Microsoft Research
FacultySummit

KEX: Knowledge Enabled Experiences and Semantics

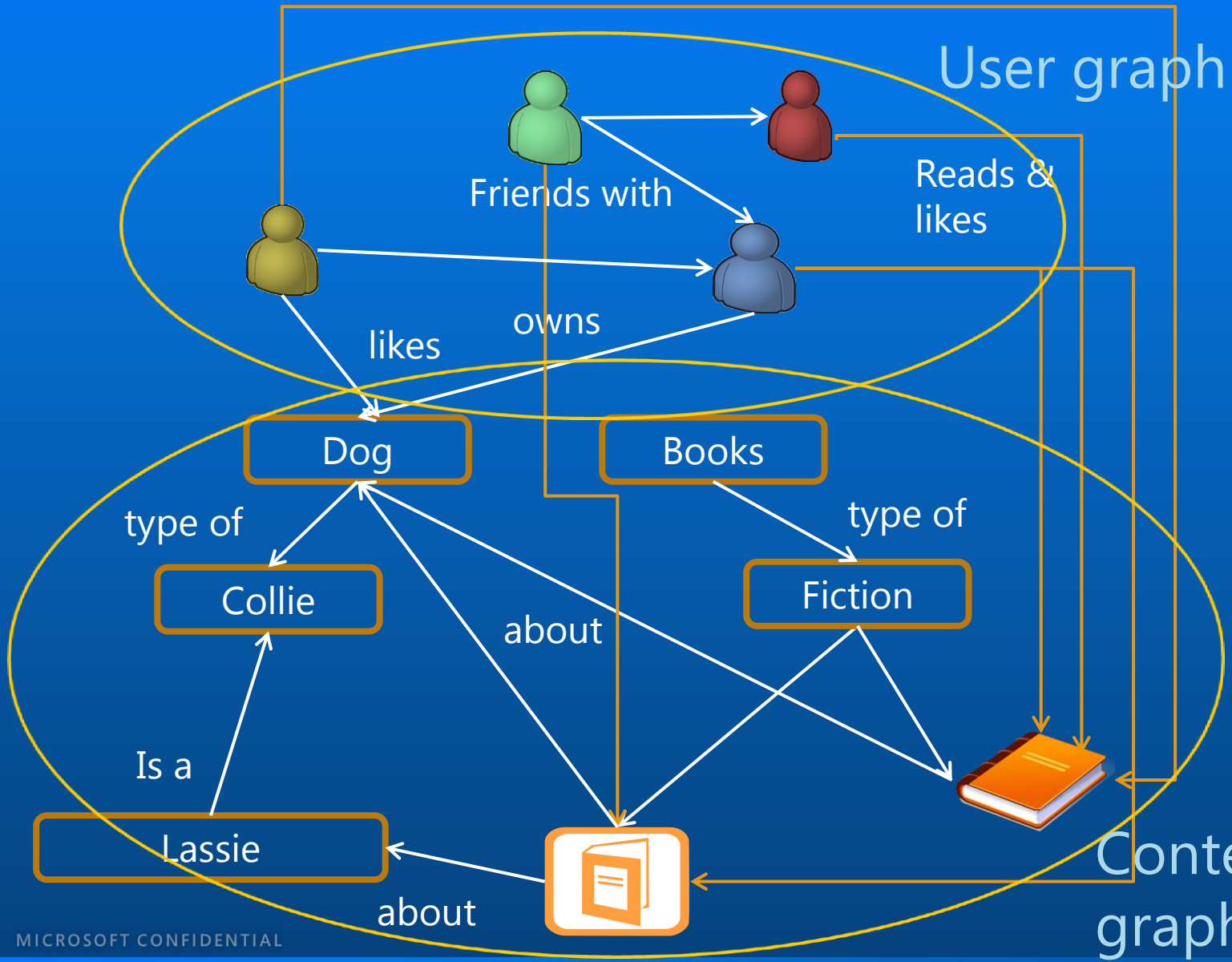
Munirathnam Srikanth
Applied Researcher
Microsoft Corporation

