The number of professional jobs that require working across collections of documents and data continues to grow, from social media marketing through cyber security to health and wellness. Many of these new jobs contain underspecified problems and unknown metrics, unlike the metrics found in industries such as stock trading, online advertising, and sales forecasting. Additionally, the volume and characteristics of relevant data, including noise and lack of structure, create problems for query planning, assessing true quantities and distributions, understanding structure at several scales, and detecting key elements of a data-generating process. This presentation will cover several approaches to these problems, including using automated methods to infer representations, enabling scalable navigation through new visualizations, and building modular and disposable systems for a wide range of users at low financial cost. The talk will include motivation from previous experience at DARPA, detailed real-world examples, and software demonstrations. It will conclude by covering next directions for special projects at Microsoft.