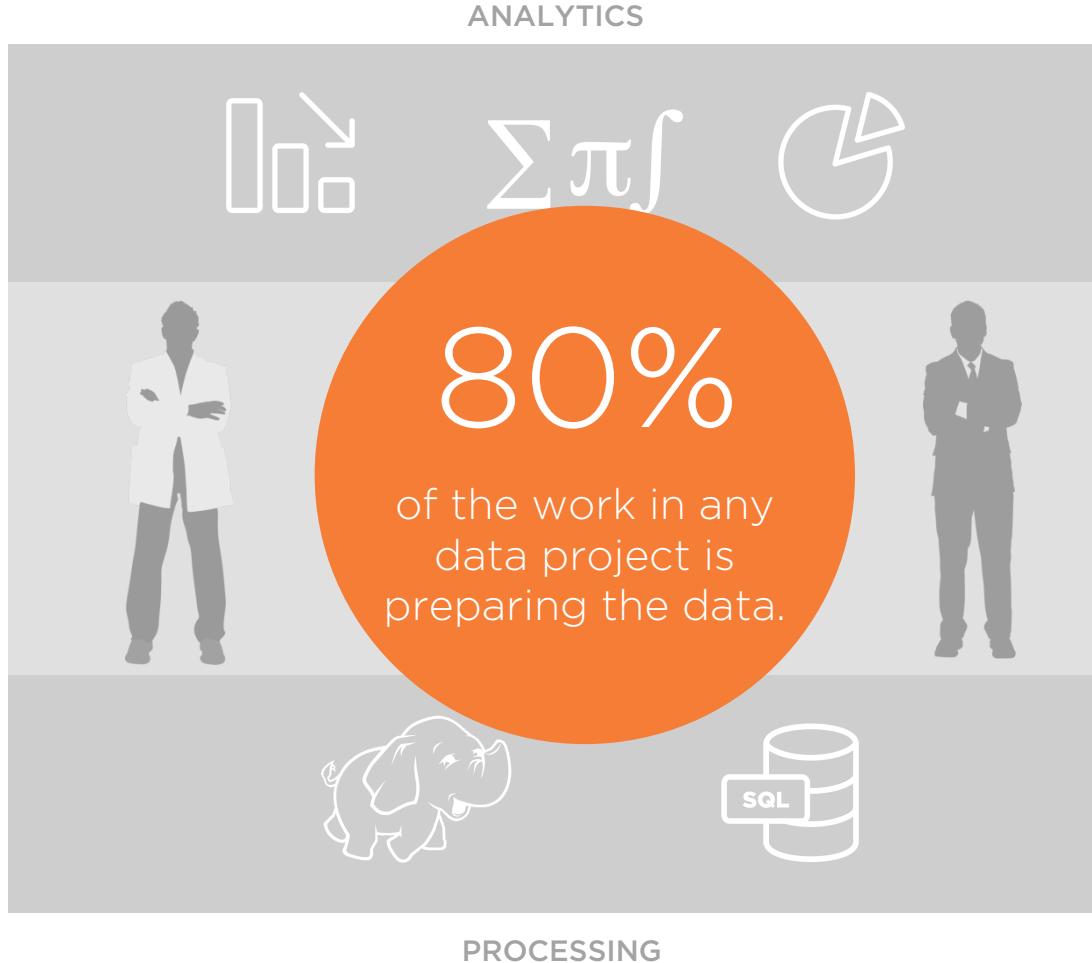
A photograph of a lioness resting in a large, gnarled tree. She is positioned in the center-right of the frame, facing slightly towards the camera with her head turned to the left. Her golden-brown fur is visible against the dark, textured bark of the tree. The background is filled with dense green foliage and branches.

Predictive Interaction for Data Transformation

Joe Hellerstein
Jeffrey Heer
Sean Kandel



WHERE DOES THE TIME GO IN ANALYTICS?



Patil, *Data Jujitsu*, 2012.

Kandel et al. "Enterprise Data Analysis and Visualization: An Interview Study", *IEEE VAST*, 2012.





**Big Data
Borat**

@BigDataBorat



Following

In Data Science, 80% of time spent prepare data, 20% of time spent complain about need for prepare data.



