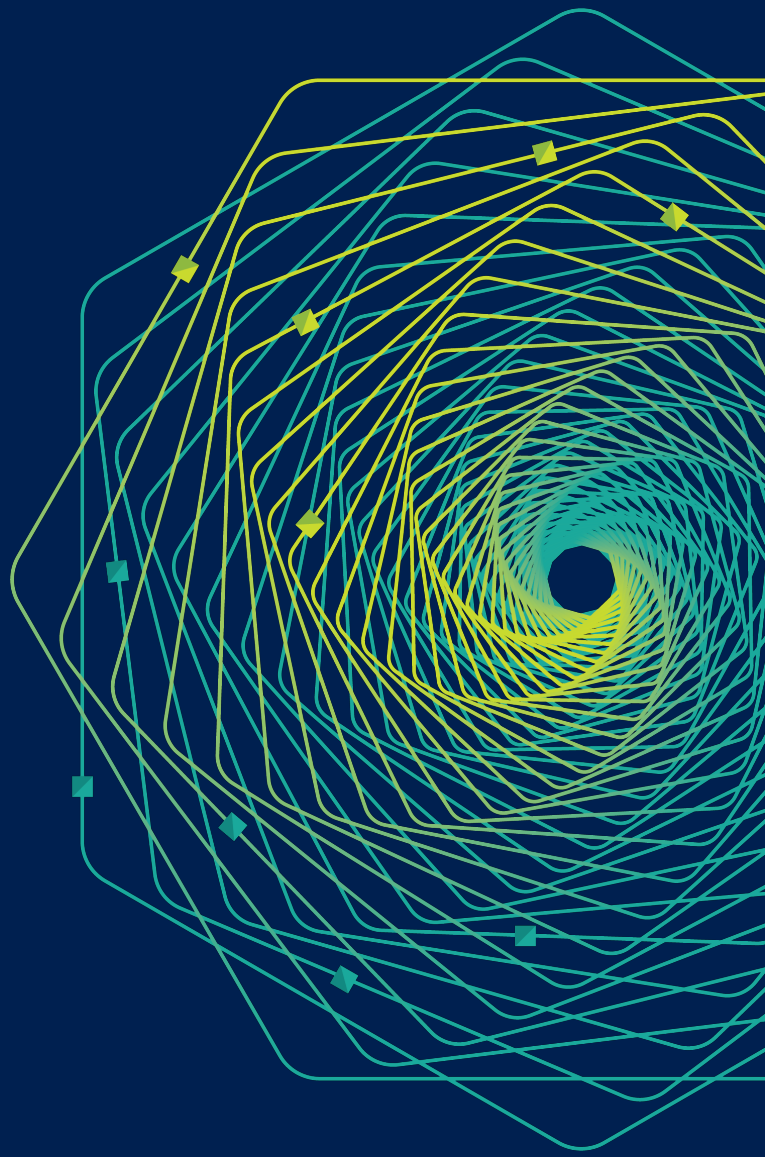
Enterprise-scale knowledge graphs

Panel: Enterprise-scale knowledge graphs

- **Yuqing Gao, Microsoft**
- **Anant Narayanan, Facebook**
- **Alan Patterson, eBay**
- **Jamie Taylor, Google**
- **Anshu Jain, IBM**

