Teaching Machines to Fish

How eBay Improves Itself

Randy Shoup
eBay Distinguished Architect

QCon San Francisco
November 21, 2008

About the Presenter

Randy Shoup is a Distinguished Architect in eBay’s Marketplace Architecture group.

Since 2004, he has been the primary architect for eBay's search infrastructure.

Prior to eBay, Randy was Chief Architect and Technical Fellow at Tumbleweed Communications, and has also held a variety of software development and architecture roles at Oracle and Informatica.

Teaching Our Machines to Fish …

"Give a man a fish and he eats for a day …
Teach a man to fish and he eats for a lifetime"

-- Lao Tzu

… in eBay's Internet-Scale Ocean

- Massive Data Volumes
  - Over 2 billion page views per day
  - 84.5 million active members in 39 markets worldwide
  - 50 TB of new, incremental data per day
  - 50 PB of data analyzed per day
- Highly Dynamic Marketplace
  - 667 million new items per quarter in 50,000 categories
  - Roughly 10% of items are listed or ended every day
- Highly Available
  - Always on, 24x7x365

Typical Problems

- Choose the “best” inventory to show for a user’s query
- Choose the “best” user experience for a user’s results
- Recommend items, categories, sellers, search terms, etc.
- Classify and cluster incoming items
- Maintain a current dictionary of eBay vocabulary

Agenda

- Machine Learning and Feedback Loops
- Typical Feedback Loop at eBay
- Extended Example: eBay Search
- Concluding Thoughts

Agenda

- Machine Learning and Feedback Loops
- Typical Feedback Loop at eBay
- Extended Example: eBay Search
- Concluding Thoughts

Why Machine Learning?

- Alternative approaches do not scale
  - Manual configuration
  - Manual or static rules
  - Static model
- Learn and improve over time without manual effort
- Adapt to changing environment more rapidly and completely
- Consider more factors and more data in decisions
- Explore solution space more thoroughly and quickly

Machine Learning in Practice

Build systems which improve themselves automatically through experience

- Machine Learning involves
  - Improving at a given task (decision, prediction)
  - With respect to a given performance metric
  - Based on experience

    [from Machine Learning, Tom Mitchell, 1997]

- Real-world requirements
  - Learn across multiple sources of input
  - Learn by active experimentation
  - Learn both predictions and decisions
  - Learn cumulatively and continually over time

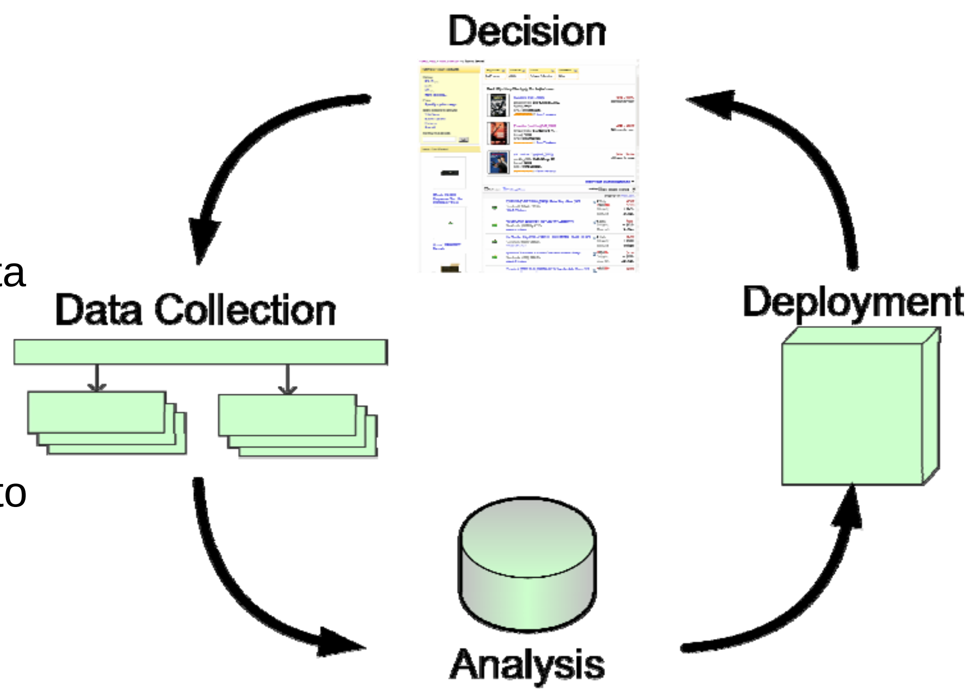
Machine Learning in Practice

- Choosing the proper performance metric is critical (!)
  - Clicks?
  - Bids?
  - Purchases?
  - Server Utilization?