FUTURE/WORLD 2011 2031

Industry Trends

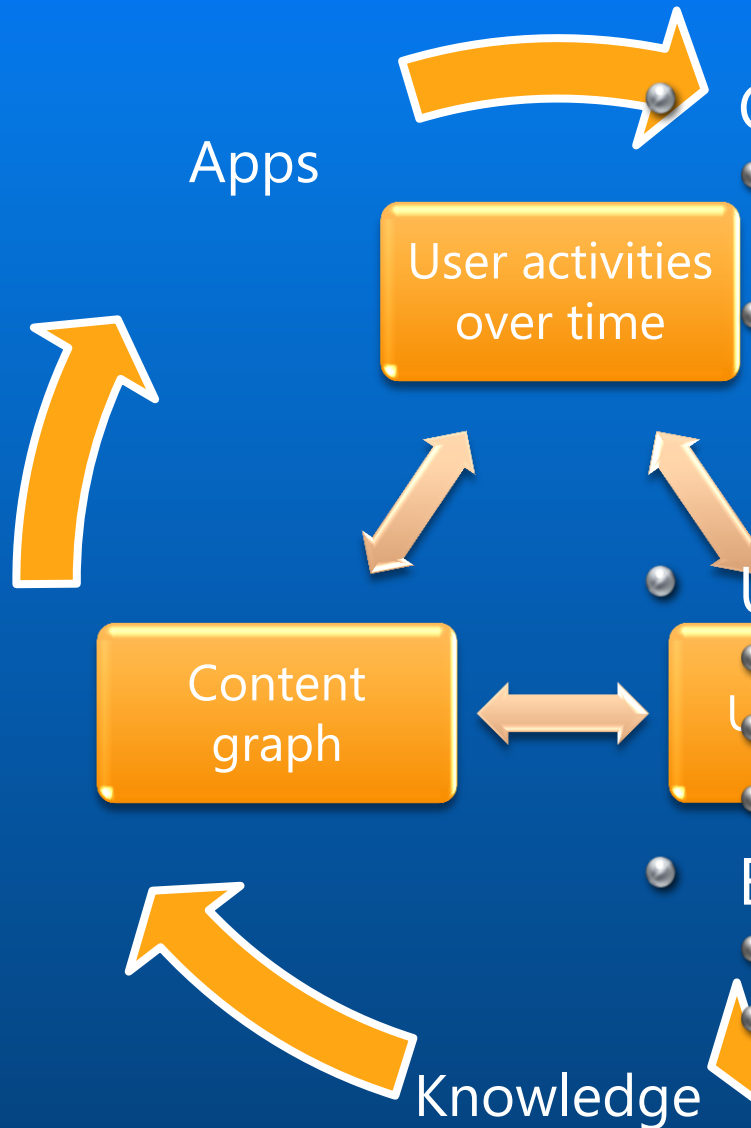
Knowledge	> Web is evolving to become knowledge centric
Structured Data	> Explosion of structured sources like Linked Data, Facebook Open Graph
Social	> Emergence of the power of social networks
Mobile & Apps	> Change in consumption behavior
Intent & Tasks	> Going beyond single queries

The Power of Connections



- Knowledge
  - Semantics
- Discovery
  - Finding connections
- Experiences
- Network effects
  - "The Long Tail", Chris Anderson

Knowledge Web



- **Content graph**
  - Concepts, Entities, their properties and relations
  - Combined statistical and semantic representation
    - supports rule-based and statistical inference
- **User Graph**
  - User Activities
  - User Interests and Models
  - Social Information
- **Experiences**
  - Entity-centric, context-aware
  - Mobile platforms

Knowledge Web

Re-imagine the next phase in the evolution of search and engineer the assets necessary towards that goal

Data Assets	Engines	Experiences
• Knowledge bases – domain focused and broad • Ontologies • Data processing pipeline	• Semantic interpretation engines • Recommendation engines • Clustering, classification engines	• New knowledge enabled experiences that leverage data assets • Discovery • New presentations on mobile devices

