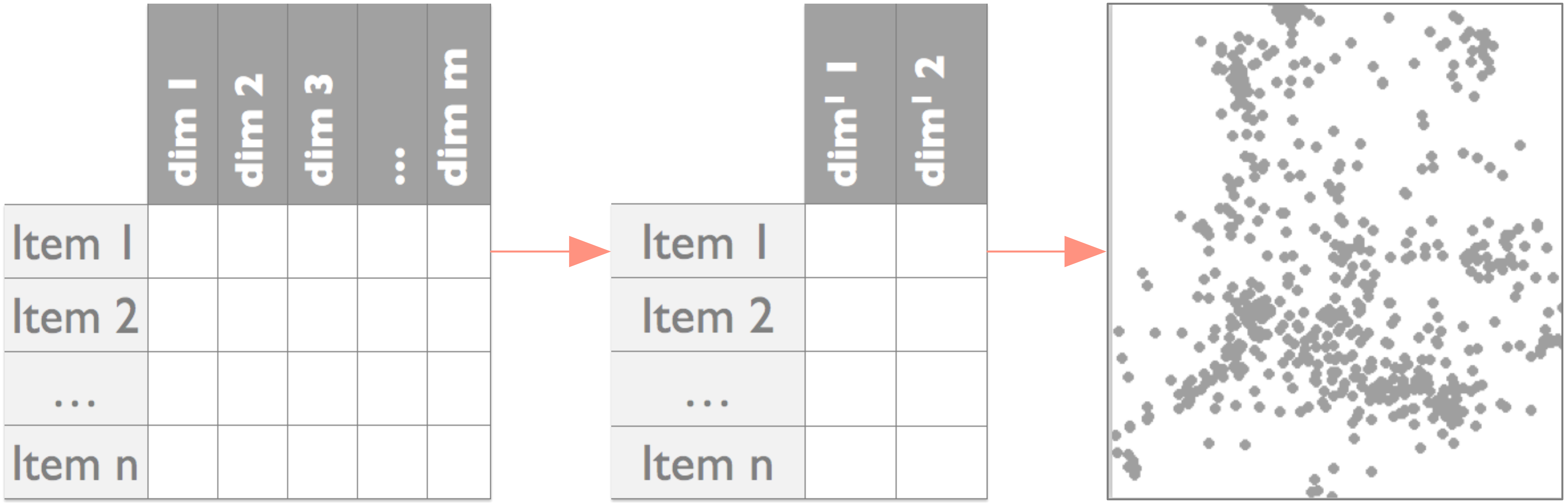
Matthew Brehmer

Michael Sedlmair

Stephen Ingram

Tamara Munzner



Visualizing Dimensionally-Reduced Data:
Interviews with Analysts and a Characterization of Task Sequences

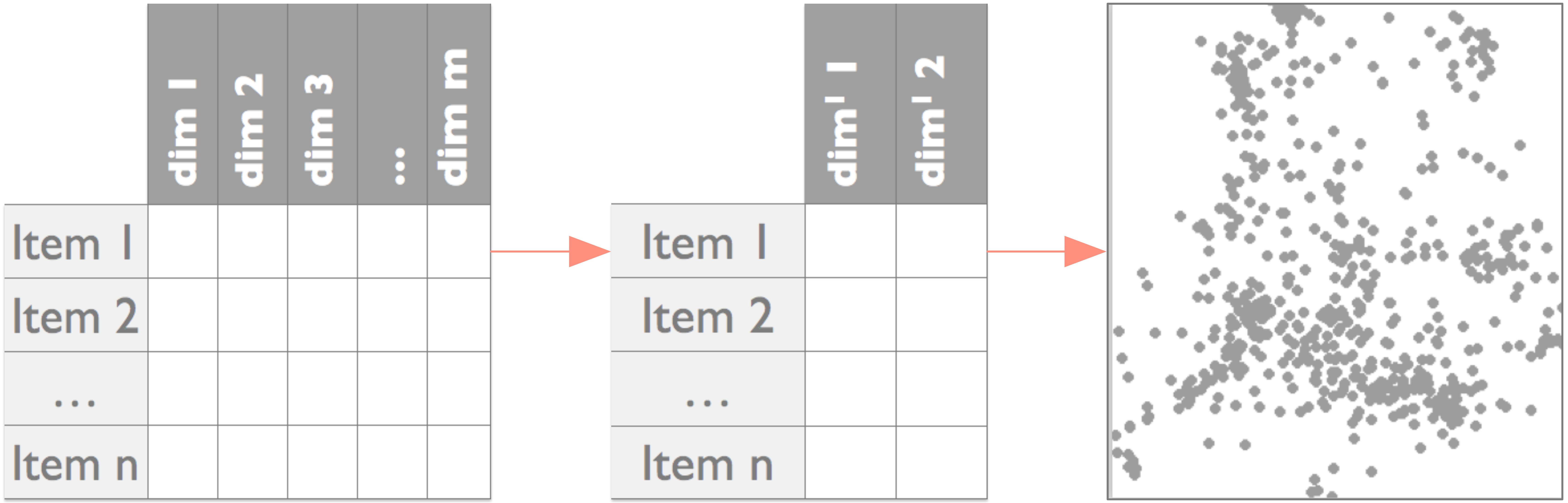
ACM BELIV Workshop
Nov 10, 2014

Matthew Brehmer

Michael Sedlmair

Stephen Ingram

Tamara Munzner



Visualizing Dimensionally-Reduced Data: Interviews with Analysts and a Characterization of Task Sequences

ACM BELIV Workshop
Nov 10, 2014

A need for abstract task characterization...