Machine Learning and Feedback Loops

All learning ultimately depends on a feedback loop

- Data Collection
  - Collect what we did
  - Collect associated user behavior
  - Collect associated business metrics
- Analysis
  - Aggregate and analyze collected data
  - Update model or metadata
- Deployment
  - Deploy updated model or metadata to online system
- Decision
  - Make prediction or decision
  - Perform action
  - Actively experiment (“perturbation”)

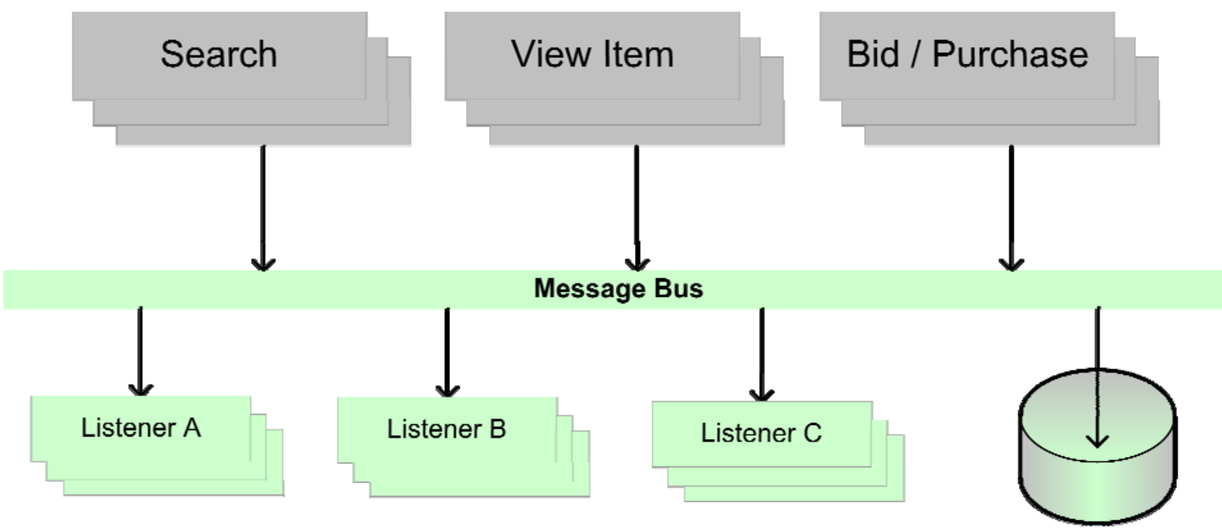


Agenda

- Machine Learning and Feedback Loops
- Typical Feedback Loop at eBay
- Extended Example: eBay Search
- Concluding Thoughts

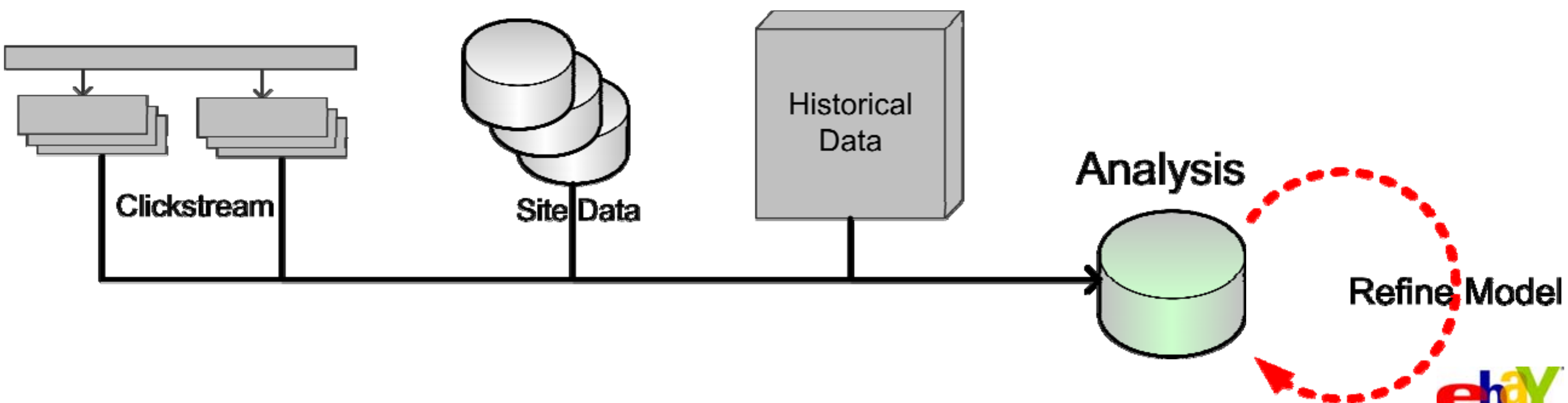
eBay Feedback Loop: Data Collection

- Application servers log user events
  - Incoming request, outgoing experience, associated data
- Events broadcast on multicast message bus
  - Partitioned horizontally by user id
- Listeners process event stream
  - Recompose clickstream from individual page events
  - Search for patterns
  - Persist for history