...

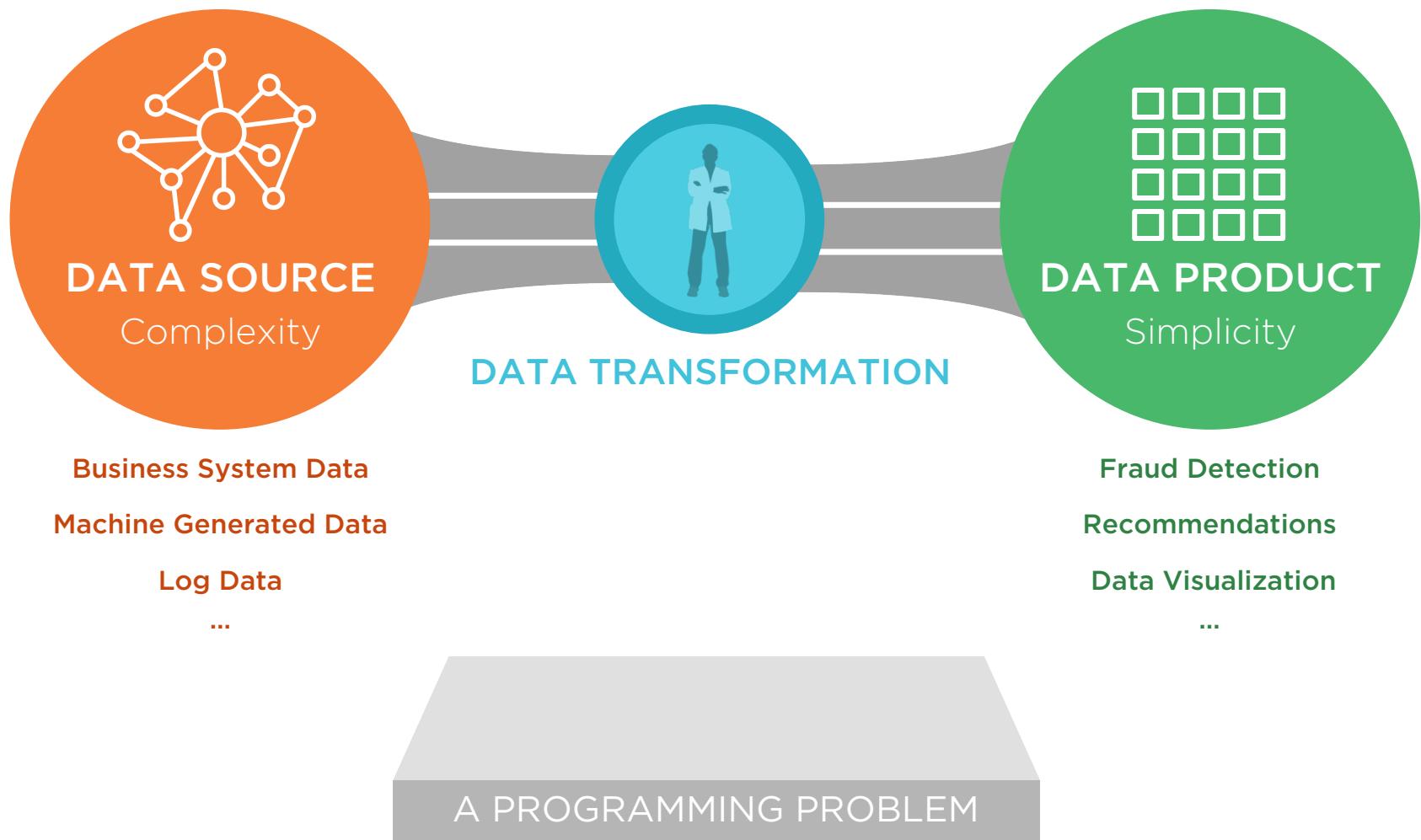


RESEARCH IMPERATIVE

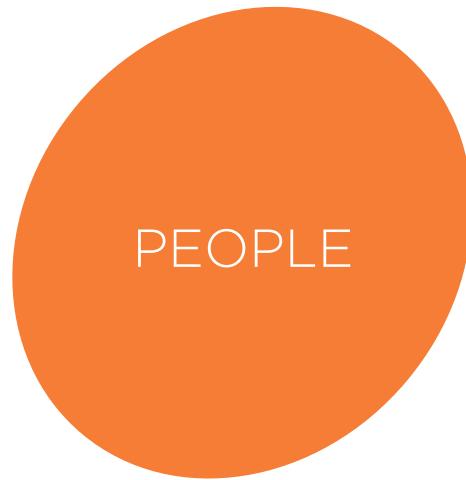
Dramatically simplify labor-intensive tasks in the data lifecycle.