...yet specific to data type

AN ABUNDANCE OF VIS TASK CHARACTERIZATION

Amar & Stasko (2004)
Amar, Eagan, & Stasko (2005)
Andrienko & Andrienko (2006)
Brehmer & Munzner (2013)
Buja et al. (1996)
Card, Mackinlay, Shneiderman (1999)
Casner (1991)
Chi & Riedl (1998)
Chuah & Roth (1996)
Dix & Ellis (1998)

Gotz & Zhou (2008)
Heer & Shneiderman (2012)
Keim (2002)
Klein, Moon, & Hoffman (2006)
Liu & Stasko (2010)
Mullins & Treu (1993)
Pike, Stasko, et al. (2009)
Pirolli & Card (2005)
Schulz et al. (2013)
Spence (2007)

Springmeyer et al. (1992)
RE Roth (2013)
Roth & Mattis (1990)
Shneiderman (1996)
Tweedie (1997)
Valiati et al. (2006)
Ward & Yang (2004)
Wehrend & Lewis (1990)
Yi, Stasko, et al. (2007)
Zhou & Feiner (1998)

DATA-TYPE SPECIFIC TASK CHARACTERIZATION

Vis Tasks for **1D, 2D, 3D, Multi-Dim, Temporal, Tree, & Network Data**

Shneiderman. (1996) IEEE Symp. Visual Languages

Vis Tasks for **Tabular Data**

Henry & Fekete. (2006) ACM **BELIV** Workshop

Vis Tasks for **Graph Data**

Lee et al. (2006) ACM **BELIV** Workshop

Vis Tasks for **Time-Oriented Data**

Lammarsch et al. (2012) EuroVA Workshop

DATA-TYPE SPECIFIC TASK CHARACTERIZATION

Vis Tasks for **1D, 2D, 3D, Multi-Dim, Temporal, Tree, & Network Data**

Shneiderman. (1996) IEEE Symp. Visual Languages

Vis Tasks for **Tabular Data**

Henry & Fekete. (2006) ACM **BELIV** Workshop

Vis Tasks for **Graph Data**

Lee et al. (2006) ACM **BELIV** Workshop

Vis Tasks for **Time-Oriented Data**

Lammarsch et al. (2012) EuroVA Workshop

...what about DR data?

Dimensionality Reduction (e.g. PCA, MDS) & Vis

	dim 1	dim 2	dim 3	...	dim m
Item 1					
Item 2					
...					
Item n					

DIMENSIONALITY REDUCTION (E.G. PCA, MDS) & VIS

	dim 1	dim 2	dim 3	...	dim m
Item 1					
Item 2					
...					
Item n					

	dim' 1	dim' 2
Item 1		
Item 2		
...		
Item n		

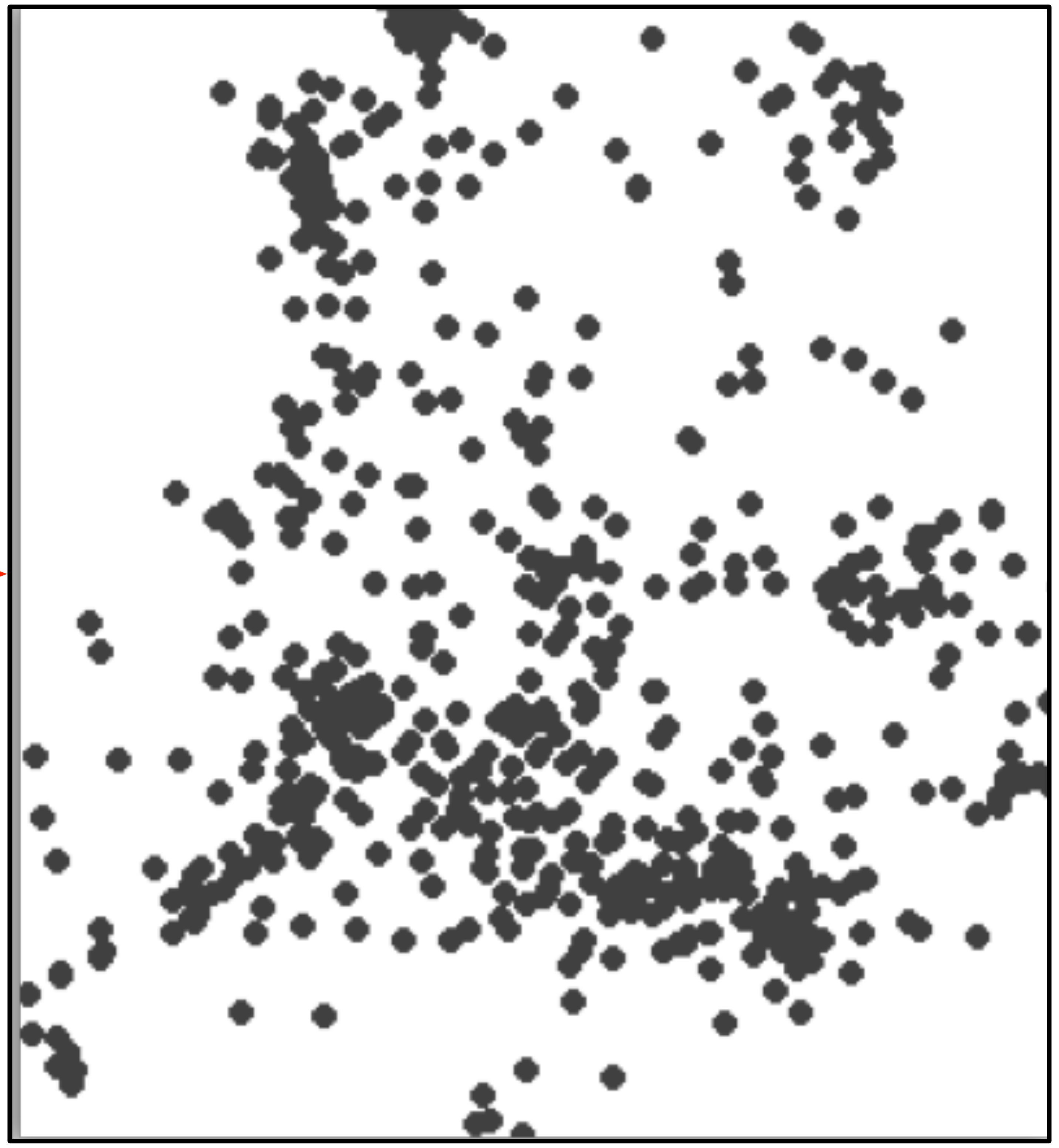
DIMENSIONALITY REDUCTION (E.G. PCA, MDS) & VIS