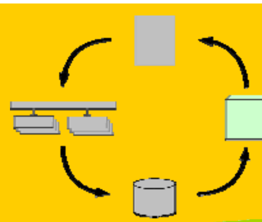


eBay Feedback Loop: Analysis

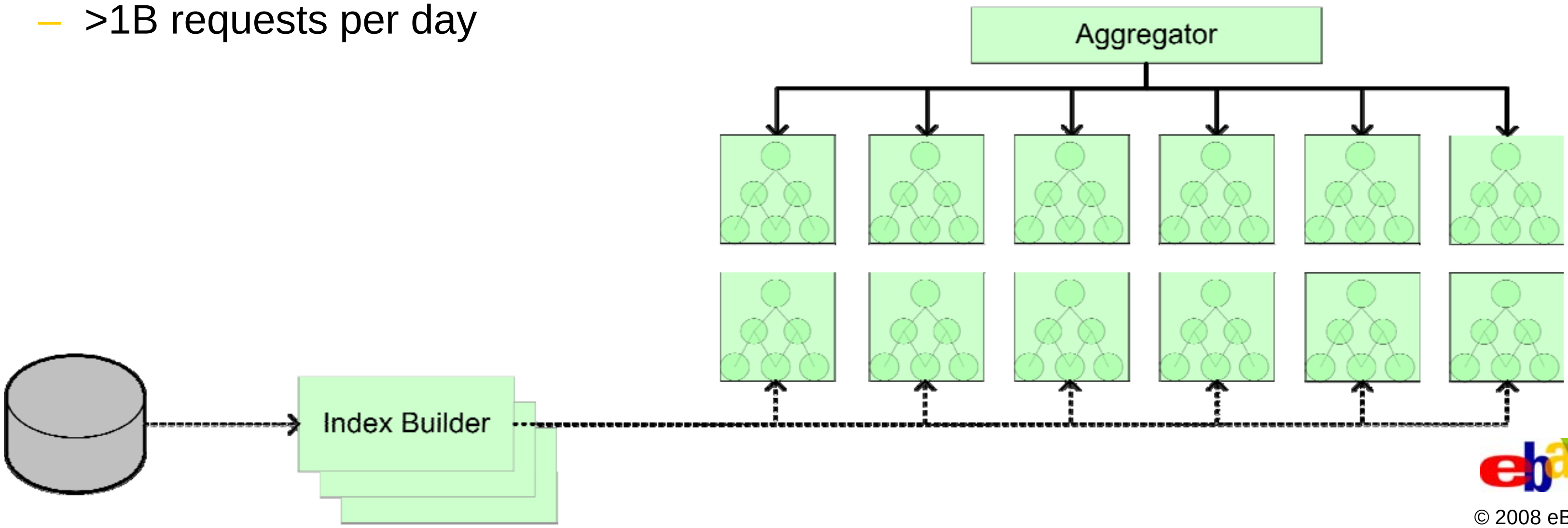
- Combine data from multiple input sources
  - User behavior
  - Attributes of user, item, etc.
  - Business metrics
- Aggregate and Analyze
  - Compute scores
  - Make predictions
- Build and refine model over time
  - Predictive variables, weights



eBay Feedback Loop: Deployment



- Build index offline from aggregate data
  - Data from multiple data sources
  - Updated periodically, typically daily or weekly
- Deploy index to online metric server
  - Fast in-memory hierarchical lookup structure for static data
  - Shared infrastructure for multiple types of static data
  - Partitioned horizontally by data
- Multiple systems query for real-time decision-making
  - >1B requests per day



Agenda

- Machine Learning and Feedback Loops
- Typical Feedback Loop at eBay
- Extended Example: eBay Search
- Concluding Thoughts

Search Challenges at eBay

- ***Goal: Match buyer's search query to best inventory on the site***

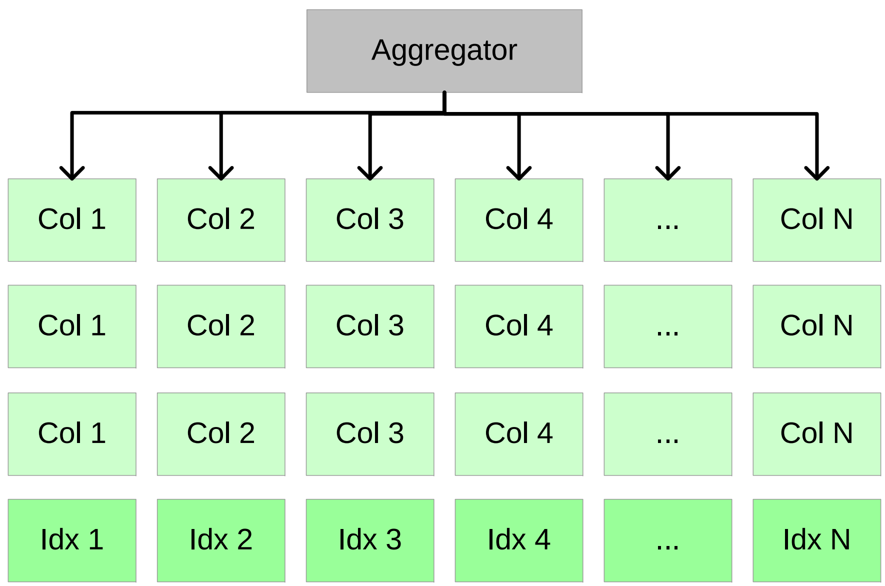
- Highly Dynamic Inventory
  - 10% of items turn over every day
- Real-time Marketplace
  - Every list / bid / purchase must be "immediately" reflected in search
- Comprehensive Results
  - Sellers and buyers expect results to contain all matching items
  - Results display and business logic require aggregate data ("histograms") across entire result set
- Navigation and Refinement
  - Need to be able to query and refine by both structured and unstructured data