Microsoft

Knowledge Systems



Yuqing Gao
AI and Research
Microsoft

Knowledge systems: World graph

Search queries, views, click throughs, …

Web pages, Web documents, Images, …

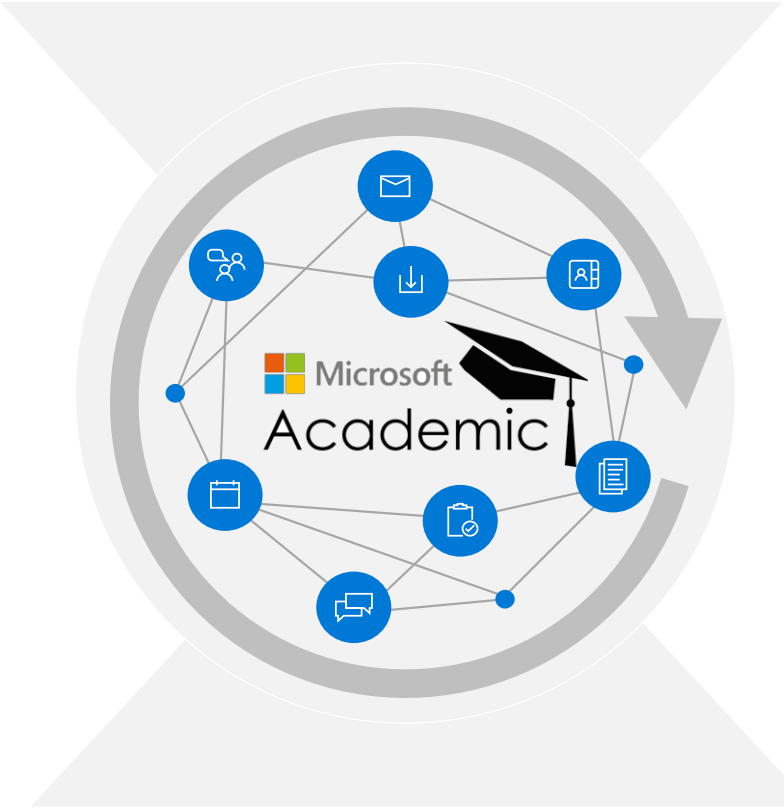
World graph

- People
- Places
- Organizations
- Things
- Actions

…

Knowledge systems: Academic graph

Knowledge acquisition, search, recommendation …



Academic graph

- People
- Publications
- Fields of Study
- Venues

…

Authors, institutions, articles, conferences …

Knowledge systems: Work graph

Messages read/sent, Document author/shared, …

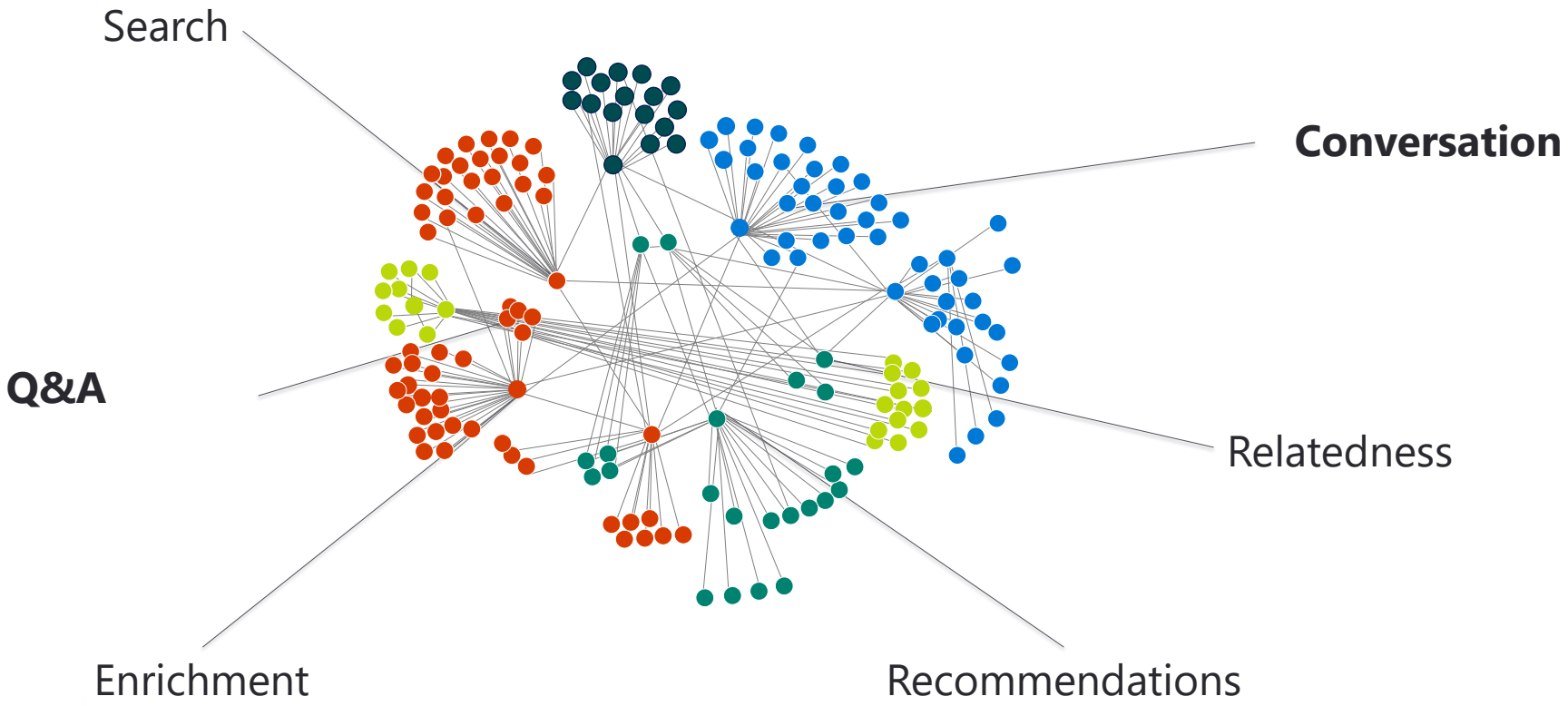
Work graph

- People
- Groups
- Messages
- Activities

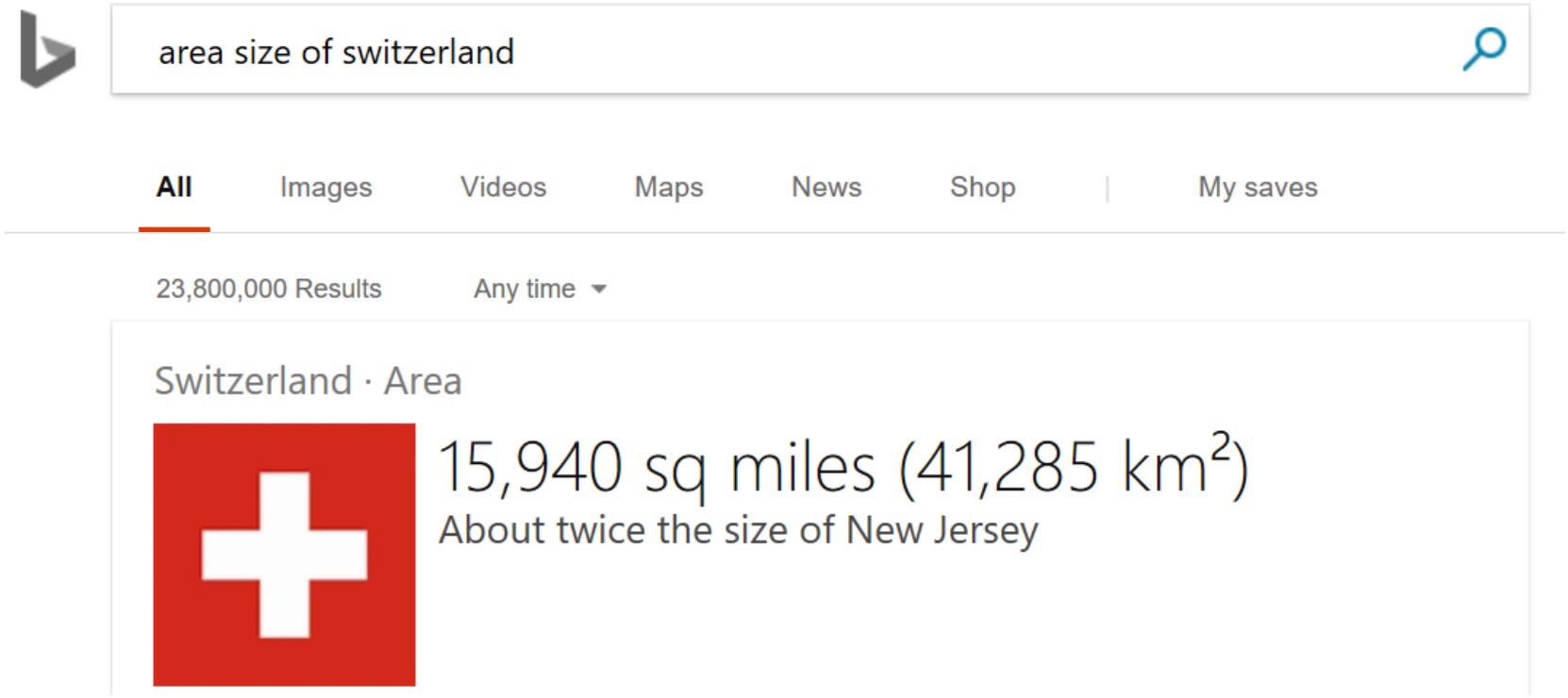
…

Emails, Messages, Documents, Meetings, …

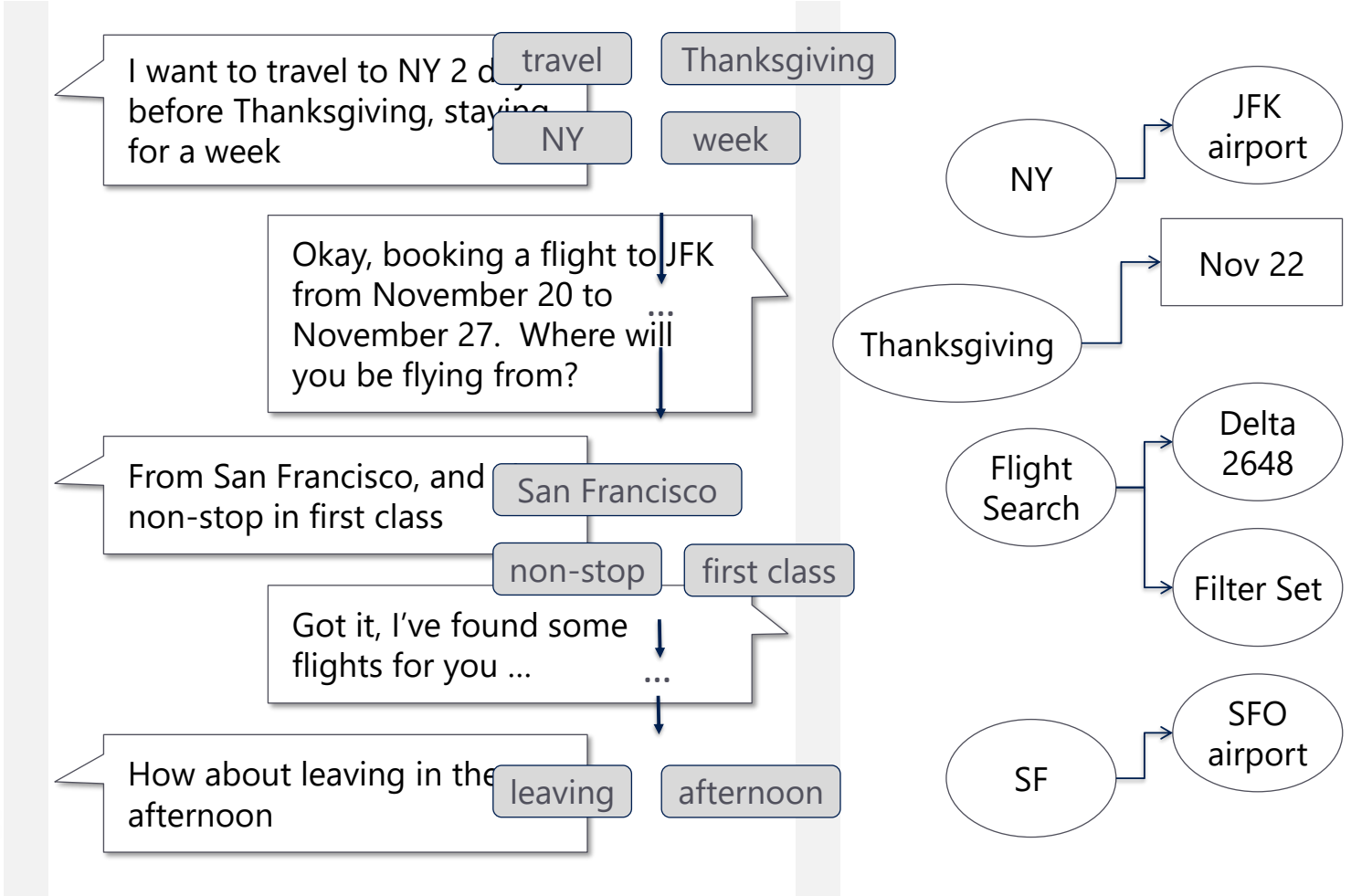
Infusing knowledge: From search to conversation