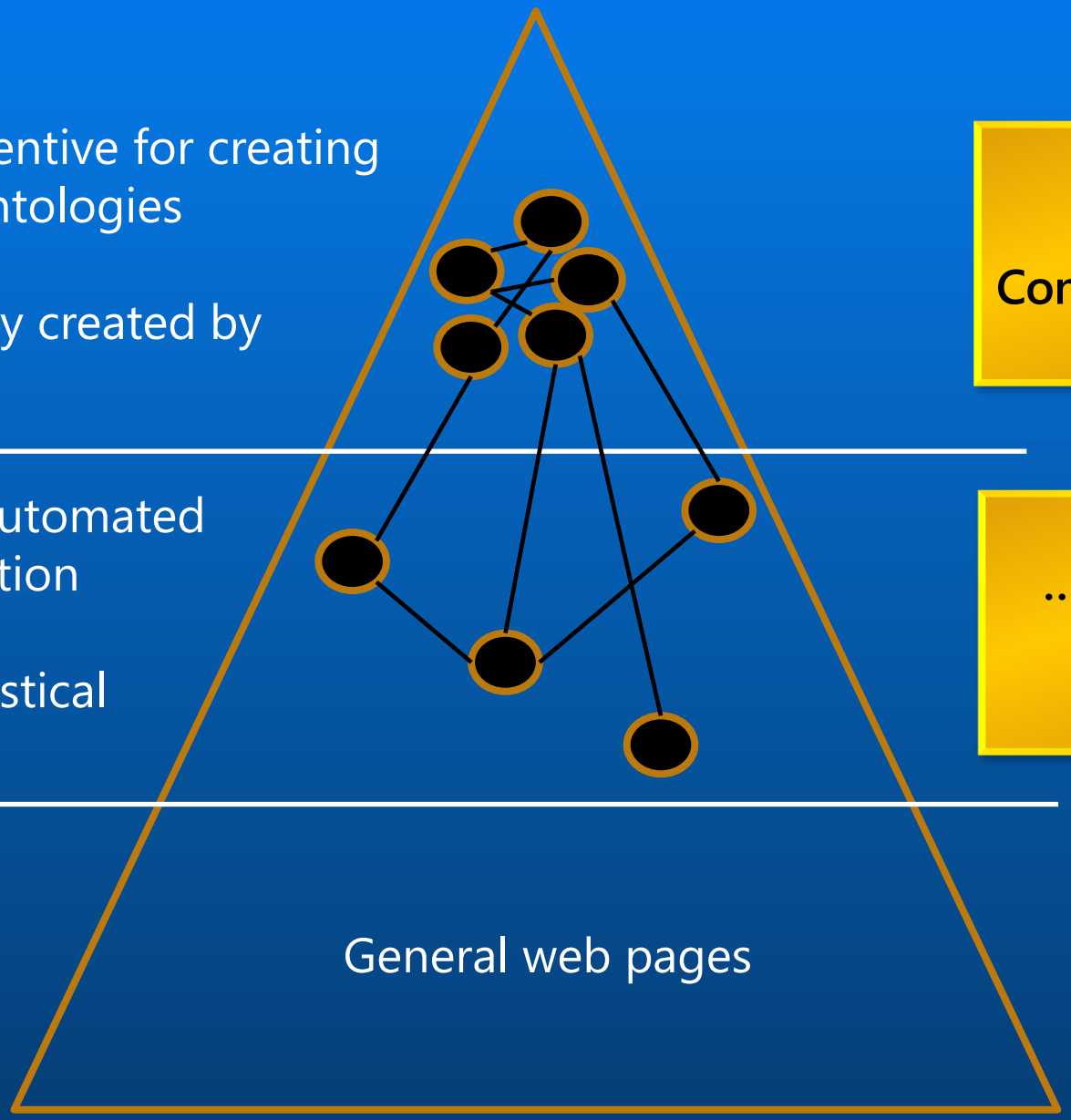
Data-Economics of Metabase

Economic incentive for creating schemas & ontologies

