

MOTIVATION

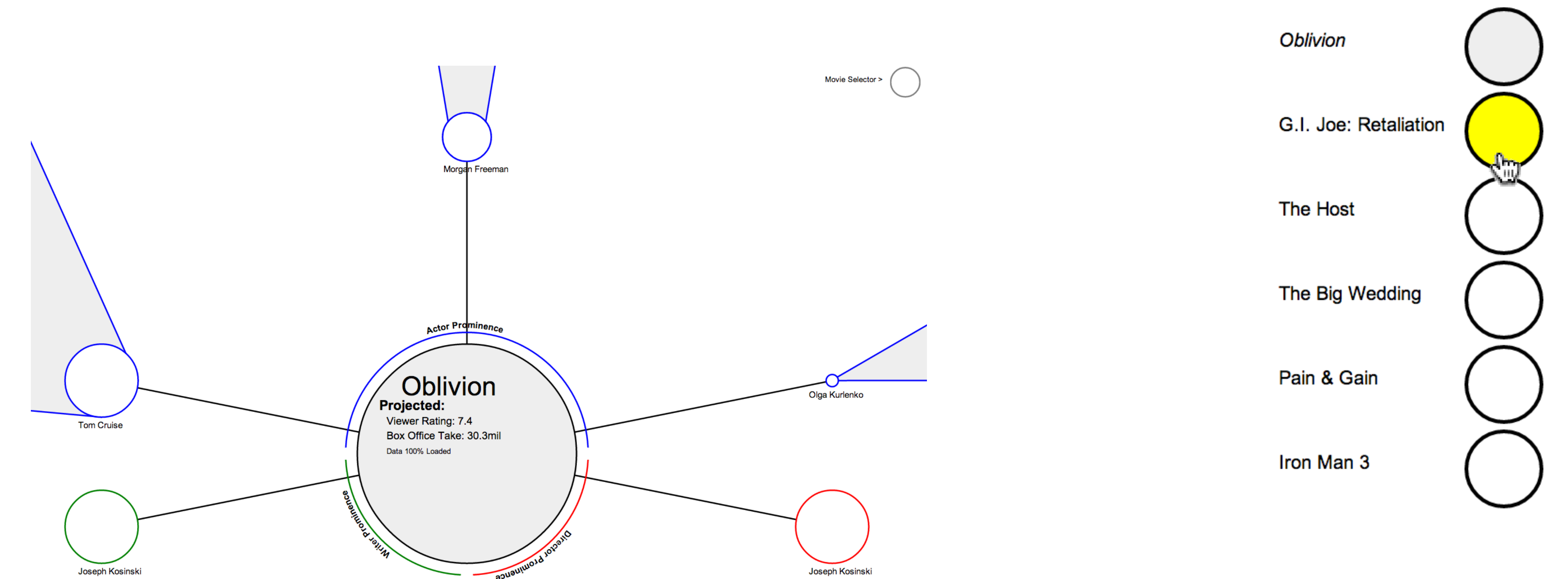
- Big Data makes computationally complex prediction heuristics necessary
- Pre-filtering of relevant data allows at-a-glance consumability
- Graph layout allows navigation of data, creative data encoding and relationship representation

WHAT IS IT?

- Node & Link Based Abstraction of Big Data
- Applicable to a Variety of Hierarchical Datasets
 - Demo shows VAST Challenge example (Movie Data → Box Office Prediction)