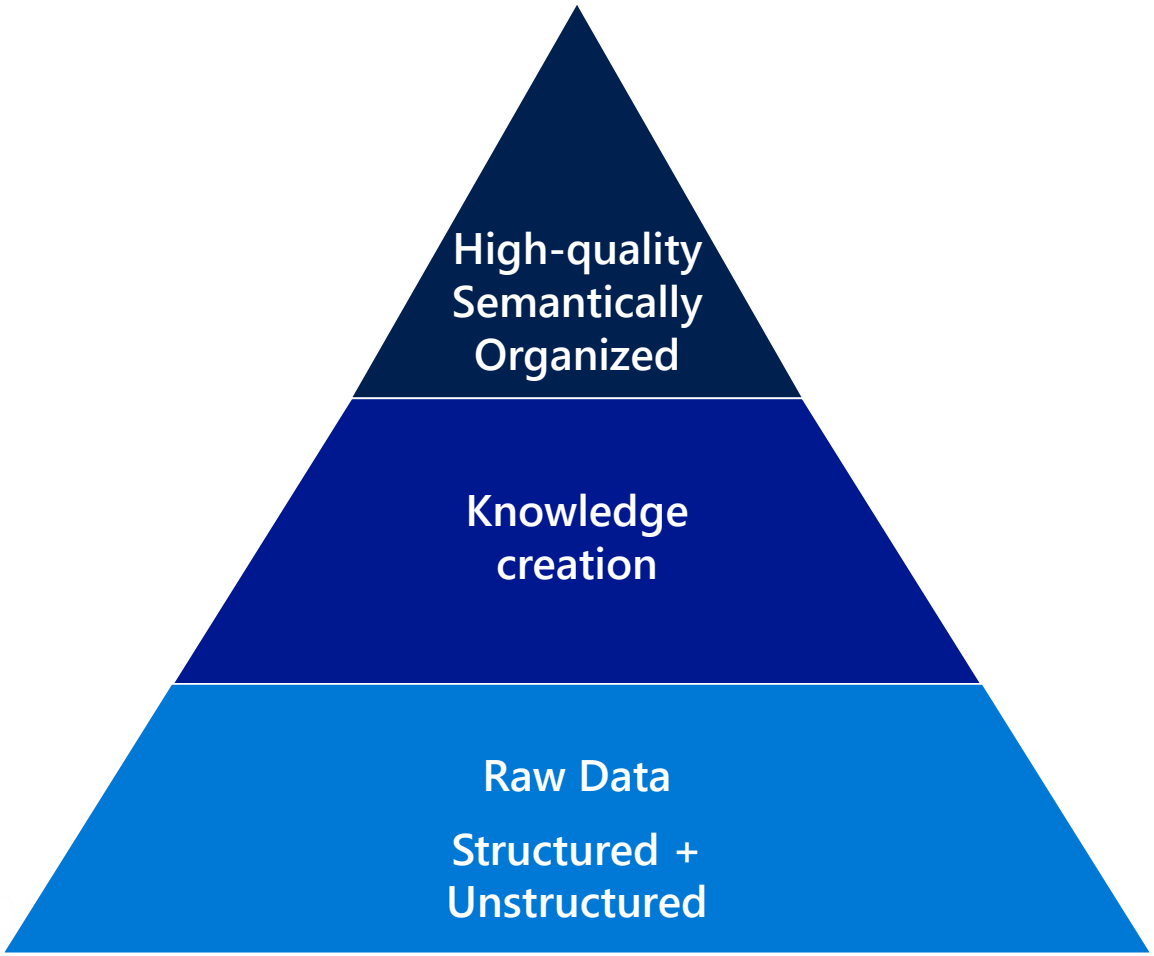
Bing – knowledge in answers



Knowledge in Conversation

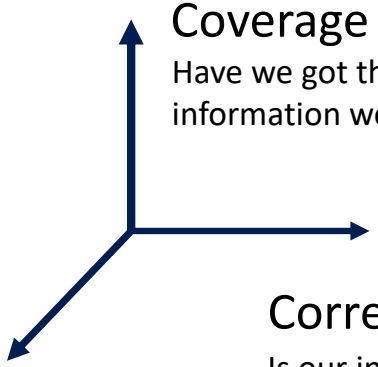


How do we bring knowledge systems to life?



Challenges of scaled KGs

Building a small KG is easy - building a vast system like Satori is a huge challenge



Coverage
Have we got the information we need?

Freshness
Is information up to date?

Correctness
Is our information accurate?

Three forces in constant conflict:

Increased freshness and coverage	⇨ Harder to ensure correctness
Increased correctness	⇨ Harder to ensure freshness and coverage

Correctness is *always* hard – what is true and correct? Particularly critical in today's world

Will Smith: Single entity, 108K facts assembled from 41 web sites.
There are 200 Will Smiths on Wikipedia alone.

Active research and product efforts in knowledge

- Unsupervised knowledge extraction from unstructured data in open domain
- Knowledge graph semantic embedding
- Autonomous knowledge inference & verification
- Real-time knowledge graph with archiving
- Large scale entity linking and disambiguation
- Ultra-scale knowledge representations
- Knowledge system for multi-lingual
- Knowledge Precision vs Comprehensiveness
- …

Thank you!


facebook

Why a Knowledge Graph for FB?

Facebook is known for the **world's largest social graph** spanning *billions* of entities and *trillions* of assertions.

We've built technology over the last decade to enable rich connections between users and we're eager to apply it to build a deeper understanding of not just people, but also **the things that people care about**.

What do we put in it?

The most **socially relevant** entities, concepts and words.