eBay’s Search Infrastructure

- Search Grid
- Feeder Pipeline
- Query Augmentation and Recommendation
- Inventory Selection
- Adaptive User Experience

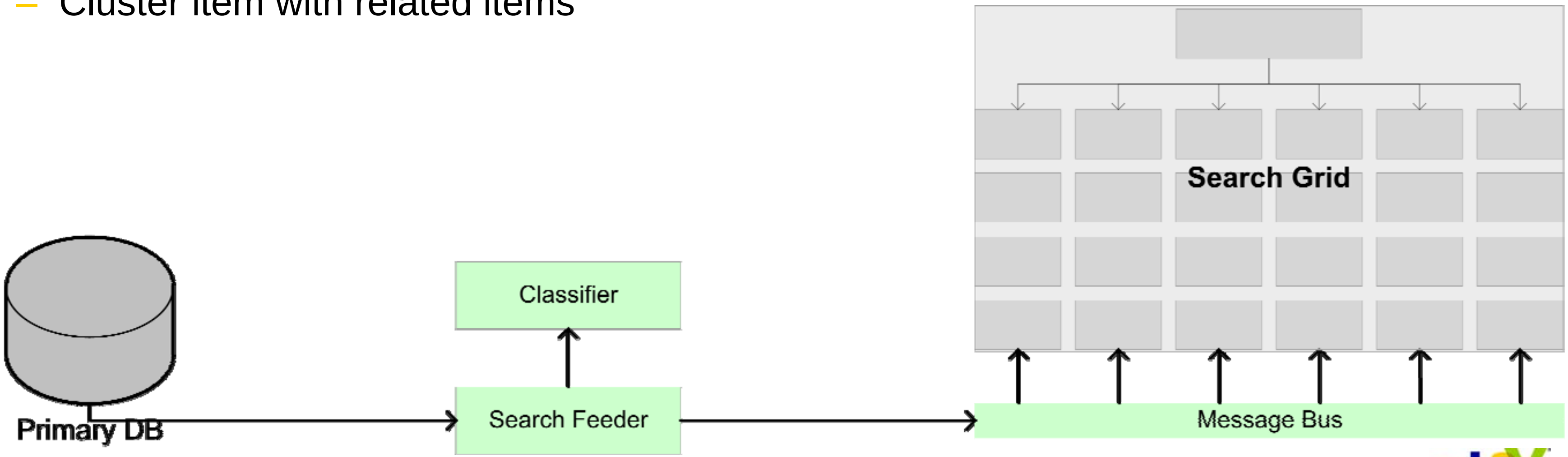
Search Grid

- In-memory search engine with real-time updates to the index
- Search index divided into grid of N shards (“columns”) by modulo of a key
- Each shard is replicated to M instances (“rows”)
- Aggregator parallelizes query to one node in each column, aggregates and post-processes results



