

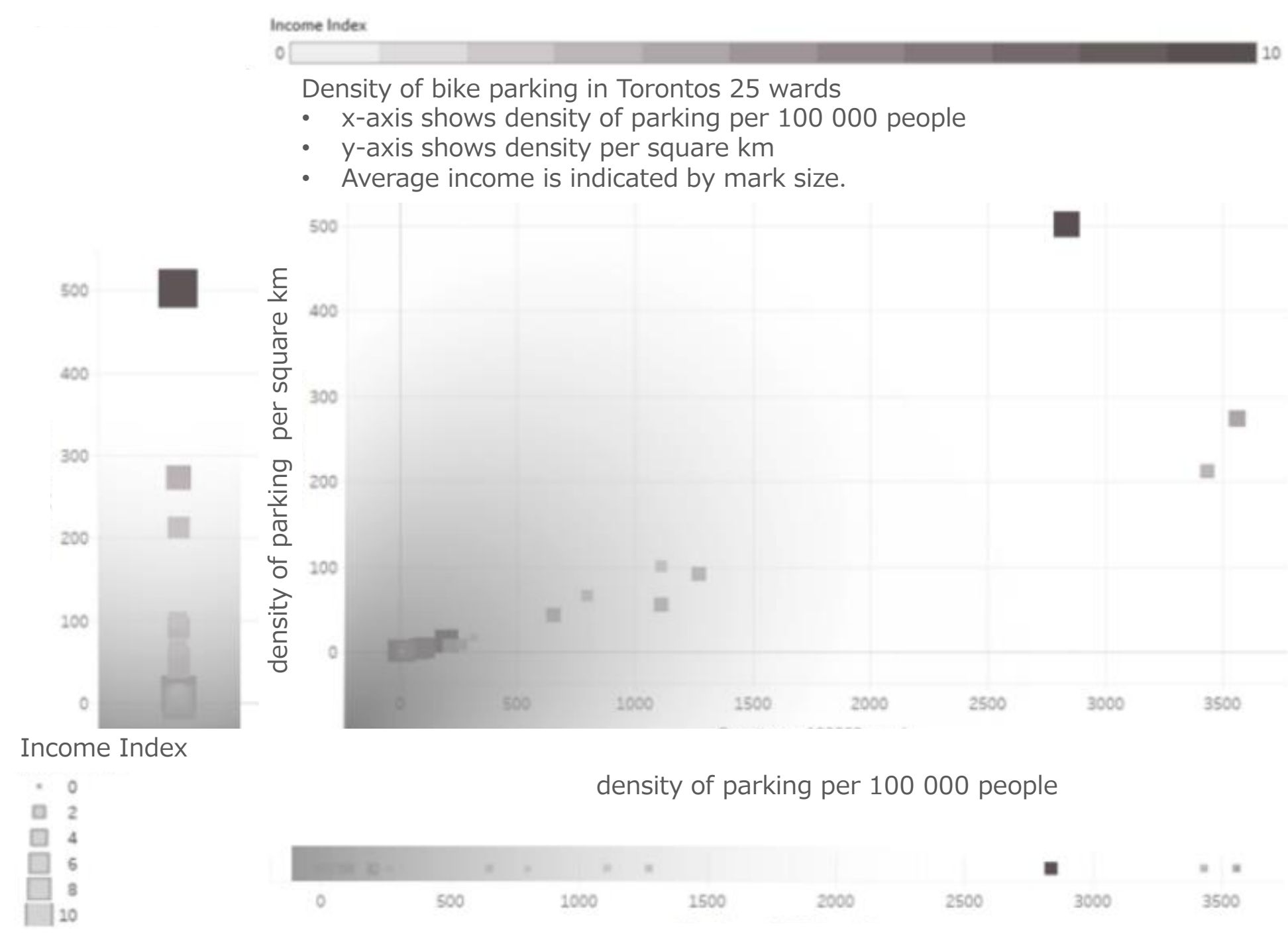
I Don't (Completely) Trust Your Data

Towards Visualization Lexicons for Ambiguous and Incomplete Data

Karly Ross & Wesley Willett

Iteration 1: Gray overlays highlight absences and ambiguities.

How: The person analyzing the data overlays shading where there are ambiguities in the data.