People, Celebrities
Places, Points of Interest
Movies, TV, Music
Sports

How large is it?

Deliberately focused and small to start with, US only.

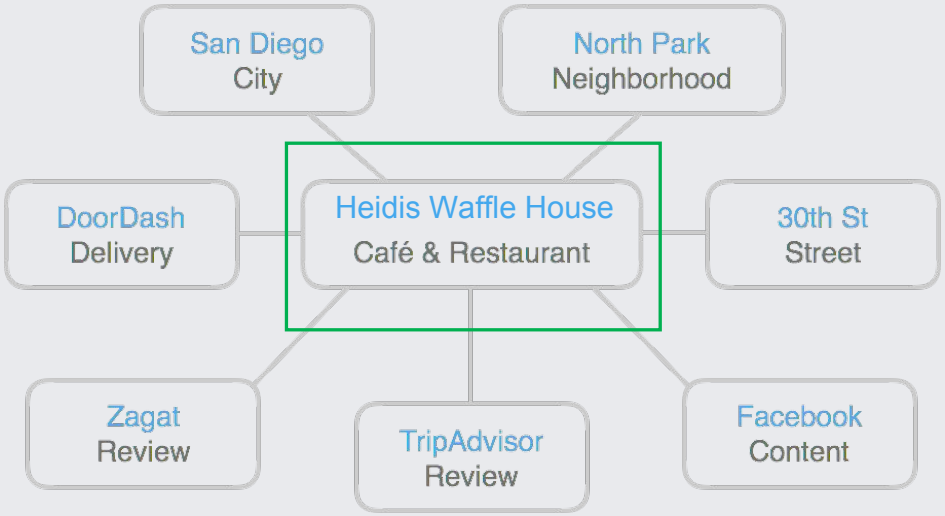
~500M assertions

~50M entities

~100s of types

Knowledge of the world

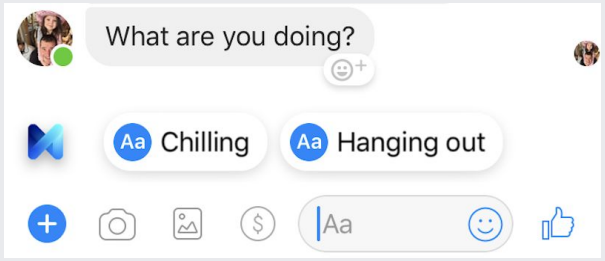
People, Places, Things



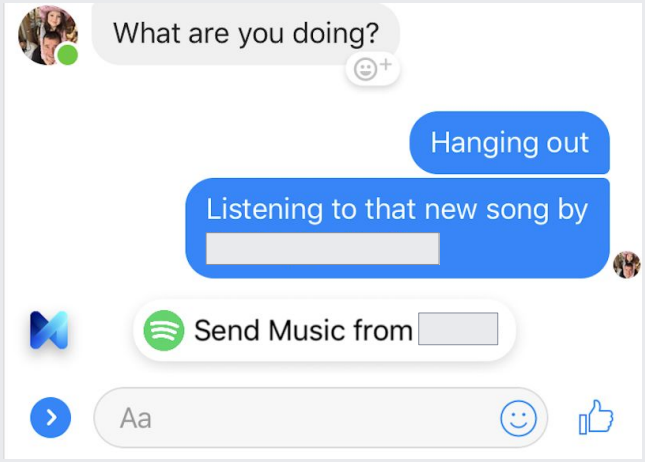
Attribute: adventurous, casual, sustainable, trendy
Dish: coffee and tea, bread, drink, parfait, cake, belgian waffle, gingerbread, liege waffle, turkey sandwich, fresh-squeezed lemonade, bacon waffle
Features: Credit cards, Takeout, Wifi, -Parking, -ADA
Meals: Breakfast, Lunch
Suggestions: liege waffle, lemonade
Telephone: (555) 987-1234
Hours: { … }
Website: http://www.heidiswafflehouse.com

Use it to provide utility...