	dim 1	dim 2	dim 3	...	dim m
Item 1					
Item 2					
...					
Item n					

→

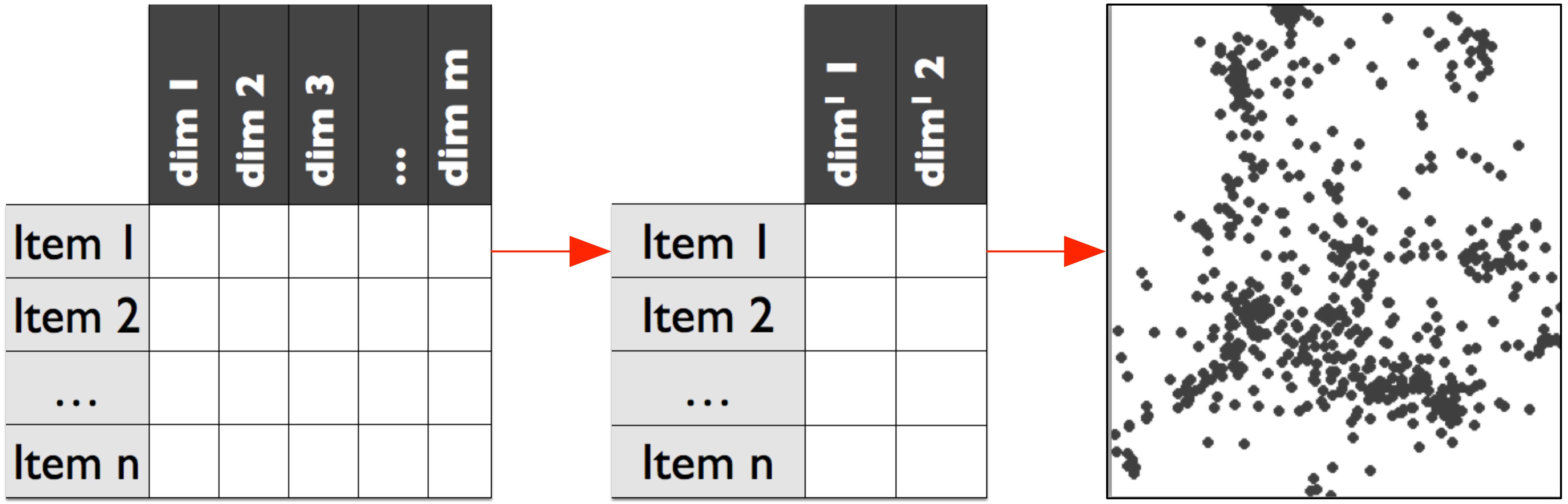
	dim' 1	dim' 2
Item 1		
Item 2		
...		
Item n		