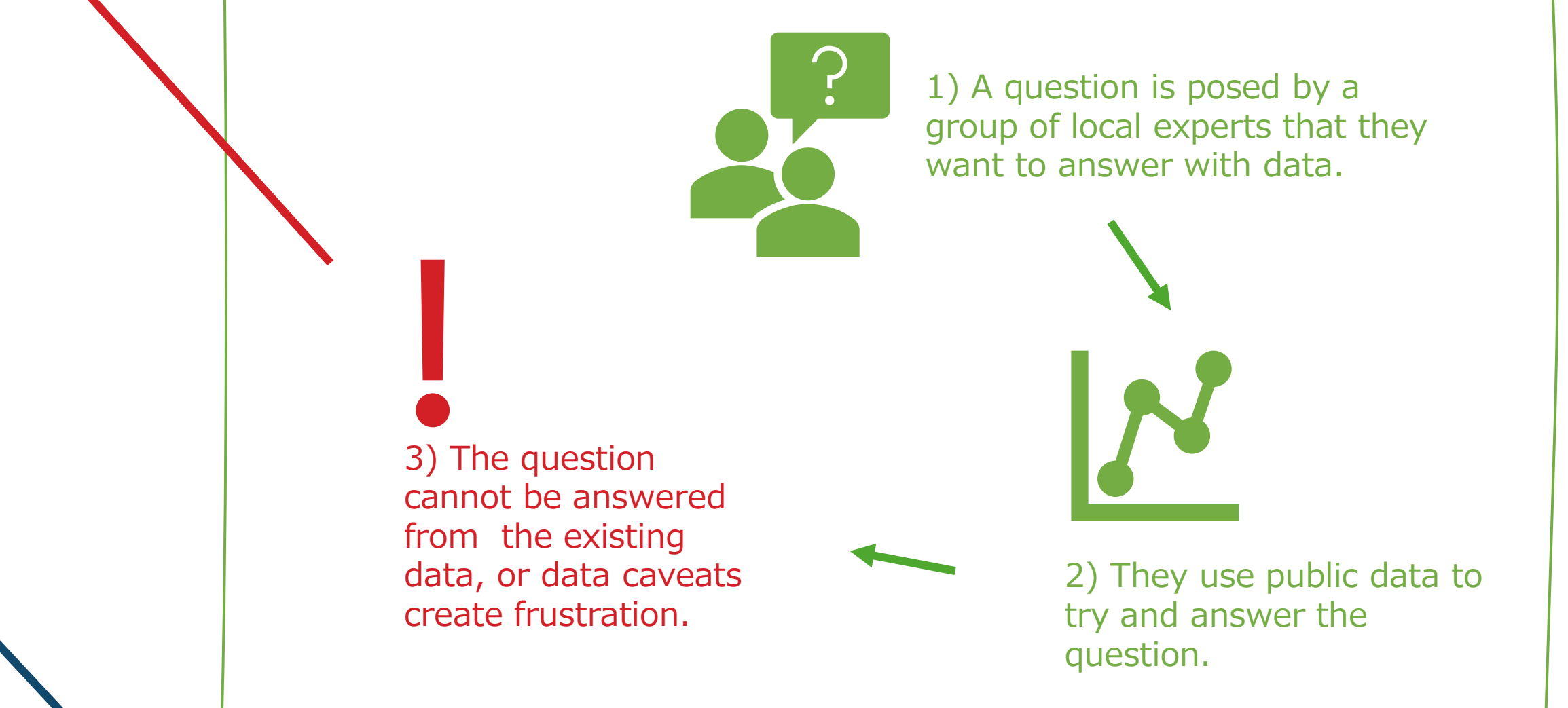


Challenge: In our case study of bike parking, users often have reliable local knowledge not captured by this process.

Gray overlays blur and add ambiguity to uncertain data.

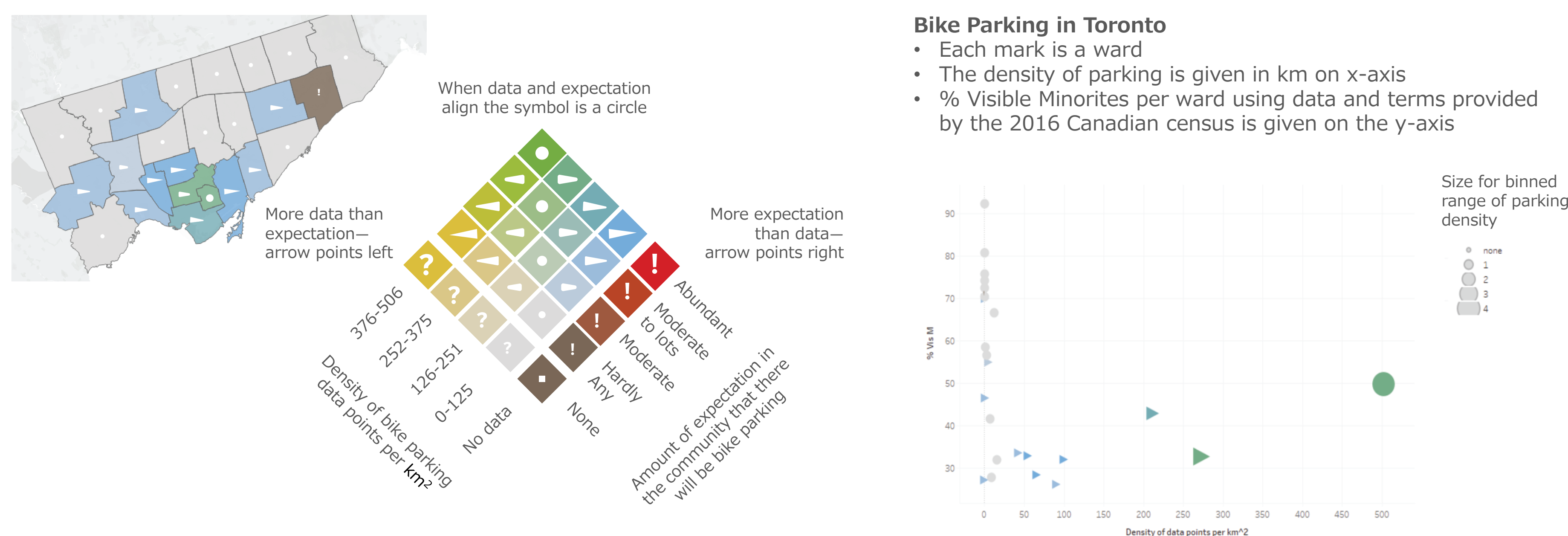
Motivation: While open data provides community access to data, the degree to which users find the provided data useful and complete can be unclear. We use a case study of bicycle parking to explore processes that address community users' data challenges.