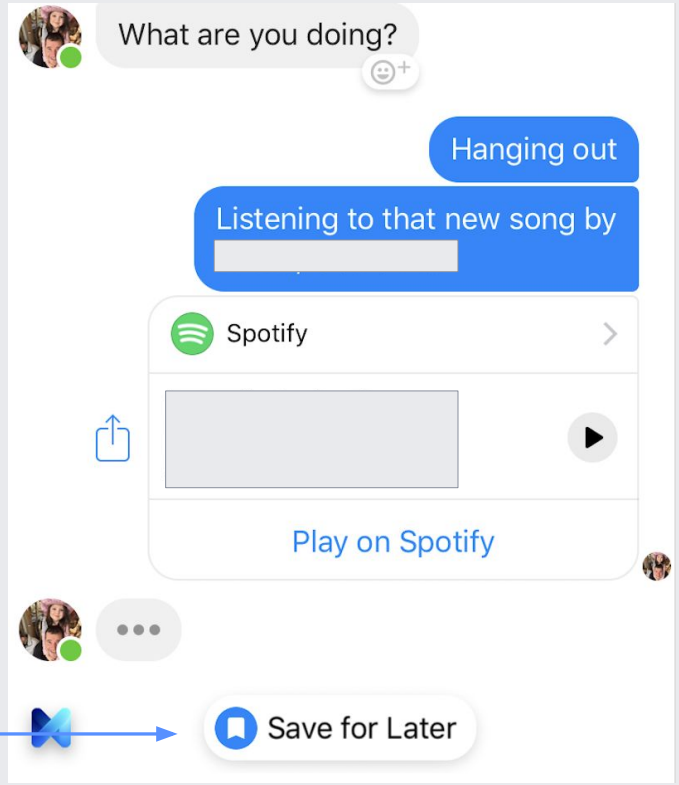
Recommend smart replies



Entity Detection

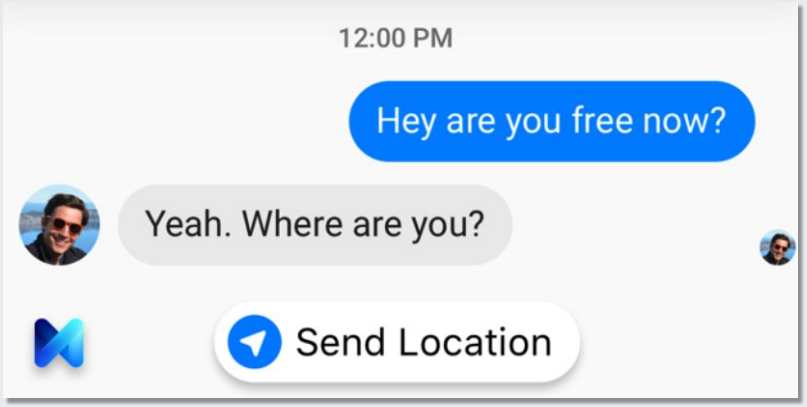


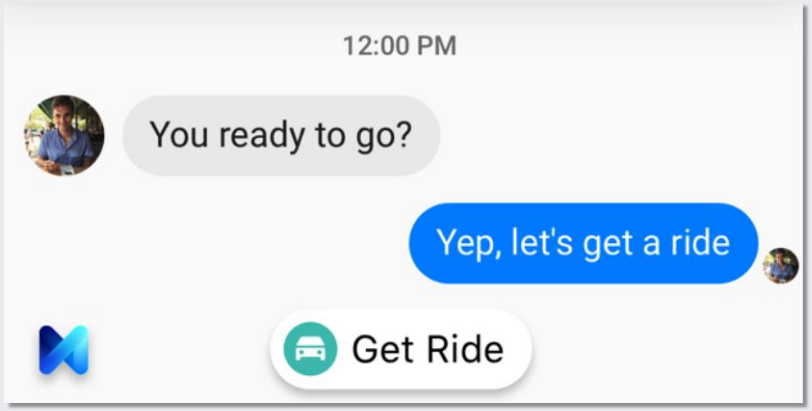
Easy sharing

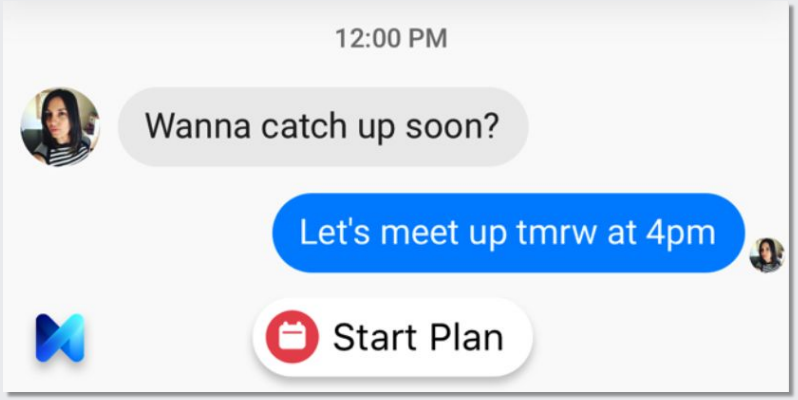


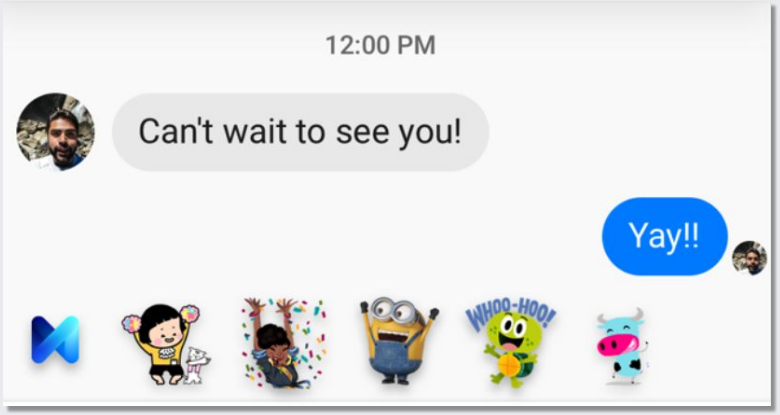
Memory

… and delight the user!







