistency of formatting/labeling, completeness, what's been excluded	

そのデータセットが役に立ちそうか決める観点

Relevance	context, coverage, original purpose, granularity, summary, time frame
Usability	labeling, documentation, license, access, machine readability, language used, format, schema, ability to share
Quality	collection methods, provenance, consistency of formatting / labeling, completeness, what has been excluded

最初にデータセットをみたときにユーザが行う最初のステップ

scrolling through / basic visual scan
 looking at headers
 looking for obvious errors
 looking at summarising statistics
 filtering relevant data (pivot tables)
 visualising to see trends / peaks
 looking at schema / format
 understanding semantics of the data
 looking at documentation / metadata
 checking the publication date and provenance
 understand the purpose of original dataset