Common Community Workflow



Iteration 2: Bivariate color and glyph encodings contrast expectations against available data

How: In this process color, shape and size are all used to indicate the degree of divergence and convergence between community and administrative data.

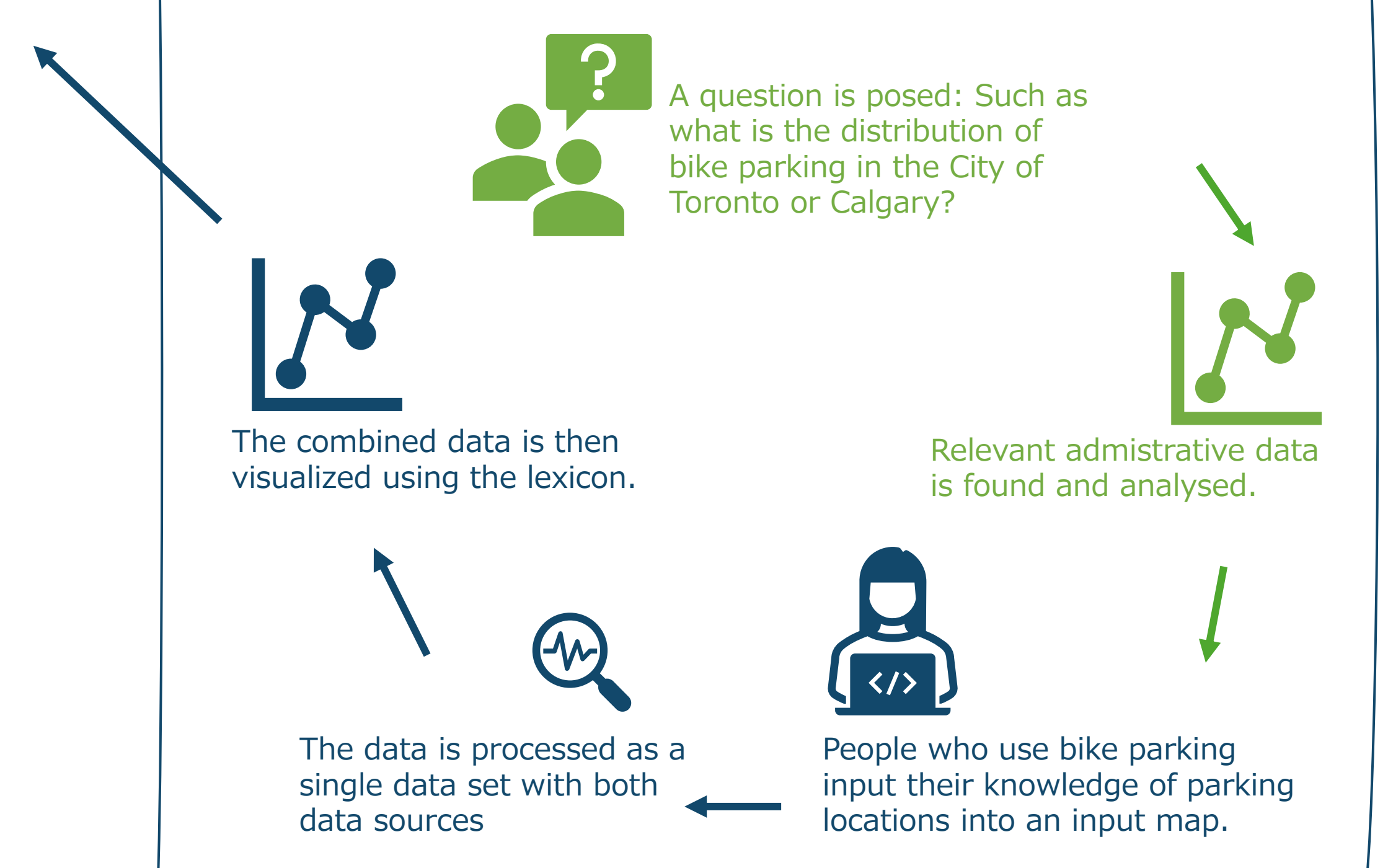


Bivariate color and glyph encodings highlight the differences between 'expected' bike parking and administrative infrastructure records across Toronto wards.

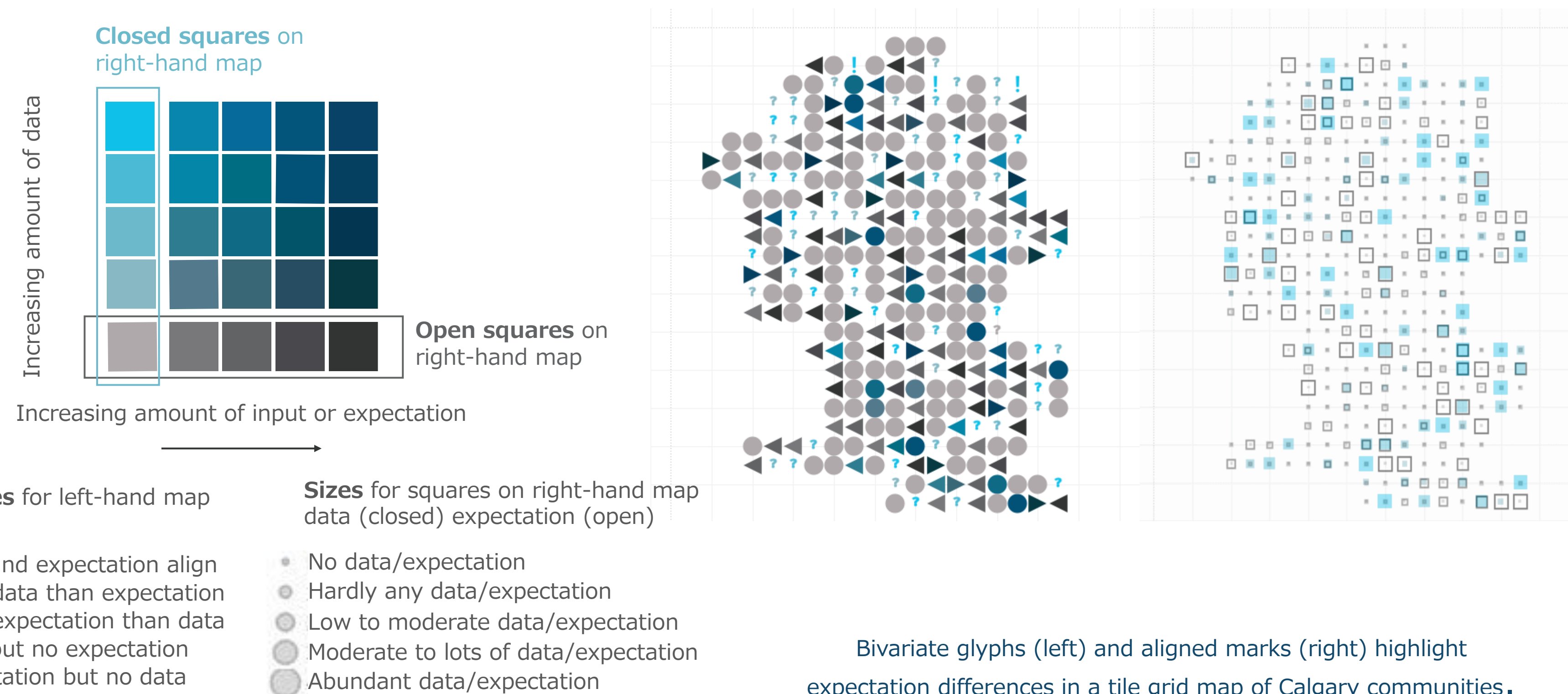
We are developing a **visual lexicon** that:

1. compares what people know (community data) with administrative data (data produced by administrative bodies), and
2. shows where there are absences in both.

Potential workflow



Iteration 3: Refined glyphs and aligned marks support data/expectation comparisons in tile grid maps.



Topics for discussion:

- The **complexity** of the visualizations.
- Value of tools to contrast user **input** with administrative data.
- Challenges in designing representations that allow community users to **view** their data *and* administrative data to enhance knowledge.
- Visualization grounded **facilitation**.