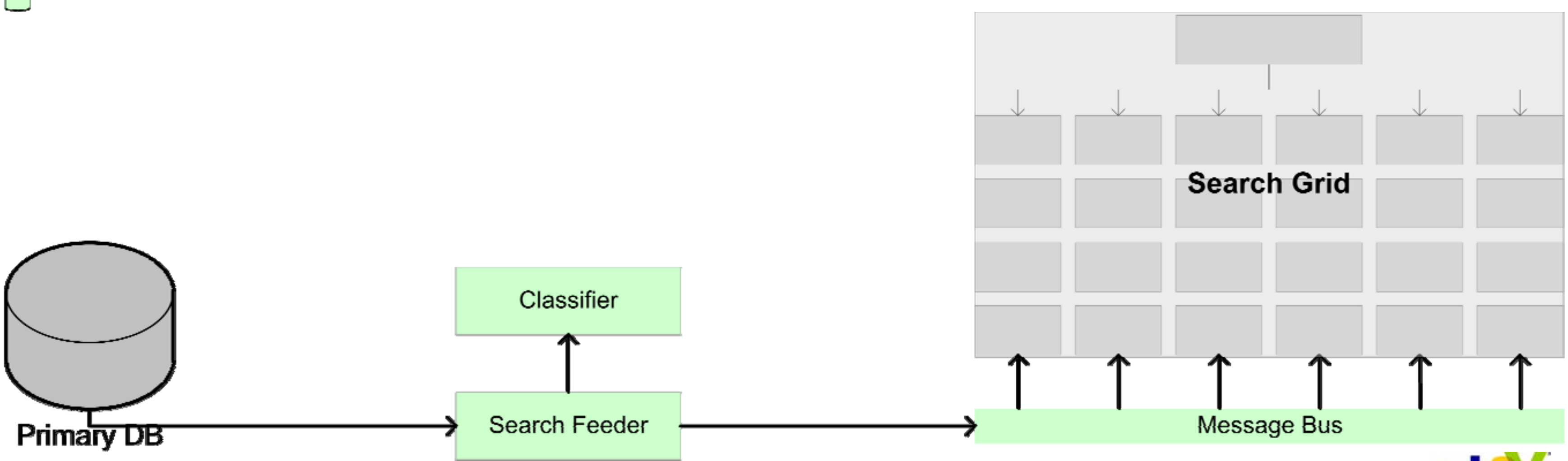
Feeder Pipeline

- Search Feeder
  - Read and transform item updates from primary database
  - Publish updates to search grid
- Classifier: improve recall and precision of search results
  - Stage in asynchronous feeder pipeline, driven by item update events
  - Augment seller-supplied metadata with additional attributes through inference
  - Extract structured concepts from text (size, color, etc.) for faceted navigation
  - Classify item to product in catalog
  - Cluster item with related items



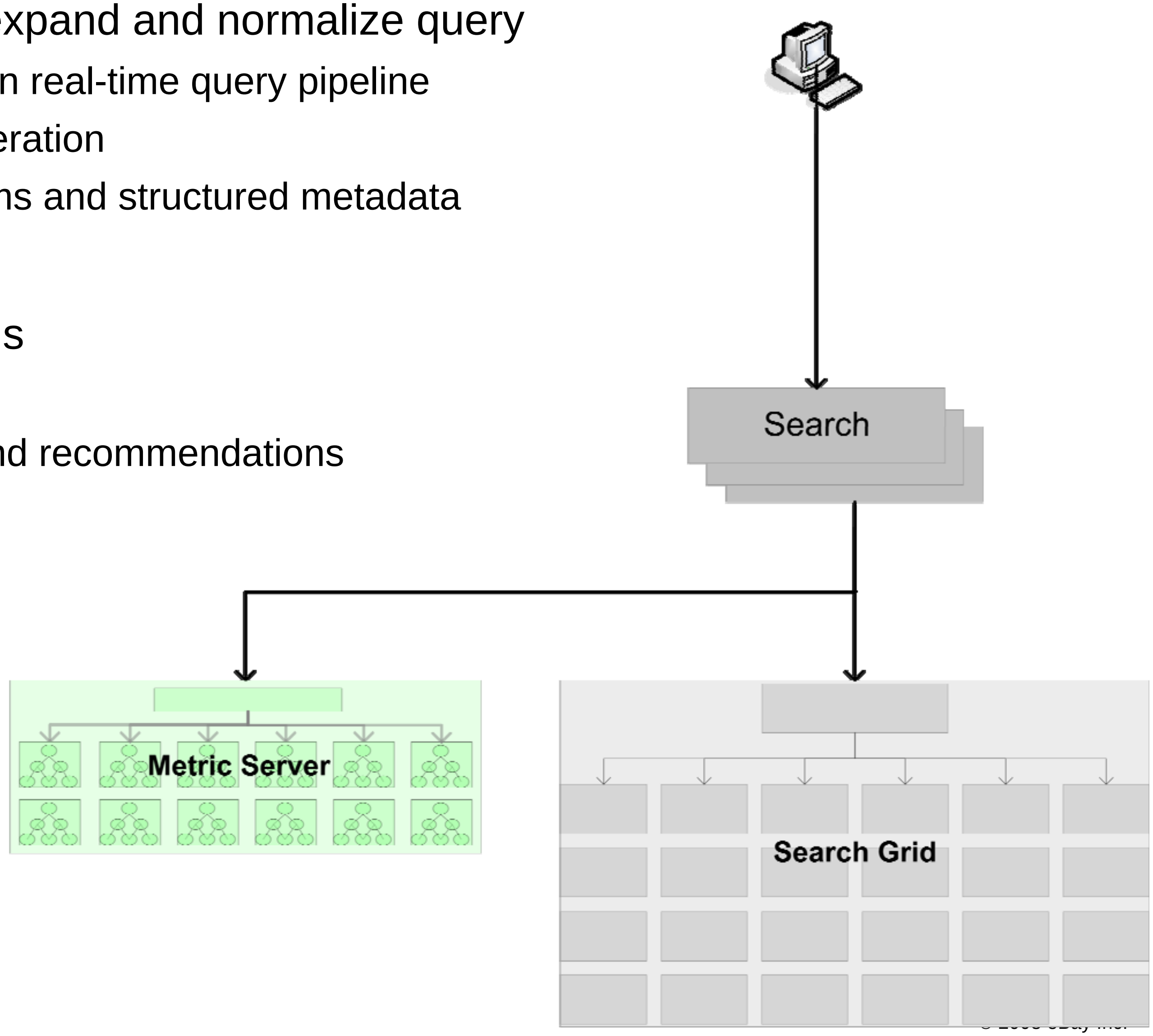
Feeder Pipeline: Feedback Loops

- Attribute extraction
  - Extract most appropriate attributes and concepts
- Product classification
  - Classify item to best product in catalog
- Item clustering
  - Cluster item with most related items