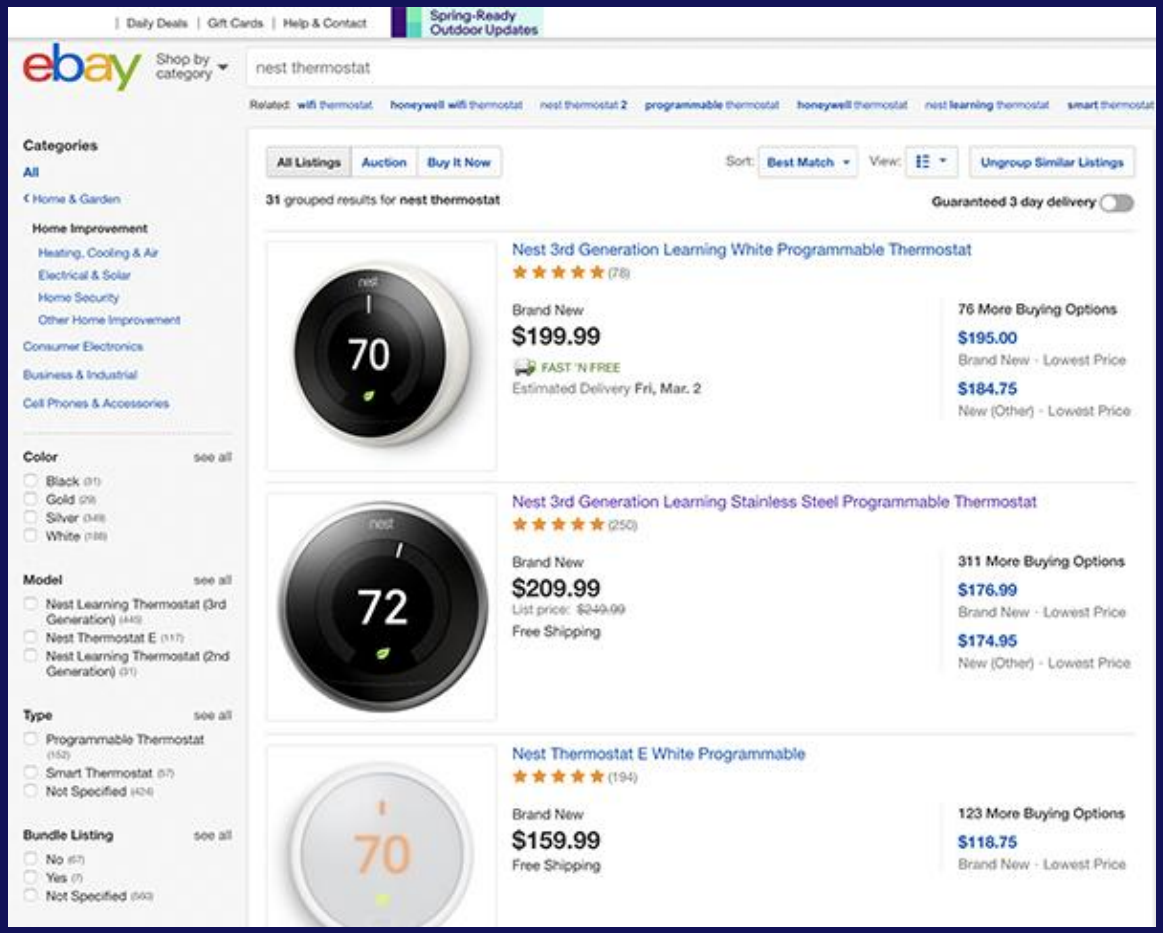


eBay Product Knowledge Graph

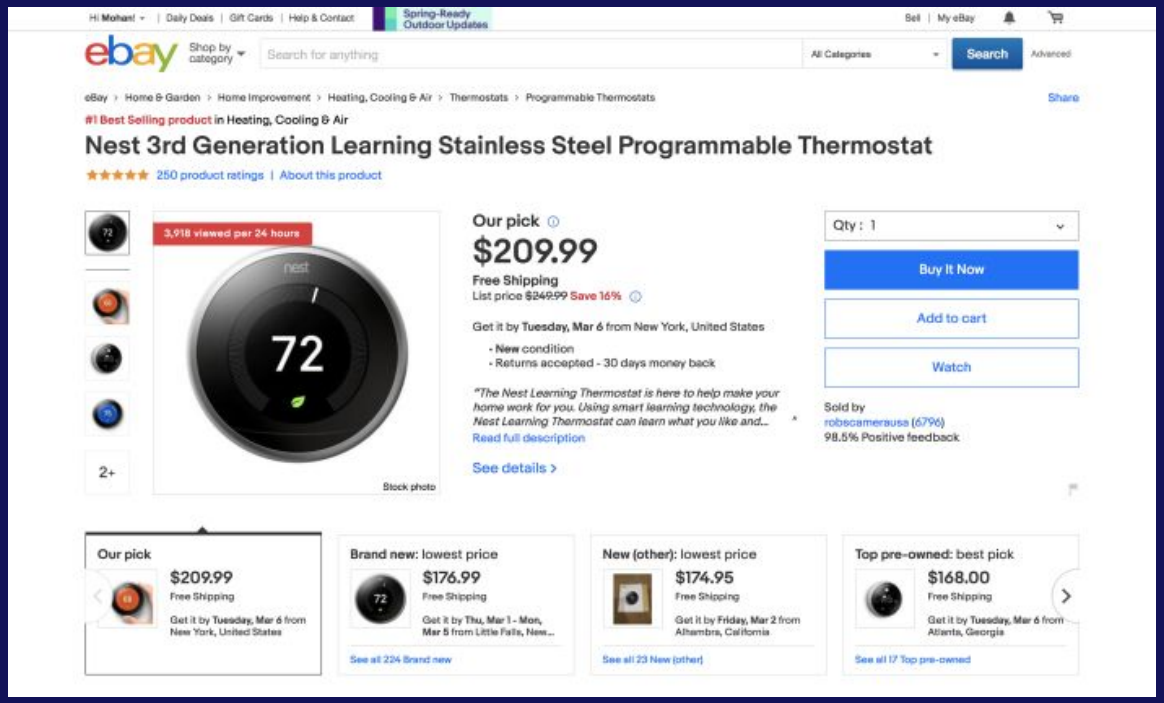
ISWC 2018

Alan Patterson, Structured Data
alpatterson@ebay.com

ebay


Daily Deals | Gift Cards | Help & Contact
Spring-Ready Outdoor Updates
ebay
Shop by category
nest thermostat
Related: wifi thermostat honeywell wifi thermostat nest thermostat 2 programmable thermostat honeywell thermostat nest learning thermostat smart thermostat
Categories
All
Home & Garden
Home Improvement
Heating, Cooling & Air
Electrical & Solar
Home Security
Other Home Improvement
Consumer Electronics
Business & Industrial
Cell Phones & Accessories
Color see all
Black (31)
Gold (29)
Silver (348)
White (188)
Model see all
Nest Learning Thermostat (3rd Generation) (445)
Nest Thermostat E (117)
Nest Learning Thermostat (2nd Generation) (31)
Type see all
Programmable Thermostat (152)
Smart Thermostat (57)
Not Specified (424)
Bundle Listing see all
No (67)
Yes (7)
Not Specified (593)
All Listings
Auction
Buy It Now
Sort: Best Match
View:
Ungroup Similar Listings
31 grouped results for nest thermostat
Guaranteed 3 day delivery
Nest 3rd Generation Learning White Programmable Thermostat
(78)
Brand New
$199.99
FAST 'N FREE
Estimated Delivery Fri, Mar. 2
76 More Buying Options
$195.00
Brand New · Lowest Price
$184.75
New (Other) · Lowest Price
Nest 3rd Generation Learning Stainless Steel Programmable Thermostat
(250)
Brand New
$209.99
List price: $249.99
Free Shipping
311 More Buying Options
$176.99
Brand New · Lowest Price
$174.95
New (Other) · Lowest Price
Nest Thermostat E White Programmable
(194)
Brand New
$159.99
Free Shipping
123 More Buying Options
$118.75
Brand New · Lowest Price

ebay


Hi Mohan! | Daily Deals | Gift Cards | Help & Contact
Spring-Ready Outdoor Updates
Sell | My eBay
ebay
Shop by category
Search for anything
All Categories
Search
Advanced
eBay › Home & Garden › Home Improvement › Heating, Cooling & Air › Thermostats › Programmable Thermostats
Share
#1 Best Selling product in Heating, Cooling & Air
Nest 3rd Generation Learning Stainless Steel Programmable Thermostat
250 product ratings | About this product
3,918 viewed per 24 hours
Stock photo
2+
Our pick
$209.99
Free Shipping
List price $249.99 Save 16%
Get it by Tuesday, Mar 6 from New York, United States
- New condition
- Returns accepted - 30 days money back
*The Nest Learning Thermostat is here to help make your home work for you. Using smart learning technology, the Nest Learning Thermostat can learn what you like and...
Read full description
See details >
Qty : 1
Buy It Now
Add to cart
Watch
Sold by
robscamerause (6796)
98.5% Positive feedback
Our pick
$209.99
Free Shipping
Get it by Tuesday, Mar 6 from New York, United States
Brand new: lowest price
$176.99
Free Shipping
Get it by Thu, Mar 1 - Mon, Mar 5 from Little Falls, New...
See all 224 Brand new
New (other): lowest price
$174.95
Free Shipping
Get it by Friday, Mar 2 from Alhambra, California
See all 23 New (other)
Top pre-owned: best pick
$168.00
Free Shipping
Get it by Tuesday, Mar 6 from Atlanta, Georgia
See all 17 Top pre-owned

ebay

Memorabilia & Collectables

Michael Jordan
Basketball
Chicago Bulls

Paris
France
Europe
Wartime

ebay

Pipeline

Graph - Products, properties, value types, fitment, relationships, standards, variations, people, places, brands, companies, events, dates.

Statistical Data Mining - Discovery (over billions of listings)

