

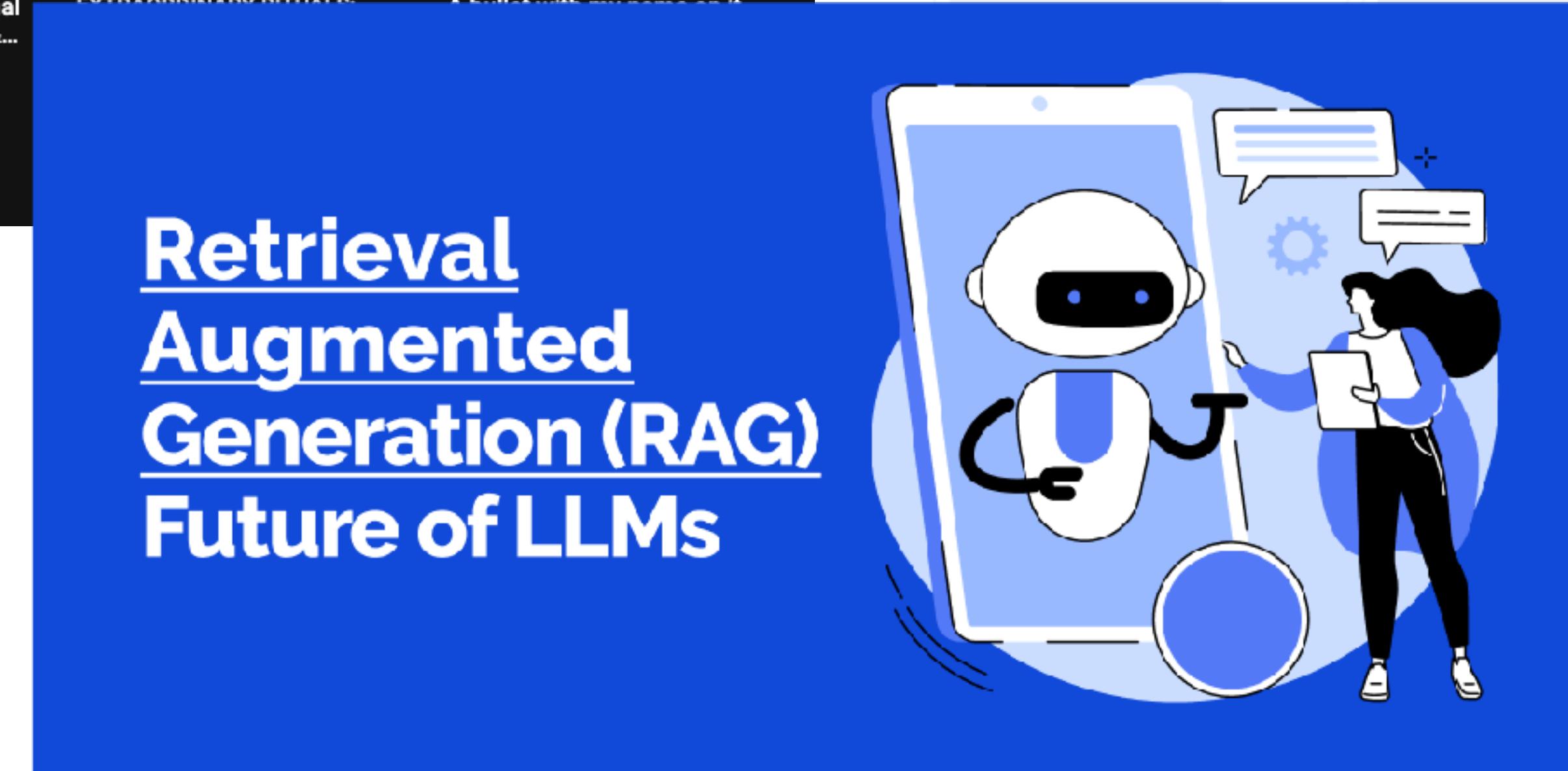
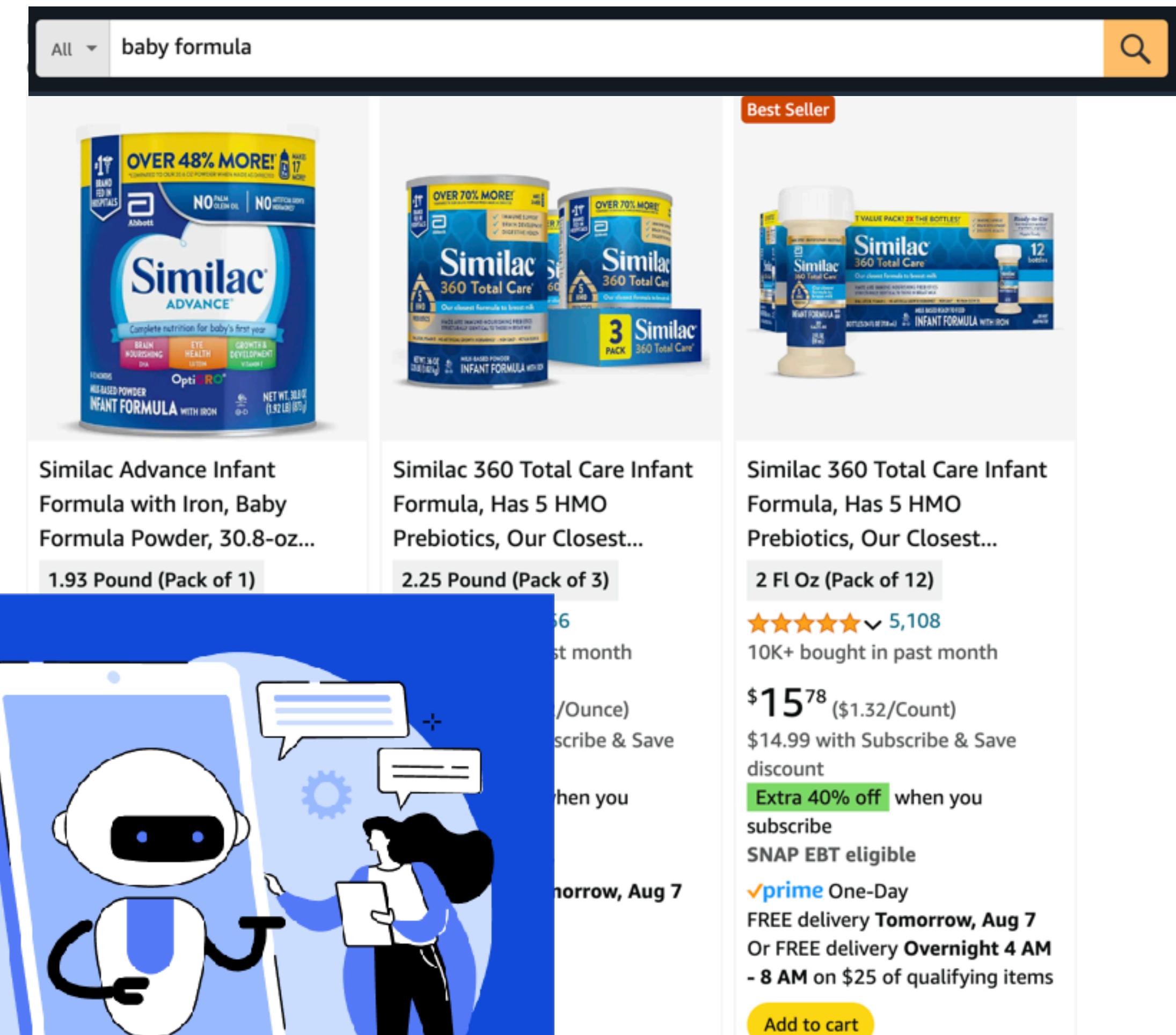