Collaboratively created by community

Incentive for automated (noisy) annotation

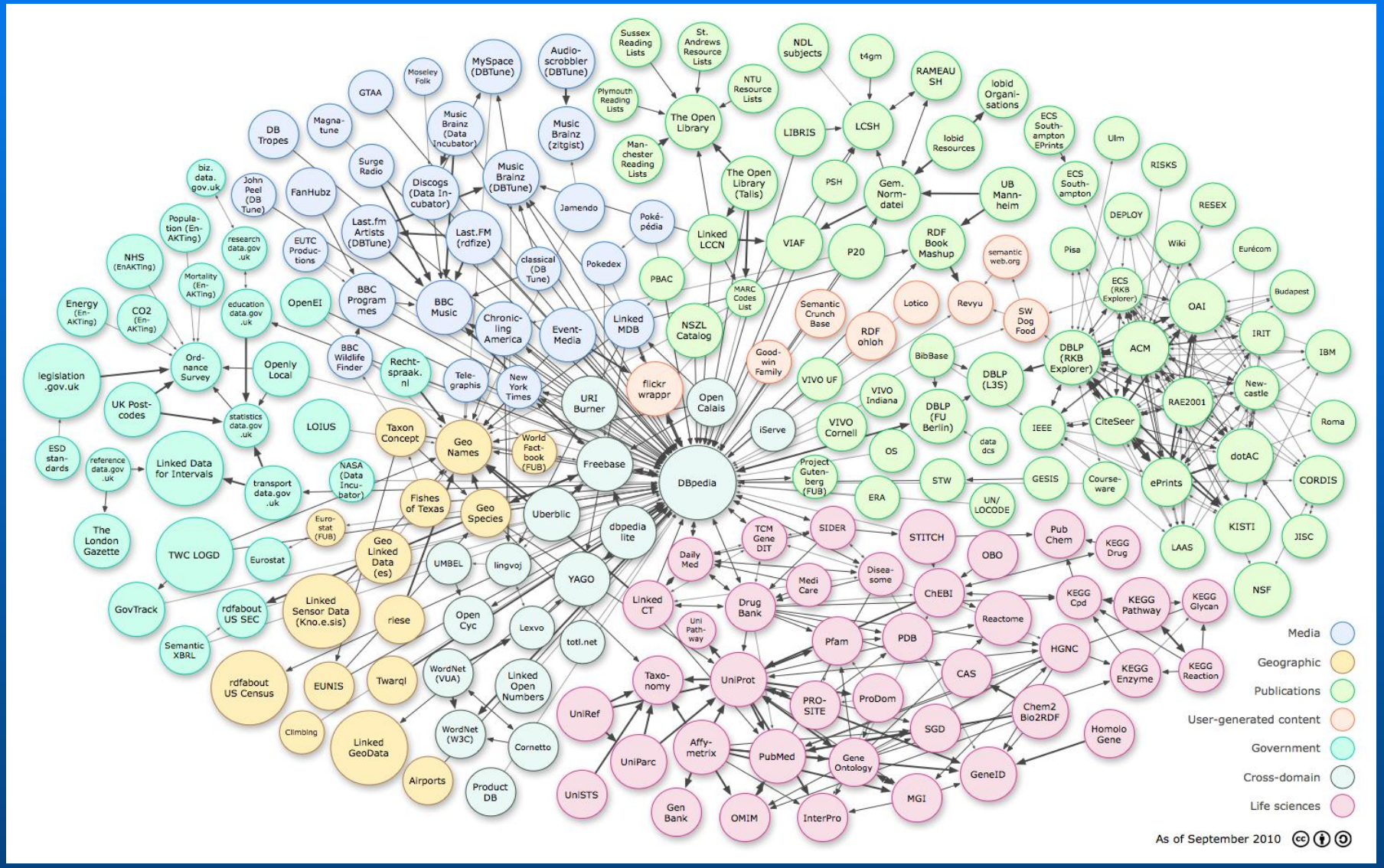