Buyer & Seller Data - Queries, clicks & conversions, seller descriptions

Knowledge Matching and Semantic Verification

Graph stats - Expect around 100M products, >1Bn triples

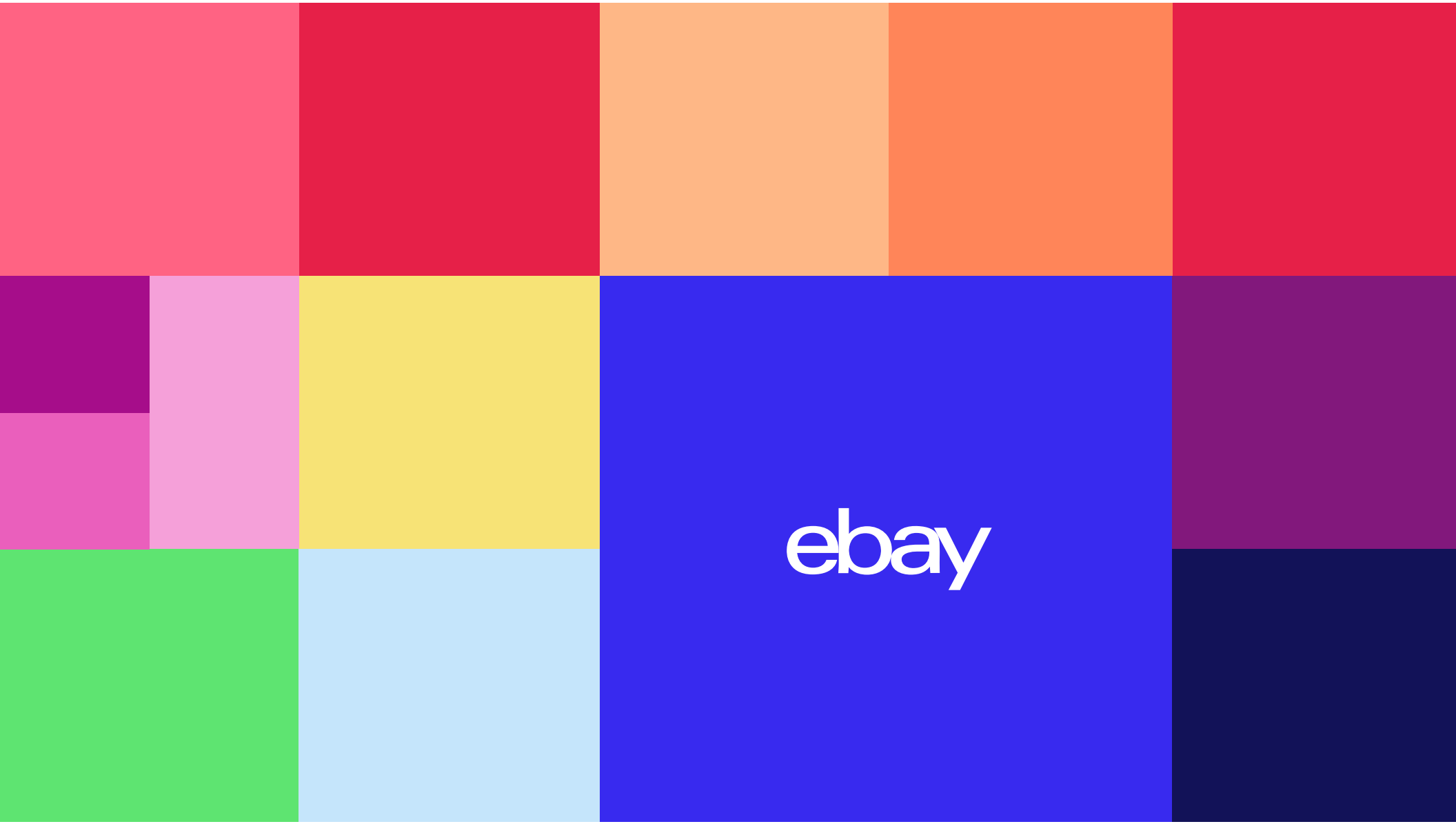
Problems

Biggest Problem: Identity

Are these listings the same product

Identity depends on who's asking

Buyer and Seller have different requirements


ebay

Knowledge Graph

Jamie Taylor
Google - Knowledge Graph Schema Team

ISWC 2018

Google

Knowledge Graph

1 Billion Things

70 Billion Facts

What is it used for

Ocean sunfish

Fish

The ocean sunfish or common mola is the heaviest known bony fish in the world. Adults typically weigh between 247 and 1,000 kg. The species is native to tropical and temperate waters around the globe. It resembles a fish head with a tail, and its main body is flattened laterally. Wikipedia

Mass: 2,200 lbs (Adult)

Scientific name: Mola mola

Length: 5.9 ft. (Adult)

Speed: 2 mph (Maximum, Adult)

Family: Molidae

Did you know: By basking on its side at the surface, the sunfish also allows seabirds to feed on parasites from its skin. eol.org

People also search for View 15+ more


Blobfish


Sunfishes


Sunfish


Bluegill


Molas

Sunfish

Sailboat

The Sunfish sailboat is a personal size, beach launched sailing dinghy utilizing a pontoon type hull carrying a lateen sail mounted to an un-stayed mast. Sunfish was developed by Alcort, Inc. and first appeared around 1952 as the "next generation" improvement on their original boat, the Sailfish. Wikipedia

Hull weight: 120 pounds (54 kg)

Rig: Lateen

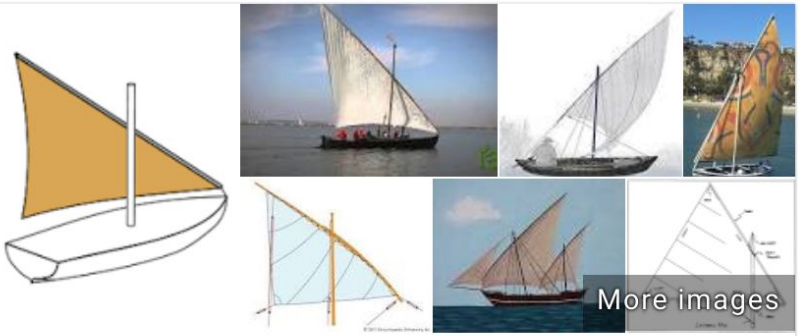
Total sail area: 75 square feet (7.0 m^2)

Keel: Daggerboard

Draft: 2 feet 11 inches (0.89 m)

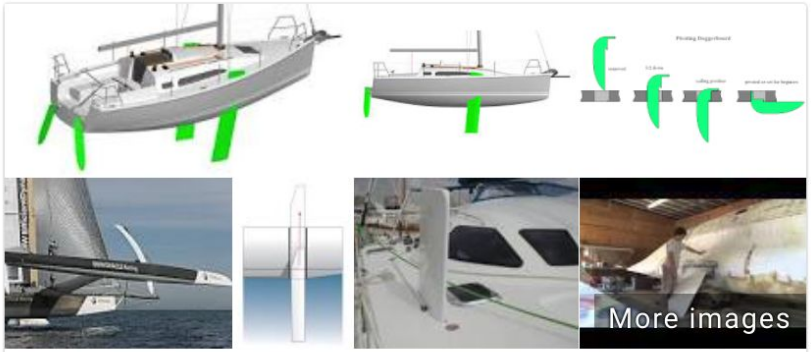
Year: 1953

Construction: Fiberglass



Lateen

A lateen or latin-rig is a triangular sail set on a long yard mounted at an angle on the mast, and running in a fore-and-aft direction. Dating back to Roman navigation, the lateen became the favorite sail of the Age of Discovery, mainly because it allows a boat to tack "against the wind."
Wikipedia



Daggerboard

A daggerboard is a retractable centreboard used by various sailing craft. While other types of centreboard may pivot to retract, a daggerboard slides in a casing. The shape of the daggerboard converts the forward motion into a windward lift, countering the leeward push of the sail.
Wikipedia




2018 Mazda MX-5 Miata

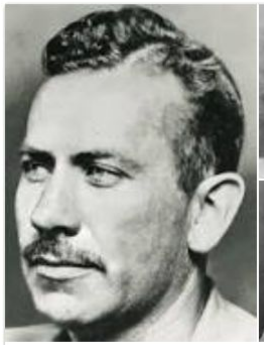
Sports car


Monterey

City in California

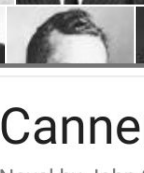
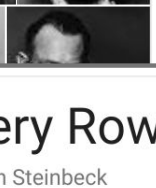
Euro

Currency







John Steinbeck

American author

Metaweb Technologies, Inc.

Company

METAWEB

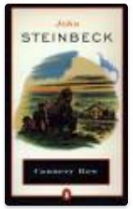
Cannery Row

Novel by John Steinbeck

Book preview

16/106 pages available

PREVIEW








Printing press

Google

Challenges

Challenges

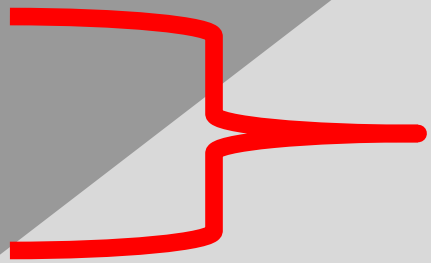
What is personally challenging

- **Long evolution**
- **Bigger than my brain**
- **So many stakeholders**

Challenges

What is personally challenging

Long evolution
Bigger than my brain
So many stakeholders



Wonderful Problems!