→

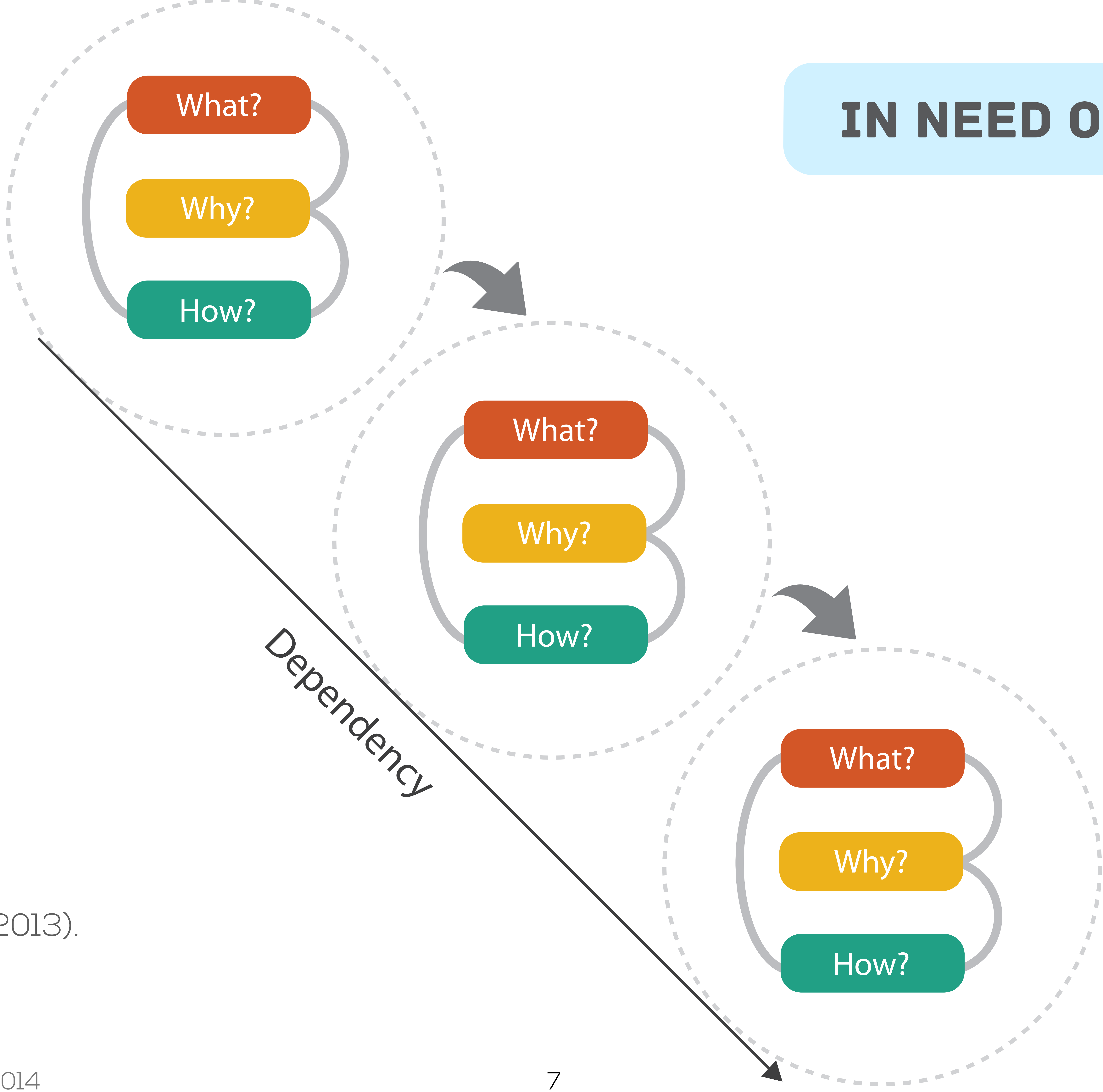


10 ANALYST INTERVIEWEES, 6 DOMAINS

Human computer interaction (x3)
Bioinformatics (x3)
Policy analysis
Computational chemistry
Social network analysis
Investigative journalism