DATA TRANSFORMATION TODAY



APPROACH #1: NEW USER INTERFACES

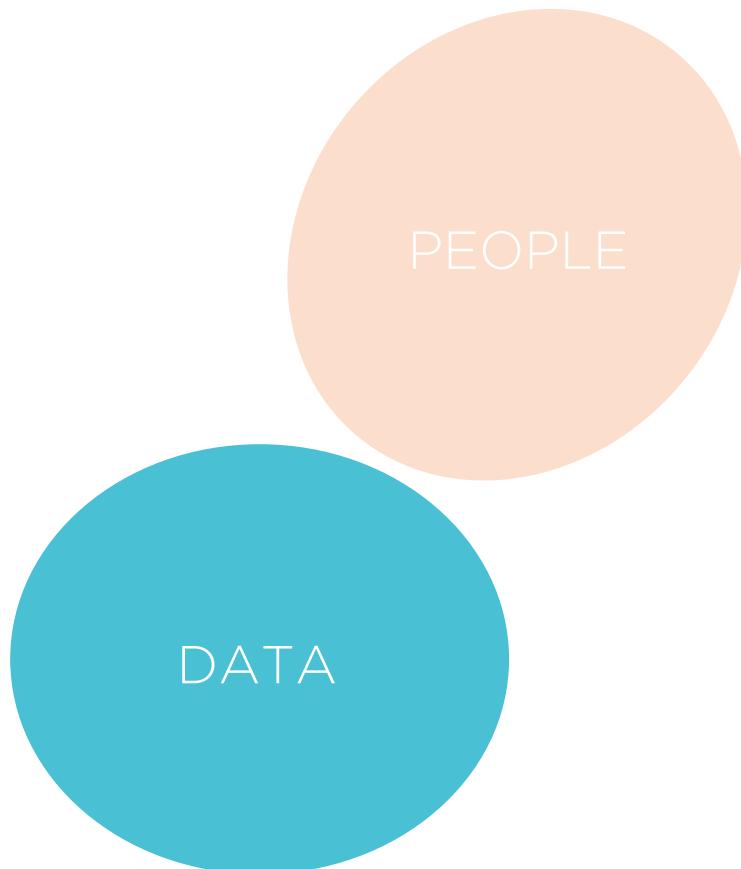


New User Interfaces

- Graphical query & xform
- Data visualization



APPROACH #2: DATABASE DSLs



New User Interfaces

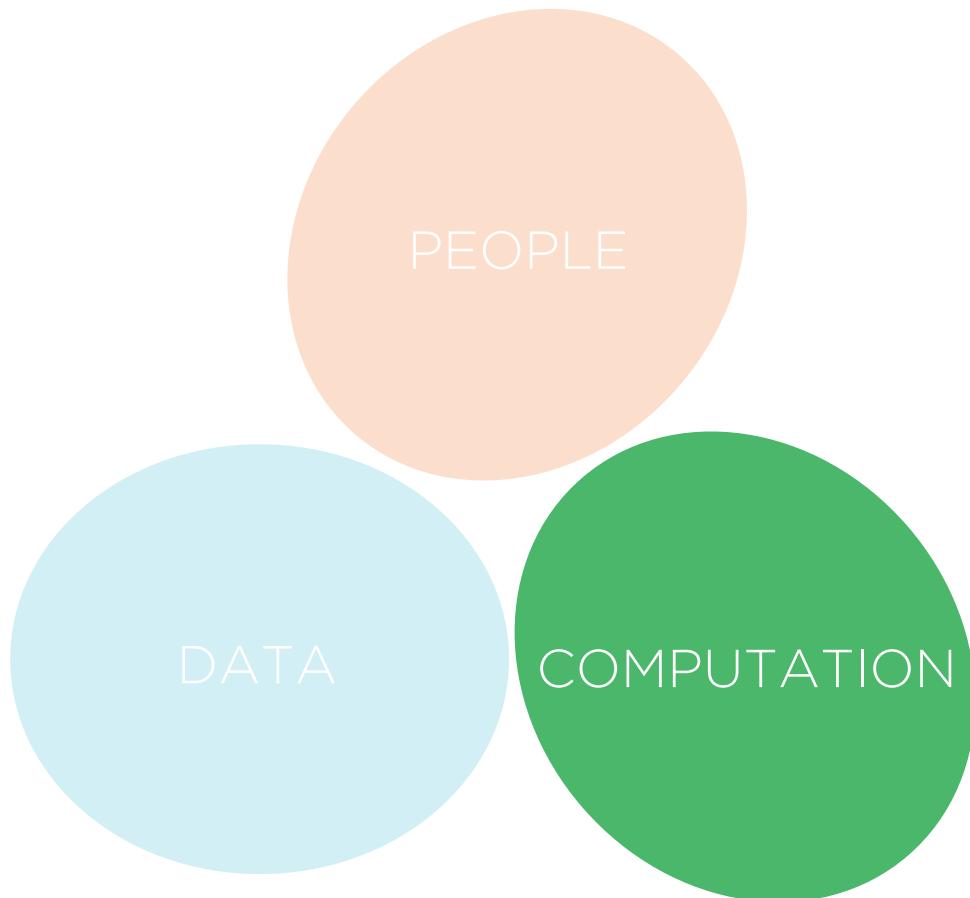
- Graphical query & xform
- Data visualization