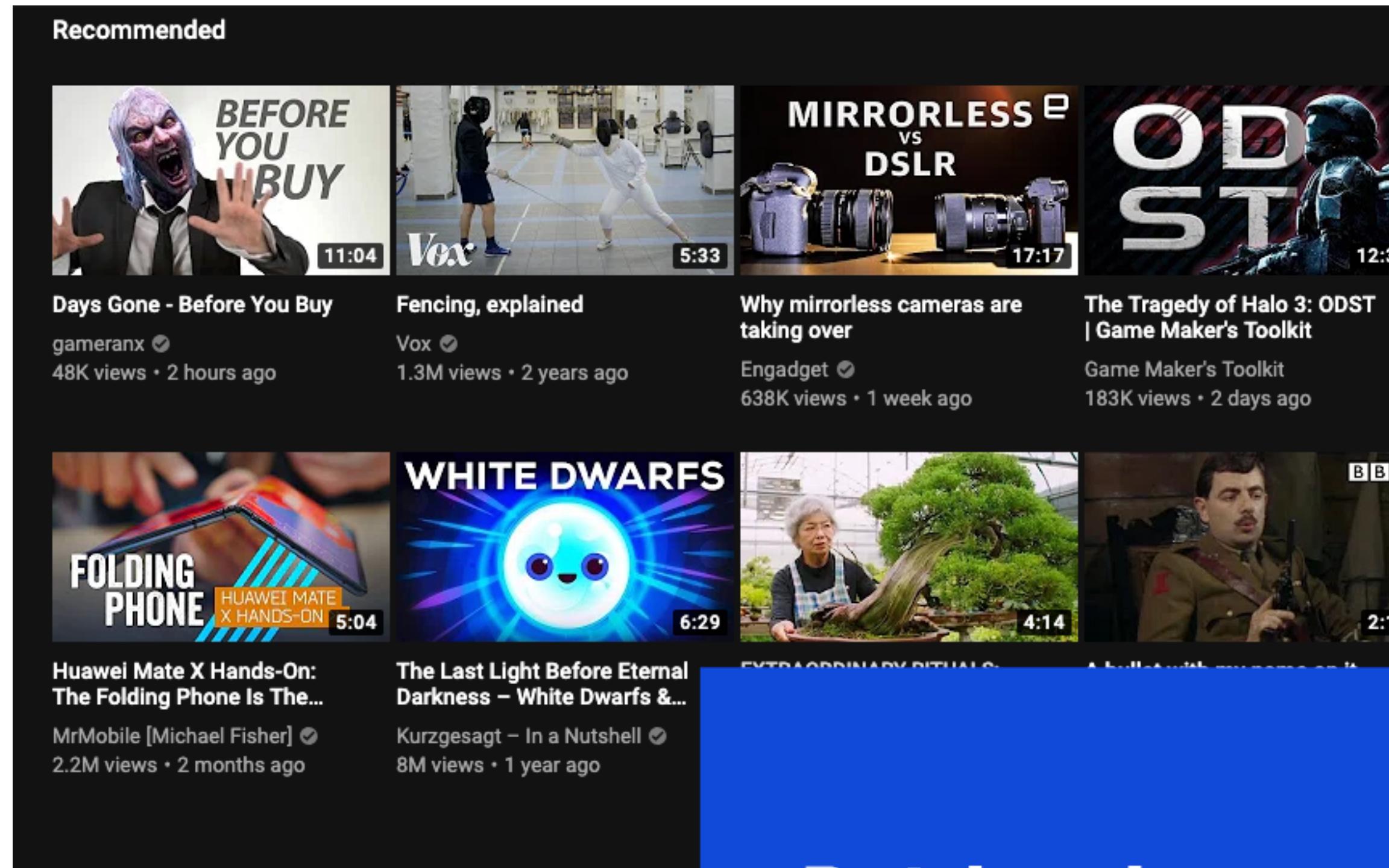
# Orbit: A Framework for Designing and Evaluating Multi-objective Rankers