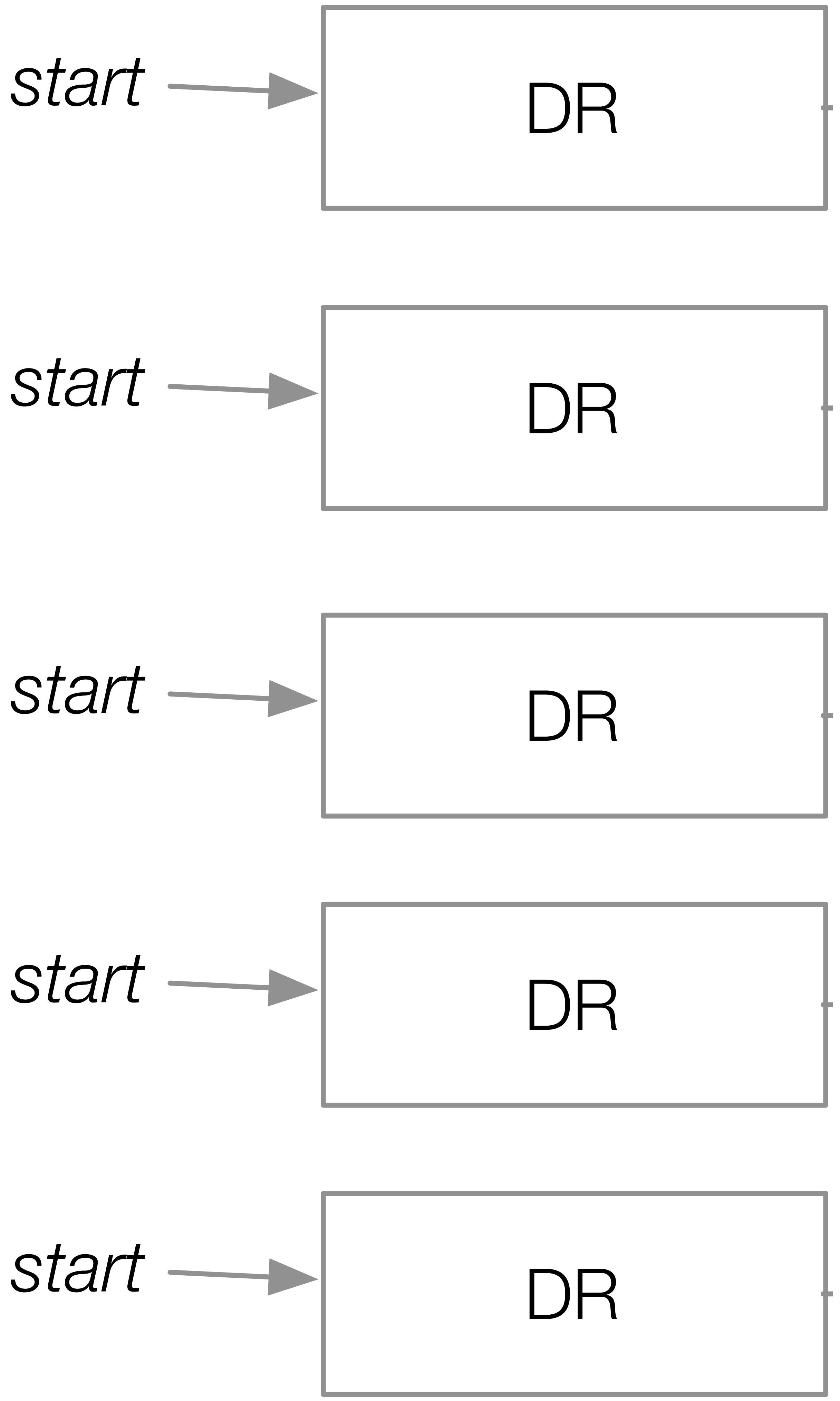
IN NEED OF A FRAMEWORK



Brehmer & Munzner.
IEEE TVCG / Proc. InfoVis (2013).

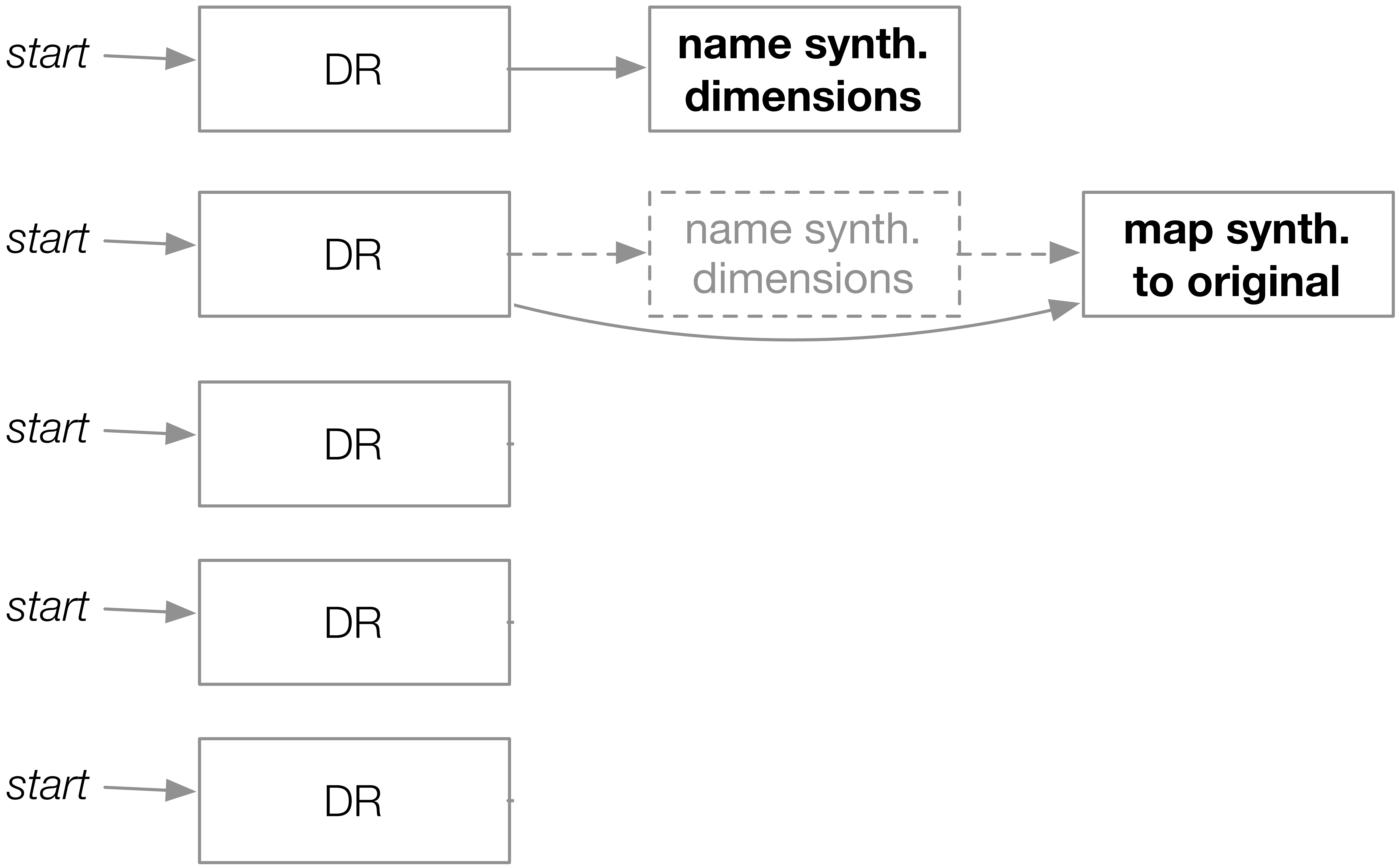
Munzner (2014)