Database DSLs

- 2nd-order algebras & calculi
- Proximity & clustering
- Constraints



APPROACH #3: AUTOMATION



New User Interfaces

- Graphical query & xform
- Data visualization

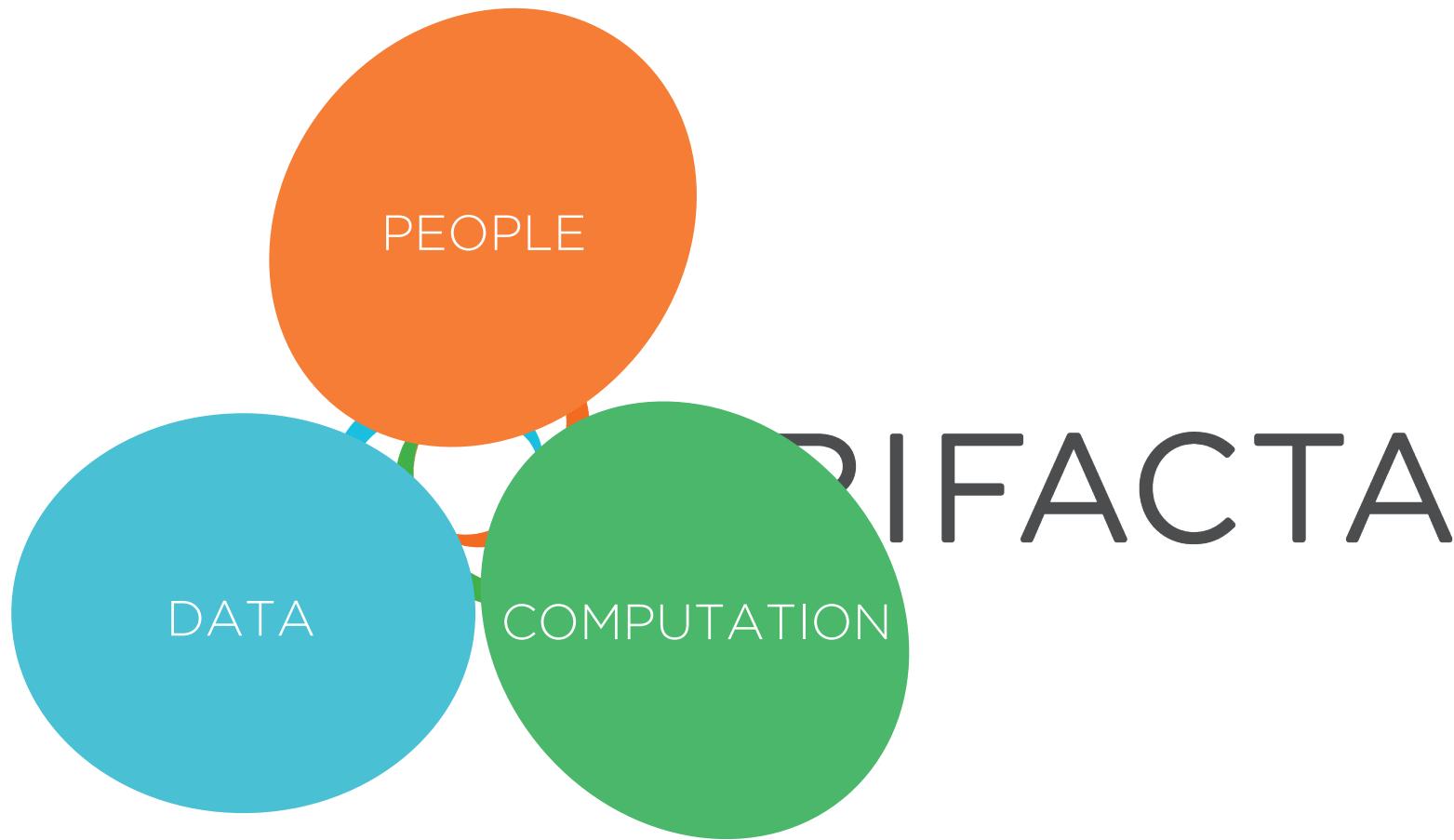