**Chenyang Yang, Tesi Xiao, Michael Shavlovsky, Christian Kästner, Sherry Tongshuang Wu**



# Ranking models are prevalent in AI applications



# Ranking models need to consider multiple objectives

All  baby formula

Product	Price	Rating	Purchase Count	Offer
Similac Advance Infant Formula with Iron, Baby Formula Powder, 30.8-oz.	\$35.98 (\$1.17/Ounce)	4.5	1,917	Extra 40% off when you subscribe
Similac 360 Total Care Infant Formula, Has 5 HMO Prebiotics, Our Closest...	\$175.25 (\$1.62/Ounce)	4.5	356	Extra 40% off when you subscribe
Similac 360 Total Care Infant Formula, Has 5 HMO Prebiotics, Our Closest...	\$15.78 (\$1.32/Count)	4.5	5,108	SNAP EBT eligible



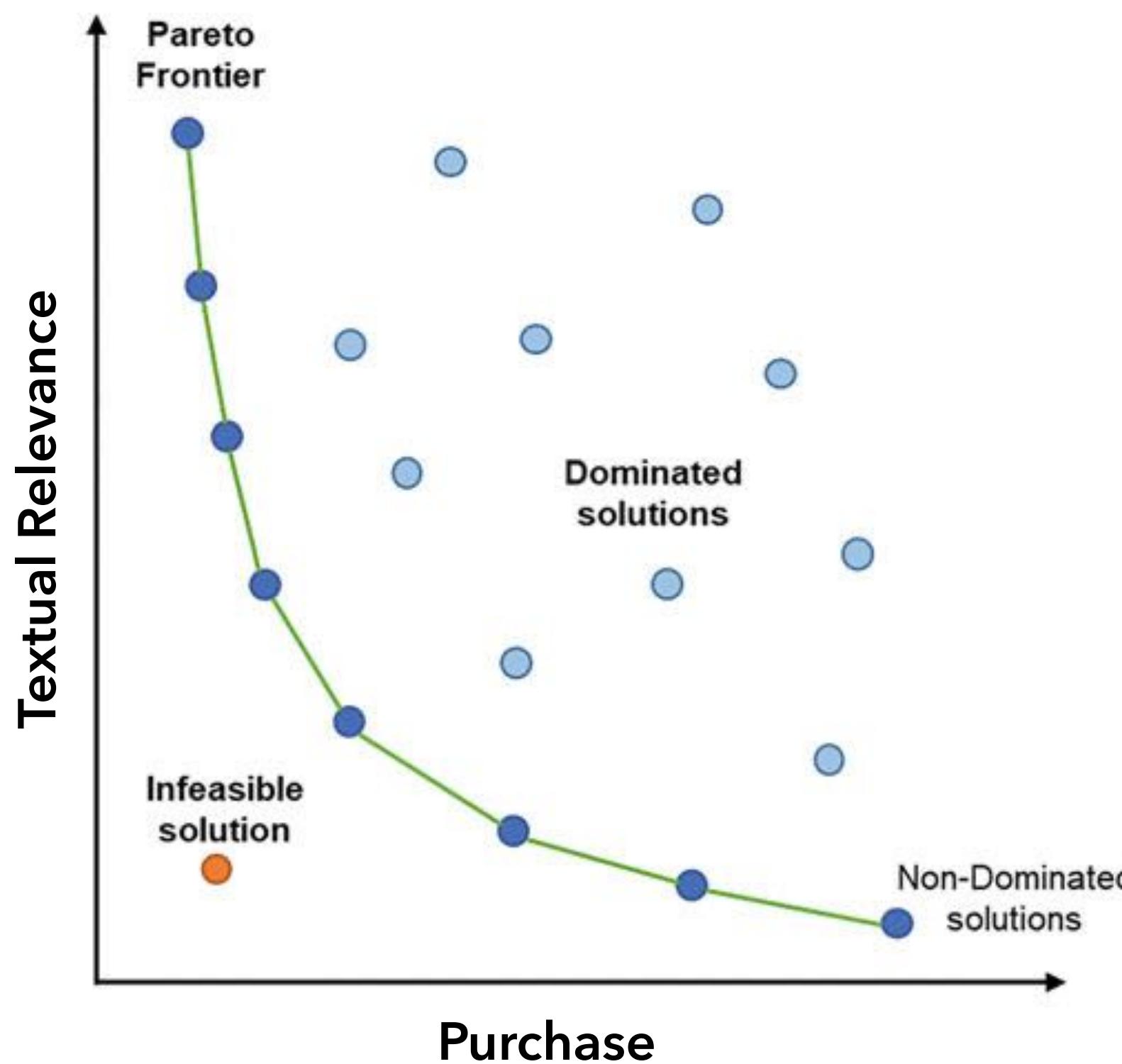
How do we train a model optimized for all these objectives 🤔?

# The Reality of Multi-objective Optimization

For multi-objective optimization, there is **no single “best” solution**.

Every solution needs to consider **trade-offs**.

With changing environment, there are **constantly new dimensions** (and trade-offs) to consider.



Multi-objective ranking is a **wicked problem!**

*“a wicked problem is a problem that is difficult or impossible to solve because of incomplete, contradictory, and changing requirements that are often difficult to recognize.”*