Textual & Statistical semantics



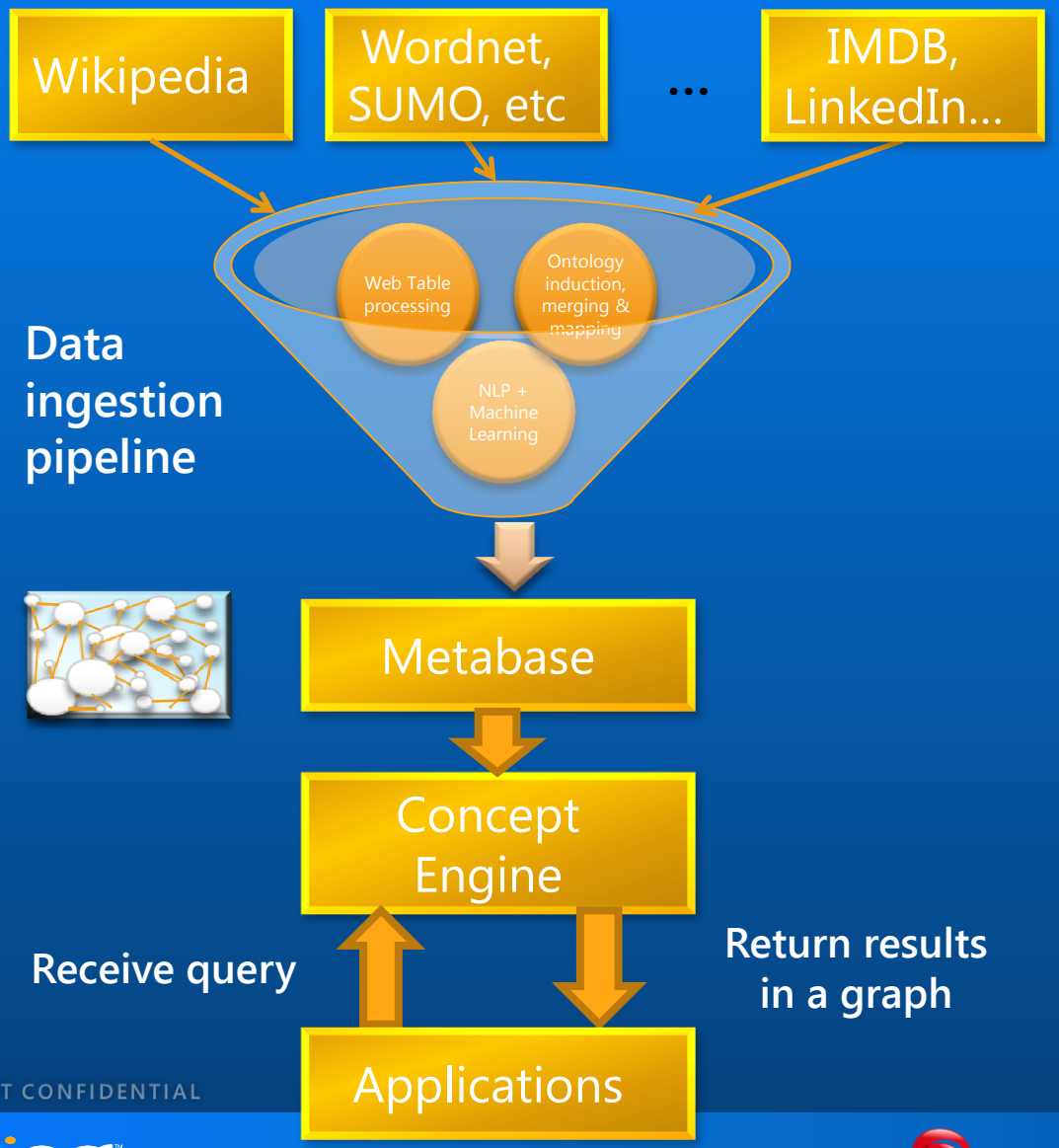
Highest "bang for the buck"
Connecting data increases value at top