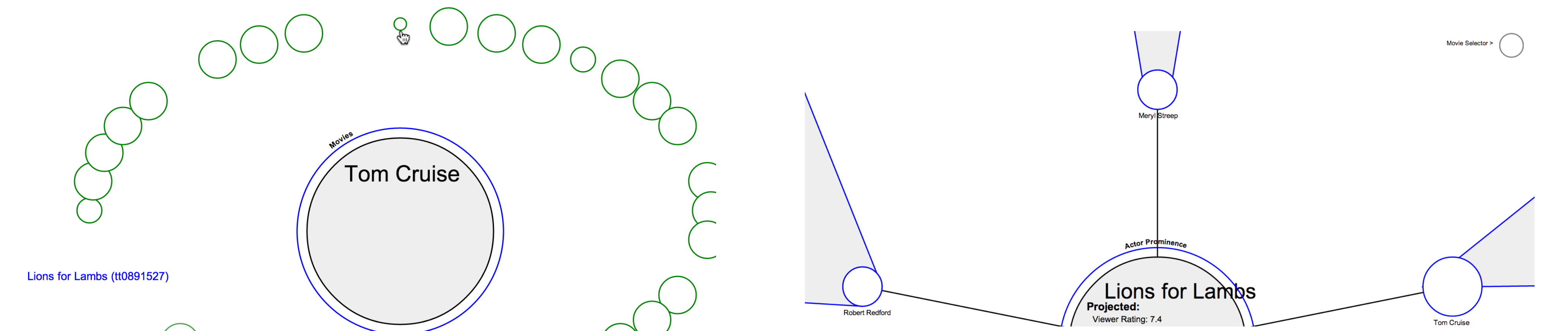
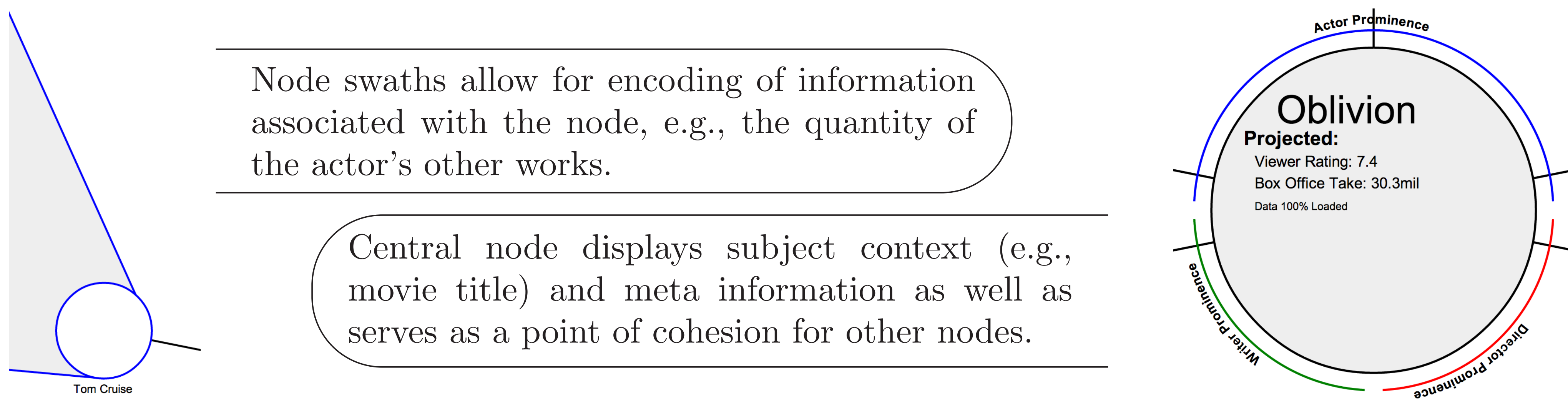
DATA ACQUISITION & SANITATION

1. Data from IMDB .list files (unwieldy), IMDBAPI.org (unofficial), and web scraping (disallowed)
2. Programmatically saved to whitespace delimited text files for re-use in the vis
 - Local storage prevents online dependency and reduces likelihood of being blocked
3. Data cleaned using standard methods and RegExp



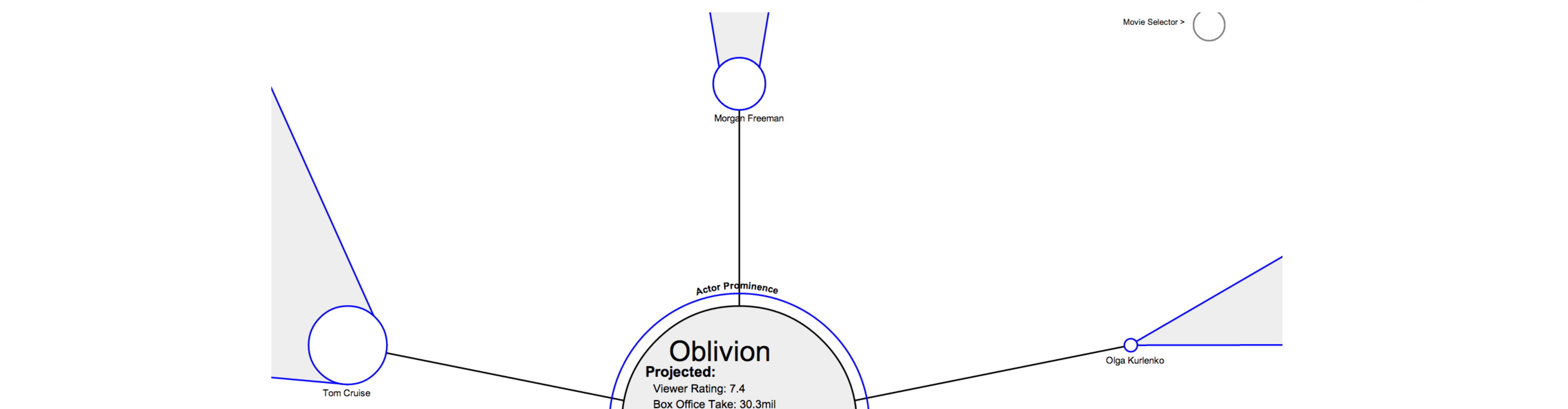
The movie context of the visualization provides a relationship between the different parameters affecting the prediction, displayed in the central node.

Data can be included with the vis. or, if not provided, will be fetched, cleaned and cached at runtime into a form reusable by the visualization on future runs.



Actor selected, context reversed. Tom Cruise's past performances were fetched and filtered based on first 3 billing (customizable) with a "drop off" animation. Size of movies encodes billing order.

Selecting a movie in the secondary (Actor) context, returns the user to the original context with the new central node. Here, Tom Cruise is billed third but his impact on the movie's performance is still evident in his swath size.



The weight an entity contributes to the prediction is represented by clockwise ordering and node size.

PREDICTION

- Portable visual architecture allows for a variety of prediction schemes to be integrated
- Feedback and rationale of prediction available on central node hover

INTERACTIVITY

- Contextual feedback when user hovers and selects
- Overflow data handled gracefully with naïve assumption of input quantity