Database DSLs

- 2nd-order algebras & calculi
- Proximity & clustering
- Constraints

Automation

- Probability & Logic
- Crowdsourcing



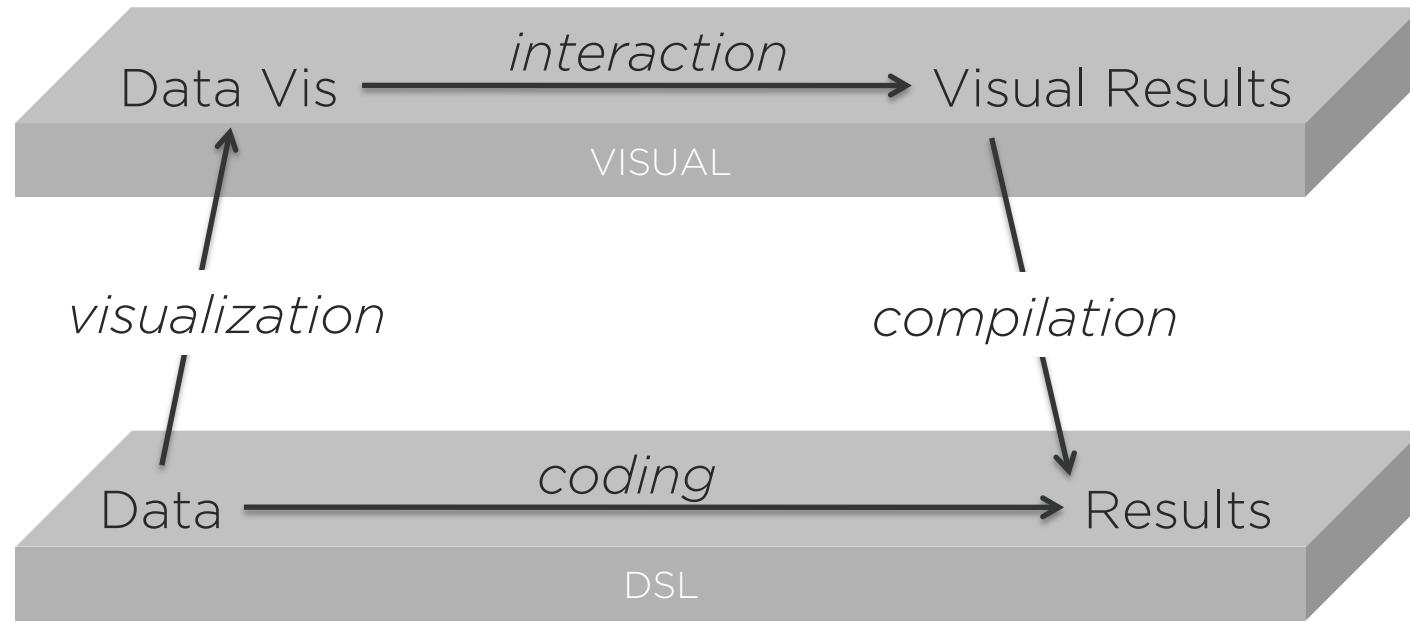




TRIFACTA