… and elevates value in the middle
More value with less processing

Sources of Structured Data & Metadata

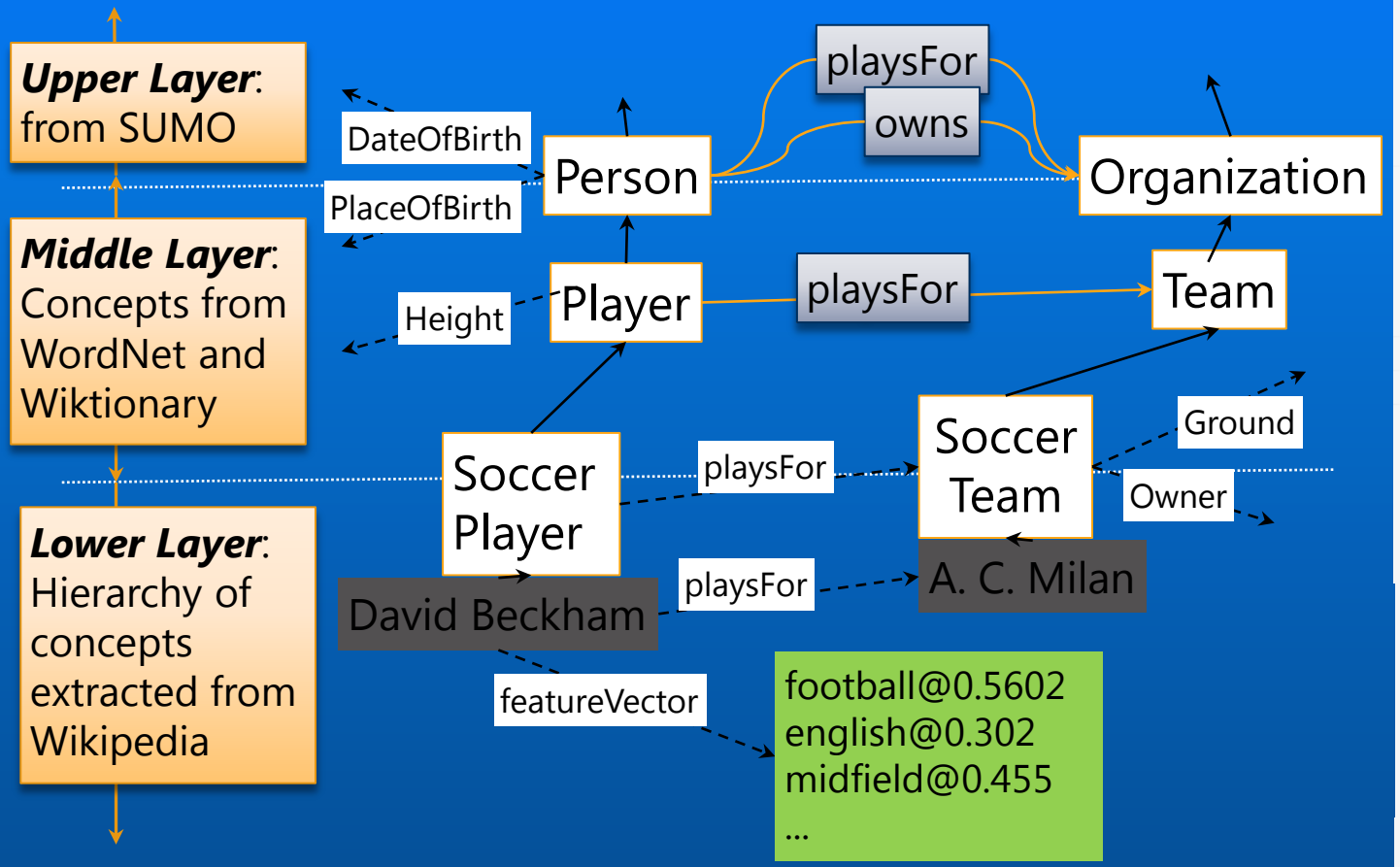


Building Concept Graph



- Ingest data from Structured and semi-structured data sources
  - Domain Specific Knowledge Bases: Finance
  - Broad-coverage Knowledge Base: Metabase
    - Sourced from Wikipedia, WordNet, Wiktionary, SUMO, etc.
- Common Information Architecture
  - Layered representation of ontologies and knowledge base
  - Represent and organize semantic and statistical information
  - Ability to represent and project different interpretations of concepts
- Engines for Knowledge-enabled Experiences
  - Support applications with APIs for different knowledge access and interpretation
  - E.g. find entities in a query, their connections and related concepts

Metabase: Layered Ontologies



Metabase Statistics	
SUMO	20K concepts, 70K axioms or rules
WordNet	147,306 Words, 117,659 Concepts, 990,149 facts
Wiktionary	236,258 Words, 197,866 Concepts, 236K synonym relations
Wikipedia	3.6M Entities, 60M facts, 305K category-WordNet links, 200M triples in feature graph

Challenges: Ontology and Knowledge Representation

Ontology Induction	> Representation that yields well for learning from data > Projecting knowledge to find gaps in information > Learning from anomalies found in the data
Ontology Alignment	> Mapping and Merging of Ontologies and Knowledge bases
Entity Resolution	> Collapsing Entities from different sources and different associated attributes > Identifying new Entities
Validation and Verification	> Correctness and completeness of source data and ontology > information in the Knowledge base > Identifying and filling missing gaps in the KB

Social Semantics and Mobile Experiences

User Graph: Capture, represent and understand users, their connections and interactions with other users and content

Sources

- Social Networks: Facebook, Twitter, etc.
- Friend, Follow, Like, Check-in
- Semantics
  - Open Graph
  - Schema.org

Representation & Models

- Genome of Users and Things
  - Music Genome (Pandora)
- Capturing User Context
  - Current Environment and Situation, Information Need
  - Short- and long-term interests
- User Interests and Models
  - Project as graph linking concepts and users