CONTRIBUTION: VIS TASK SEQUENCES FOR DR DATA



CONTRIBUTION: VIS TASK SEQUENCES FOR DR DATA

2 dimension-oriented sequences



CONTRIBUTION: VIS TASK SEQUENCES FOR DR DATA

IMPLICATIONS FOR VIS EVALUATION

understanding work practices
evaluating visual data analysis & reasoning
evaluating communication through vis
evaluating collaborative data analysis
evaluating user performance
evaluating user experience
evaluating vis algorithms

"Seven Scenarios": Lam et al. IEEE TVCG 2012.

Empirical Studies in Information Visualization: Seven Scenarios

Heidi Lam, Enrico Bertini, Petra Isenberg, Catherine Plaisant, and Sheelagh Carpendale

Abstract—We take a new, scenario-based look at evaluation in information visualization. Our seven scenarios, evaluating visual data analysis and reasoning, evaluating user performance, evaluating user experience, evaluating environments and work practices, evaluating communication through visualization, evaluating visualization algorithms, and evaluating collaborative data analysis were derived through an extensive literature review of over 800 visualization publications. These scenarios distinguish different study goals and types of research questions and are illustrated through example studies. Through this broad survey and the distillation of these scenarios, we make two contributions. One, we encapsulate the current practices in the information visualization research community and, two, we provide a different approach to reaching decisions about what might be the most effective evaluation of a given information visualization. Scenarios can be used to choose appropriate research questions and goals and the provided examples can be consulted for guidance on how to design one's own study.

Index Terms—Information visualization, evaluation.

1 Introduction

Evaluation in information visualization is complex since, for a thorough understanding of a tool, it not only involves assessing the visualizations themselves, but also the complex processes that a tool is meant to support. Examples of such processes are exploratory data analysis and reasoning, communication through visualization, or collaborative data analysis. Researchers and practitioners in the field have long identified many of the challenges faced when planning, conducting, and executing an evaluation of a visualization tool or system [10], [41], [54], [63]. It can be daunting for evaluators to identify the right evaluation questions to ask, to choose the right variables to evaluate, to pick the right tasks, users, or data sets to test, and to pick appropriate evaluation methods. Literature guidelines exists that can help with these problems but they are almost exclusively focused on methods—"structured as an enumeration of methods with focus on *how* to carry them out, without prescriptive advice for *when* to choose between them." ([54, p.1], author's own emphasis).

This paper takes a different approach: instead of focusing on evaluation methods, we provide an in-depth discussion of evaluation scenarios, categorized into those for understanding data analysis processes and those which evaluate visualizations themselves.

The scenarios for understanding data analysis are

- Understanding environments and work practices (UWP),
- evaluating visual data analysis and reasoning (VDAR),
- evaluating communication through visualization (CTV), and
- evaluating collaborative data analysis (CDA).

The scenarios for understanding visualizations are

- Evaluating user performance (UP),
- evaluating user experience (UE), and
- evaluating visualization algorithms (VA).

Our goal is to provide an overview of different types of evaluation scenarios and to help practitioners in setting the right evaluation goals, picking the right questions to ask, and to consider a variety of methodological alternatives to evaluation for the chosen goals and questions. Our scenarios were derived from a systematic analysis of 850 papers (361 with evaluation) from the information visualization research literature (Section 5). For each evaluation scenario, we list the most common evaluation goals and outputs, evaluation questions, and common approaches in Section 6. We illustrate each scenario with representative published evaluation examples from the information visualization community. In cases where there are gaps in our community's evaluation approaches, we suggest examples from other fields. We strive to provide a wide coverage of the methodology space in our scenarios to offer a diverse set of evaluation options. Yet, the "Methods and Examples" lists in this paper are not meant to be comprehensive as our focus is on choosing among evaluation scenarios. Instead, we direct the interested reader

- *H. Lam is with Google, Inc, Mountain View, CA. E-mail: heidi.lam@gmail.com.*
- *E. Bertini is with the Department of Computer and Information Science, University of Konstanz, Box 78, Konstanz 78457, Germany. E-mail: enrico.bertini@uni-konstanz.de.*
- *P. Isenberg is with INRIA, Université Paris-Sud, Team Aviz, Bat 650, Saclay, Orsay Cedex 91405, France. E-mail: petra.isenberg@inria.fr.*
- *C. Plaisant is with the University of Maryland, 2117C Hornbake South Wing, College Park, MD 20742. E-mail: plaisant@cs.umd.edu.*
- *S. Carpendale is with the Department of Computer Science, University of Calgary, 2500 University Dr. NW, Calgary, AB T2N 1N4, Canada. E-mail: sheelagh@ucalgary.ca.*

Manuscript received 8 Sept. 2010; revised 6 Nov. 2011; accepted 9 Nov. 2011; published online 30 Nov. 2011.
Recommended for acceptance by C. North.
For information on obtaining reprints of this article, please send e-mail to: tvcg@computer.org, and reference IEEECS Log Number TVCG-2010-09-0224.
Digital Object Identifier no. 10.1109/TVCG.2011.279.