Challenges

What is personally challenging

Long evolution
Bigger than my brain
So many stakeholders

Survival Mechanism:
A Principled Approach

Managing Identity

Google

Managing Identity

Santos FC Before Pelé

Sailing

Cuba

Managing Identity

Santos FC Before Pelé

Sailing

Cuba

Maintaining Semantic Stability

Managing Identity

Santos FC Before Pelé

Sailing

Cuba

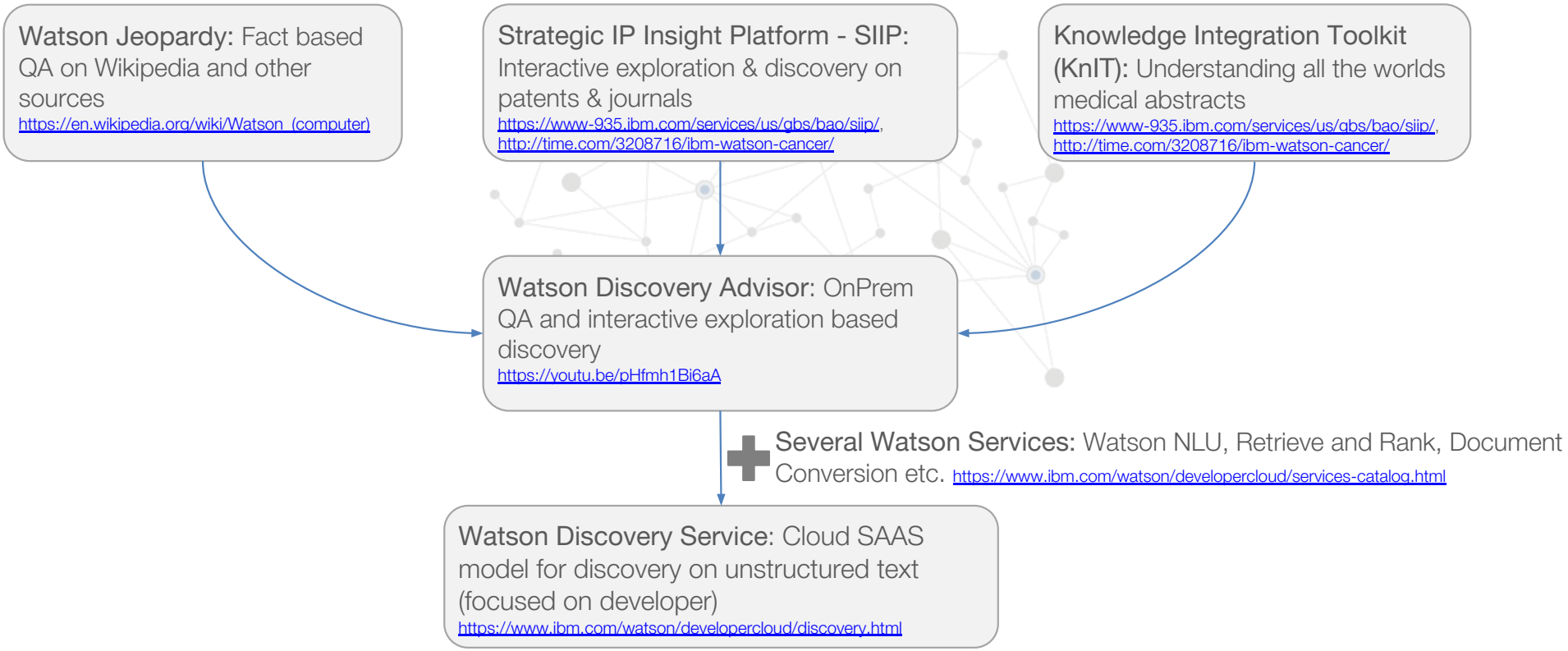
Maintaining Semantic Stability

E-Sports & Sports

Movies & Television

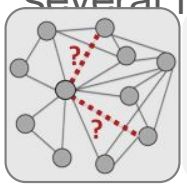
Broadcast, Podcast & Streaming

Knowledge Graph @ Watson - Context

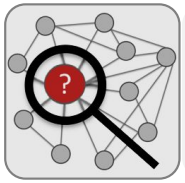


Discovery is about the non-obvious

Discovery creates new knowledge (hitherto unknown / unrecorded in domain docs, data sources). New knowledge is surprising and anomalous. It could be formally abstracted in several forms:



New link between entities: A new side effect of a drug, a new potential emerging company as an acquisition target or sales lead, a non-obvious but relevant case / law applicable in a particular trial / proceeding, a person of interest in a terrorist attack. (Link Prediction, Relationship Discovery, Relationship Ranking)



A potential new important entity in the domain: e.g. a new material for display technologies, a new investor for a particular investment area etc. (Entity Discovery, Entity Recommendation, Entity Ranking)

Changing significance of an existing entity: it's changing relationships/attributes/metrics. e.g. increasing stake by an investor in an organization, increasing interaction between a person of interest and some criminal in an intelligence gathering scenario, decreasing number complaints for a particular product or service in a retail / consumer scenario. (Trend analysis, Distribution Analysis, Anomaly Detection)

Search and Exploration look for knowledge already available in the knowledge sources available to the system. They are necessary for Discovery but not sufficient

Confirmation of facts known to an SME of a domain is not discovery (listing all possible side effects of a drug which may have been mined from structured or unstructured data)

In a nutshell

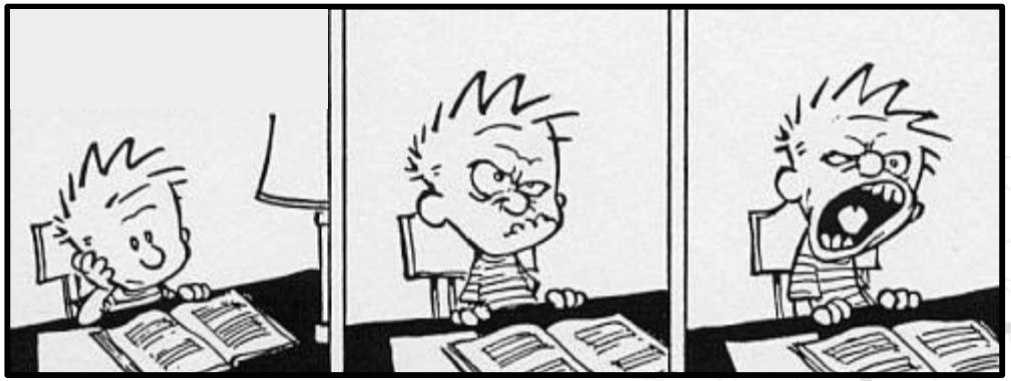
- We haven't focused on **ONE Global** knowledge graph
- We have a framework for anyone to build their domain specific KGs
- Offered as a part of Watson Discovery Advisor and now Watson Discovery Service, also consumed internally by services / solutions
- Clients In various sectors:
  - Banking and Finance
  - IT Services, Customer Service
  - Cyber Security
  - Scientific Discovery: Life Science, Oil & Gas, Chemicals & Petroleum
  - Defense
  - Space Exploration
  - Media and Entertainment

valuable things we learnt..

- *polymorphic stores is the answer, stop fighting for one store*
- *build surprise into your model, into your api*
- *evidence is primitive to the system, which every node and edge maps to*
- *push entity resolution to runtime through context*
- *every new piece of knowledge should know its cascade effect*

Detailed description can be found at talks.anshu.co

Open problems that haunt us



- *modeling and analyzing changing knowledge*
- *merge relationships discovered from unstructured data with known relationships*
- *federation of global, domain specific and customer specific knowledge*
- *incremental update of global knowledge on horizontally scaled stores*
- *scanned PDF extraction ..garbage in garbage out*

Thank you.