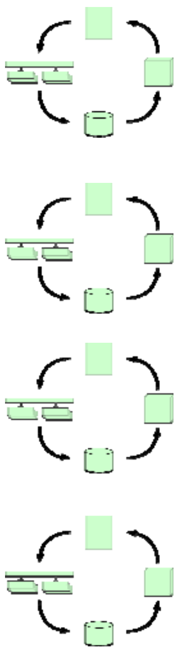


Query Augmentation and Recommendation

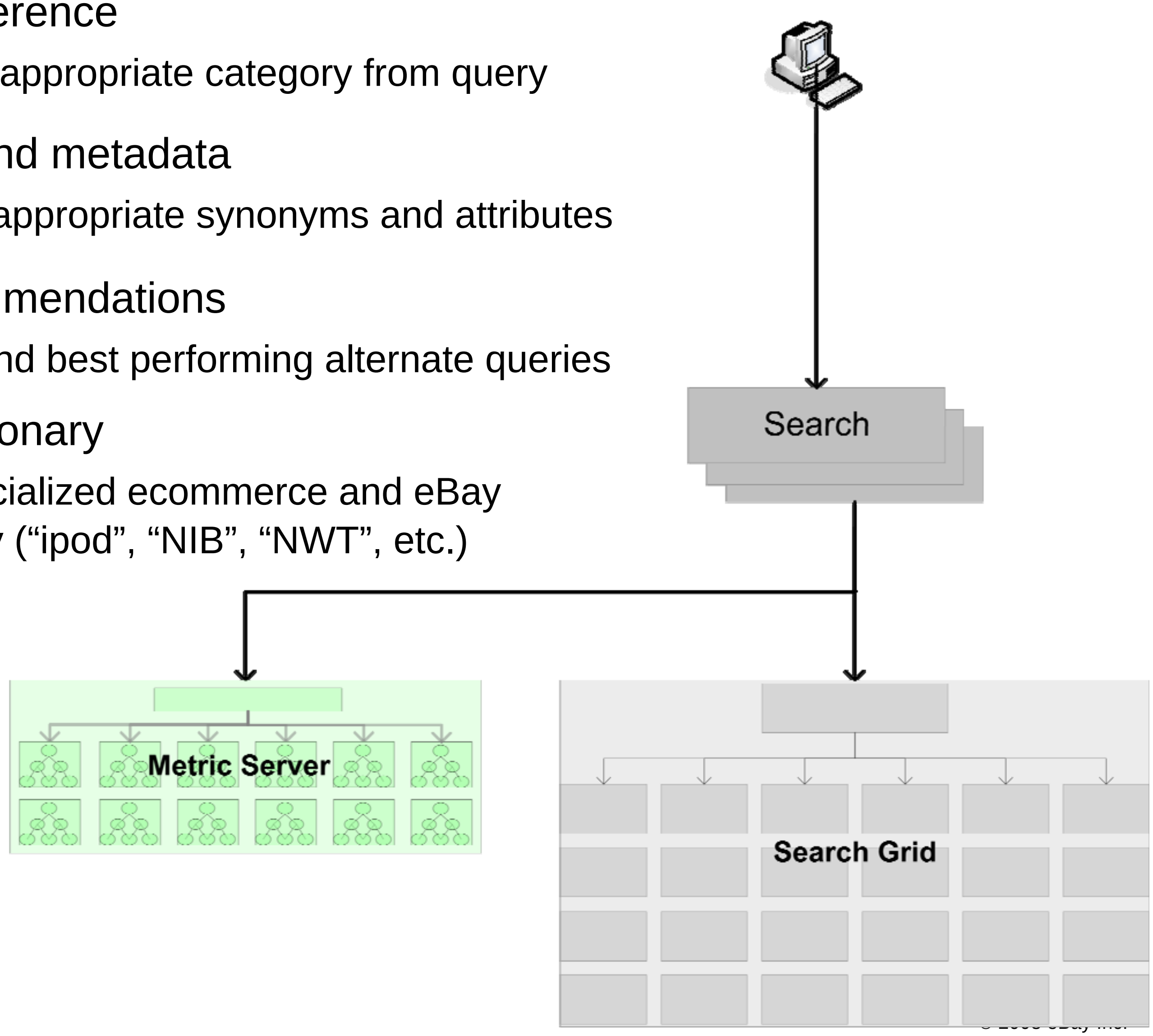
- Query augmentation: expand and normalize query
  - Pre-processing stage in real-time query pipeline
  - Stemming and transliteration
  - Augment with synonyms and structured metadata
  - Category inference
- Query recommendations
  - Spelling corrections
  - Search suggestions and recommendations



Query Augmentation: Feedback Loops



- Category inference
  - Infer most appropriate category from query
- Synonyms and metadata
  - Add most appropriate synonyms and attributes
- Query recommendations
  - Recommend best performing alternate queries
- Spelling dictionary
  - Learn specialized ecommerce and eBay vocabulary (“ipod”, “NIB”, “NWT”, etc.)