Engines & Experiences

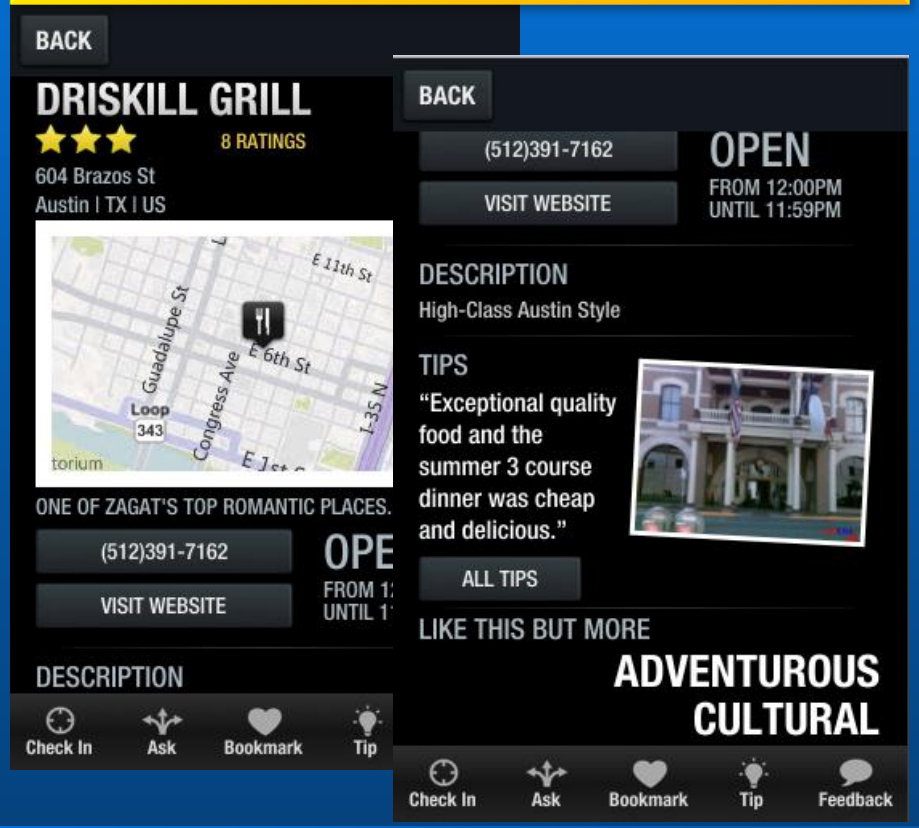
- Social in Search Engines
- Recommendation Engines
  - Hotpot (Google)
  - Simon (Microsoft) - Decision at a Glance
- Mobile experience
  - More contextual data (e.g. geo), need more contextual understanding
  - Proactive recommendations (collaborative + content + social filtering)
  - Social experience

Simon: Knowledge-enabled Mobile Experience

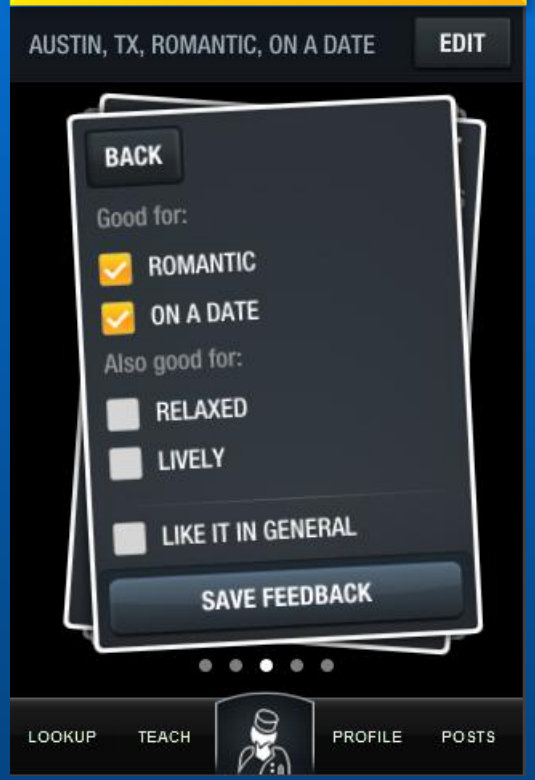
Recommend restaurants and events on mobile platform with minimal or no query box user experiences.