Published by the IEEE Computer Society

Matthew Brehmer

Michael Sedlmair

Stephen Ingram

Tamara Munzner

CONCLUSION

Domain-Agnostic and **Data-Type-Specific** Task Characterization

Tasks in **Sequence**, not in Isolation

Tasks Characterization for **BELIV**

cs.ubc.ca/labs/imager/tr/2014/DRVisTasks/

thanks: UBC InfoVis group. UBC Multimodal User Experience group

DR Vis Tasks: Supplemental

Implications for Vis Evaluation

Lam et al. IEEE TVCG 2012.

IMPLICATIONS FOR VIS EVALUATION

Lam et al. IEEE TVCG 2012.

understanding work practices

pre-design **requirements analysis**, especially in problem-driven **design studies**; use task sequences as code set

IMPLICATIONS FOR VIS EVALUATION

Lam et al. IEEE TVCG 2012.

understanding work practices	evaluating user performance
pre-design **requirements analysis**, especially in problem-driven **design studies**; use task sequences as code set	task sequences can inform **experimental design** and participant instructions when evaluating techniques combining DR + Vis

IMPLICATIONS FOR VIS EVALUATION

Lam et al. IEEE TVCG 2012.

understanding work practices	evaluating user performance	evaluating user experience
pre-design **requirements analysis**, especially in problem-driven **design studies**; use task sequences as code set	task sequences can inform **experimental design** and participant instructions when evaluating techniques combining DR + Vis	inform participant instructions in a think-aloud evaluation **focus** questionnaire **questions**…

IMPLICATIONS FOR VIS EVALUATION

Lam et al. IEEE TVCG 2012.

understanding work practices	evaluating user performance	evaluating user experience	evaluating visual data analysis & reasoning
pre-design **requirements analysis**, especially in problem-driven **design studies**; use task sequences as code set	task sequences can inform **experimental design** and participant instructions when evaluating techniques combining DR + Vis	inform participant instructions in a think-aloud evaluation **focus** questionnaire **questions**…	analyze the use of deployed DR + Vis tools ***in the wild***; use task sequences as code … **focus** diary / interview **questions**

10 Analyst Interviewees, 6 Domains

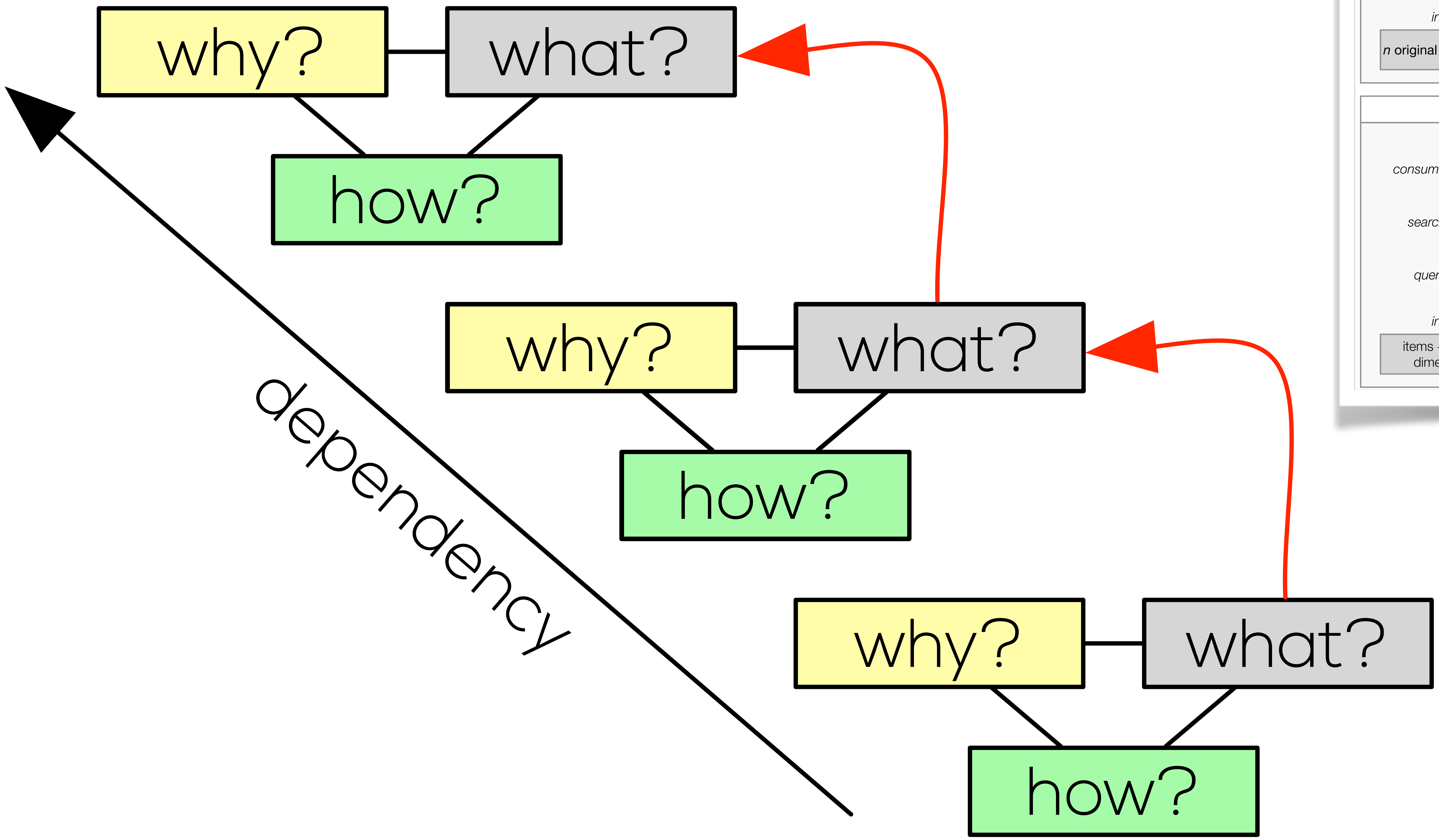
Human computer interaction (x3)
Bioinformatics (x3)
Policy analysis
Computational chemistry
Social network analysis
Investigative journalism