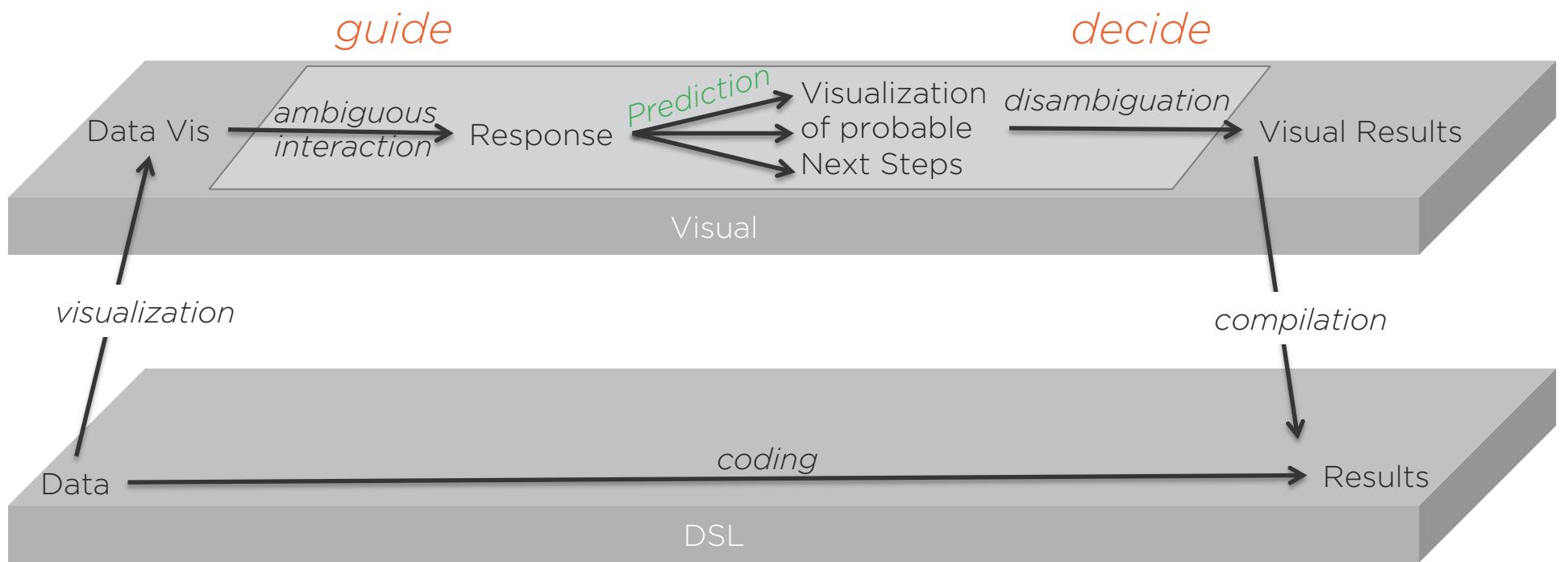
LIFTING TO THE VISUAL DOMAIN



STILL... THE BURDEN OF SPECIFICATION

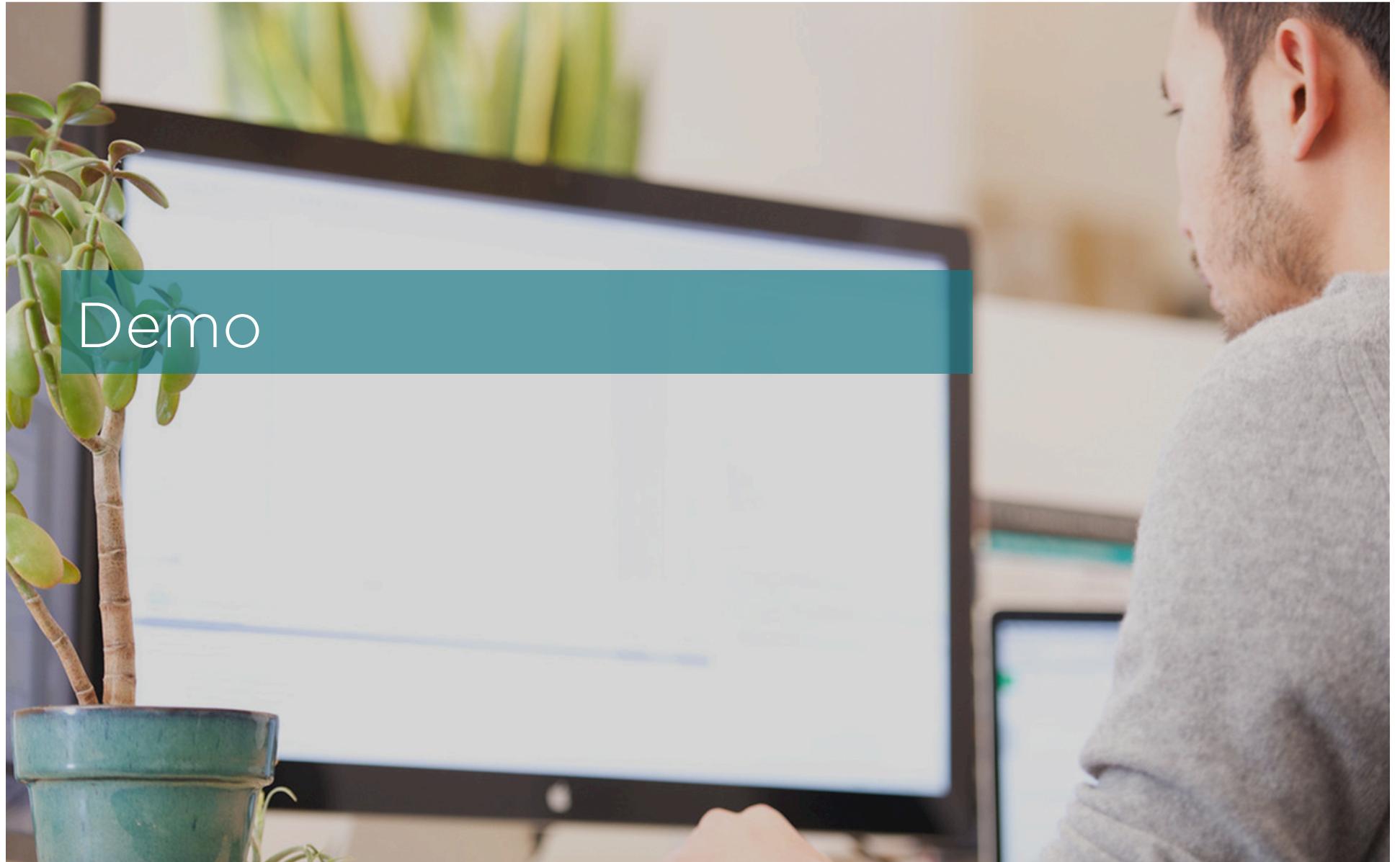


PREDICTIVE INTERACTION

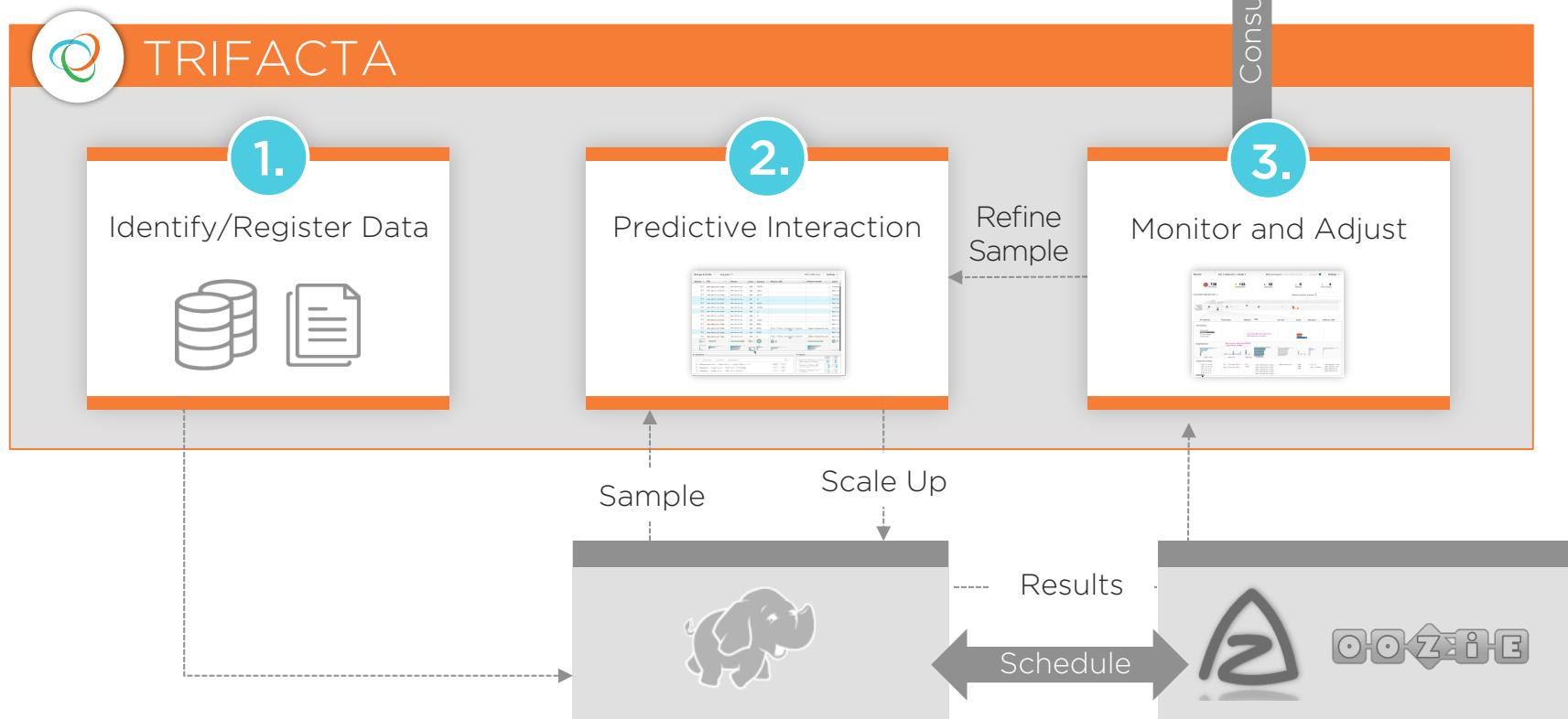




TRIFACTA



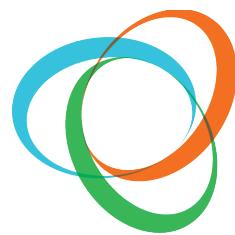
ARCHITECTURE



IN SUM

- Trifacta customer feedback: order-of-magnitude efficiency gains
- Design challenges: Subtle interplay of three aspects
 - **People:** Visualization and Interaction
 - **Data:** DSL of “actual” behaviors
 - **Computation:** prediction of DSL steps
Co-design these three!
- Paper has more on design considerations and applications





TRIFACTA
PEOPLE • TRANSFORMING • DATA