Decision at a glance

Details and contextual drill-down



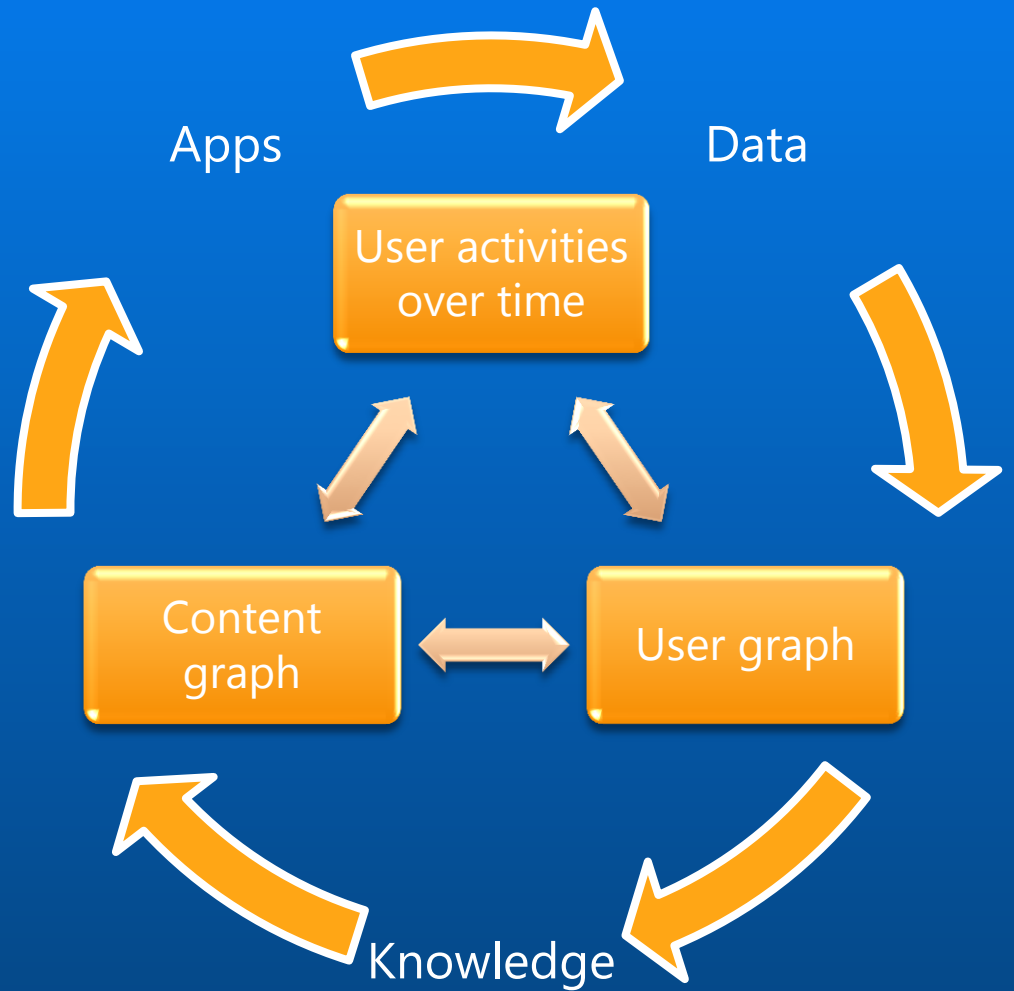
User Feedback to learn from



Challenges: Social Semantics and Mobile Experiences

Interpreting Links	Not all likes are Equal Interpreting User Links to suggest actions E.g. *Can John provide better answers for this question?*
Modeling Users and their Interests	Capturing and deriving common representation from different sources of user interaction Modeling user interests over time
User Context	Learning from different sources of user actions (social network, search logs) Context and Consumption Behavior (Mobile settings)
Recommendation	Enabling Contextual Discovery Genome of Users and Things

In Summary



- Data-Economics Pyramid
  - Exploit structured and semi-structured data
  - Focus on Information Organization and Integration
- Layered Semantic Representation
  - Ontology Alignment and Entity Resolution
  - Combine Semantics and Statistics
- Social and Mobile Experiences
  - Modeling User Context and Interests
  - Entity-centric, context-aware recommendations

Microsoft Research
FacultySummit

FUTURE/WORLD

2011 — 2031