Inventory Selection

- Determine “best” set of item results (“Best Match”)
  - Order results by weighted combination of relevance, trust, price, and other factors
  - Relevance of an item is specific to a particular user and query
- Scoring
  - Calculate static item factors asynchronously on item update events
  - Calculate query factors asynchronously from user behavior
  - Item score for a query is a combination of item, query, and seller factors

Score =

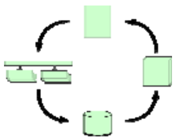
$$\omega_{i1..in}\{\text{Query Factors}\} + \omega_{j1..jn}\{\text{Item Factors}\} + \omega_{k1...kn}\{\text{Seller Factors}\}$$

Query Factors reward relevance and item attributes positively correlated with buying activity

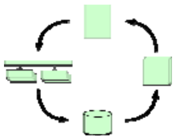
Item factors can include price, shipping cost, terms of sale, format, etc.

Seller factors can include buyer satisfaction metrics, trust, etc.

Inventory Selection: Feedback Loops



- Factors
  - Learn query, item, and seller factors which best predict good buyer experience



- Item Attributes
  - Learn most relevant item attributes for a particular user and query

Score =
$\omega_{i1..in}$ {Query Factors} + $\omega_{j1..jn}$ {Item Factors} + $\omega_{k1...kn}$ {Seller Factors}

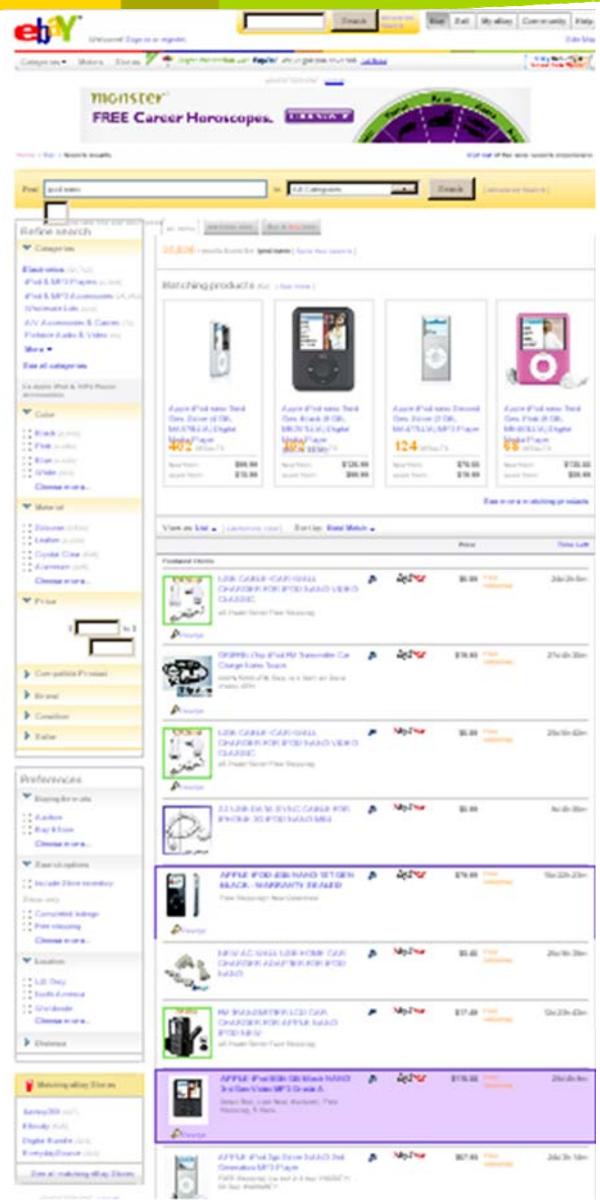
Query Factors reward relevance and item attributes positively correlated with buying activity

Item factors can include price, shipping cost, terms of sale, format, etc.

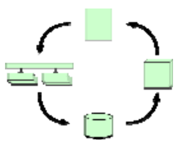
Seller factors can include buyer satisfaction metrics, trust, etc.