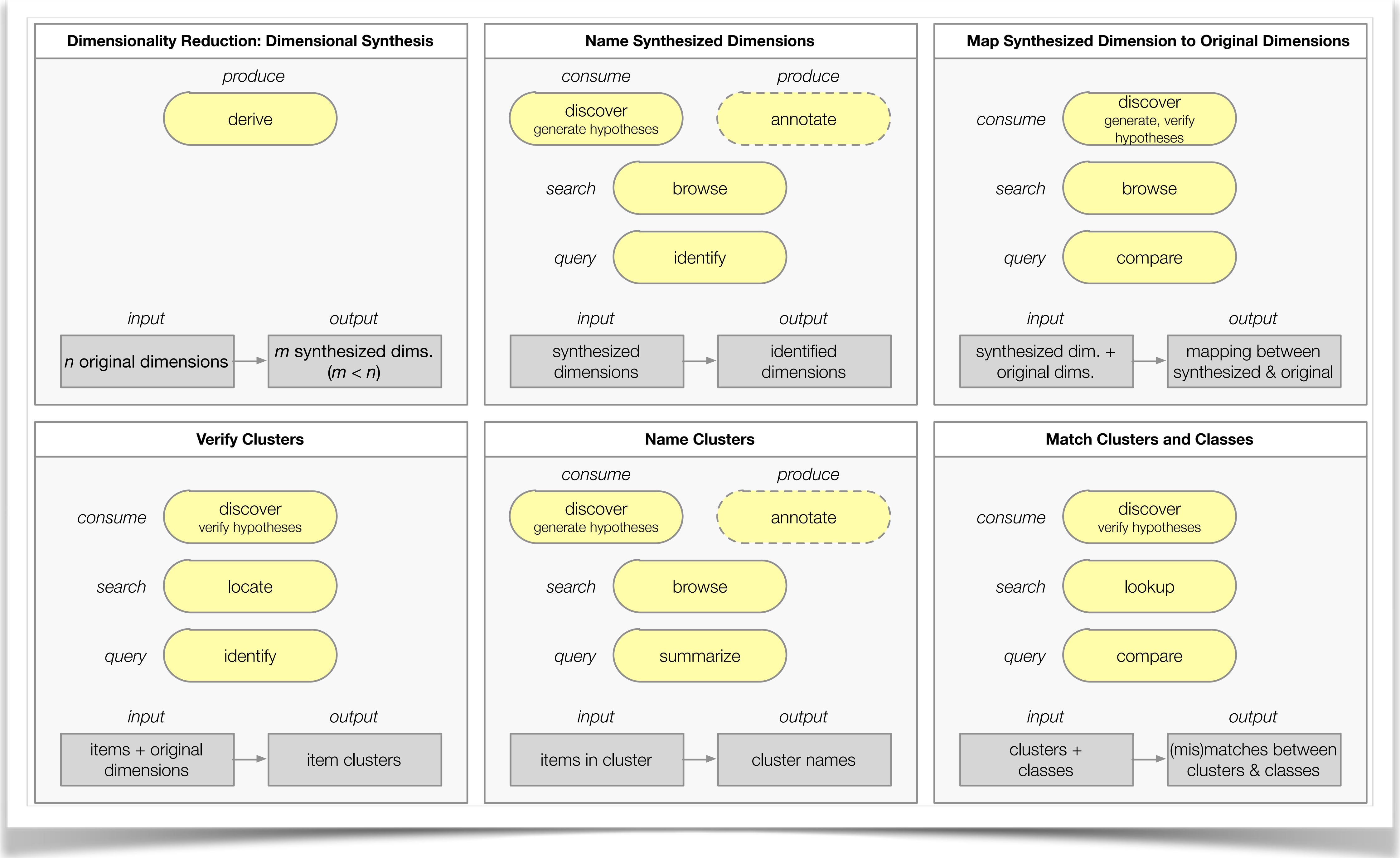
10 ANALYST INTERVIEWEES, 6 DOMAINS

Human computer interaction (x3)
Bioinformatics (x3)
Policy analysis
Computational chemistry
Social network analysis
Investigative journalism

+ 4 known use cases from DR + Vis technique papers

A COMMON LEXICON FOR ANALYSIS





domain-agnostic yet **data-type-specific** task characterization

Brehmer & Munzner. IEEE TVCG / Proc. InfoVis 2013.