Adaptive Finding Experience

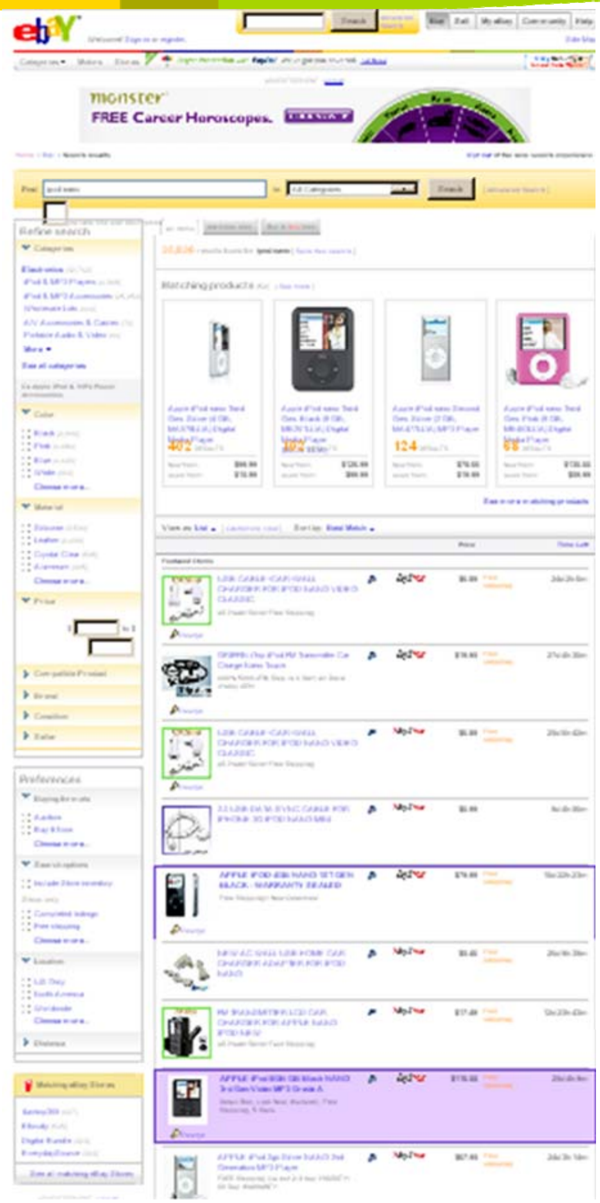
- Choose page, modules, and inventory which provide best experience for that user and query
- Display most useful facets for refinement
- Users “vote” with their activity and purchases



Adaptive Finding Experience



- Learn which options perform best for a particular user and query
  - Page
  - Modules
  - Inventory
  - Facets
- Actively experiment with alternatives for continual learning

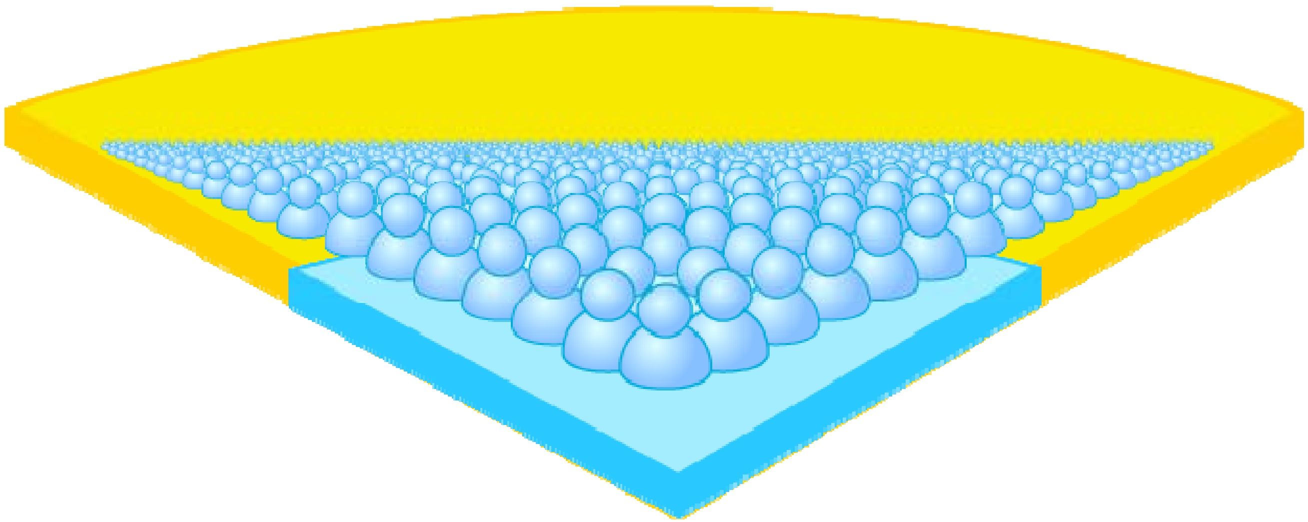


Agenda

- Machine Learning and Feedback Loops
- Typical Feedback Loop at eBay
- Extended Example: eBay Search
- Concluding Thoughts

Machine Learning Can Empower the Community

At the end of the day, the eBay community is the most important arbiter of what works on the site



An automated machine-learning system is the most complete and accurate way to reflect the diversity and richness of the eBay community

Questions?

- Randy Shoup, eBay Distinguished Architect

rshoup@ebay.com