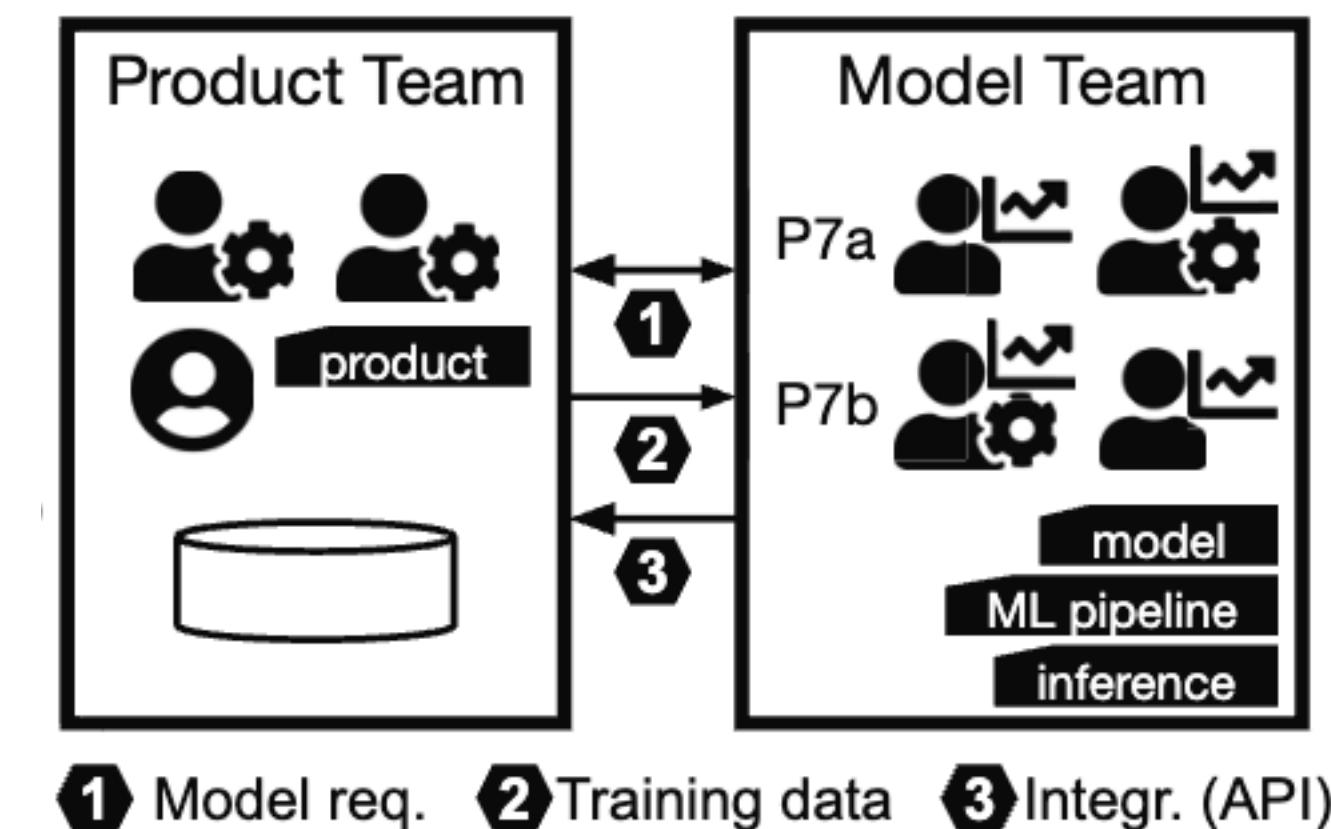
# Observation #1: Collaboration Challenges

Multi-objective ranking is a wicked problem: **Need to iterate and make trade-offs.**

**Product team** and **model team** need to **collaborate** to iterate on product rankers, but they speak different languages!

Strong domain knowledge, concrete observations, understand customer experience and needs.

Think about **examples, observations, user expectations.**



Strong experimentation skills, understand data, model, objectives, and metrics

Think about **(multi-)objectives, metrics, guardrails.**

**Struggle to provide actionable feedback & incorporate feedback!**

# Observation #1: Collaboration Challenges

With changing environment, there are **constantly new dimensions** (and trade-offs) to consider.  
**Model team and product team** need to collaborate to iterate on product rankers, but they speak different languages!

How to provide a shared language for practitioners to  
**communicate** 💬 and **collaborate** 🤝?

➊ Model req. ➋ Training data ➌ Integr. (API)

*Struggle to provide actionable feedback & incorporate feedback!*

# Observation #2: Design & Evaluation Challenges



**Huge design space to iterate over**

**Lots of evaluation information to track**

Many different potential objectives, and different ways to combine them into the model.

- Aggregated metrics for overall trend
- Individual examples for concrete customer experience
- Data slices for more detailed analysis & refined insights

# Observation #2: Design & Evaluation Challenges

The diagram shows a sequence of interactions starting from a search results page for "baby formula". The first interaction is a "Click" represented by a cursor icon over a product thumbnail. This leads to a second "Click" over a shopping cart icon, followed by a "Purchase" step. A large arrow points from the initial search results to the final purchase stage, labeled "Relevance".

How to support practitioners efficiently **explore** 🔎 the design space and **evaluate** 📈 their explorations?

Huge ways to combine them into the model.

*Lots of evaluation information to track*

- Aggregated metrics for overall trend
- Individual examples for concrete customer experience
- Data slices for more detailed analysis & refined insights

# Orbit : A Framework for Designing and Evaluating Multi-objective Rankers

# Key idea: Objectives-centric Design & Evaluation

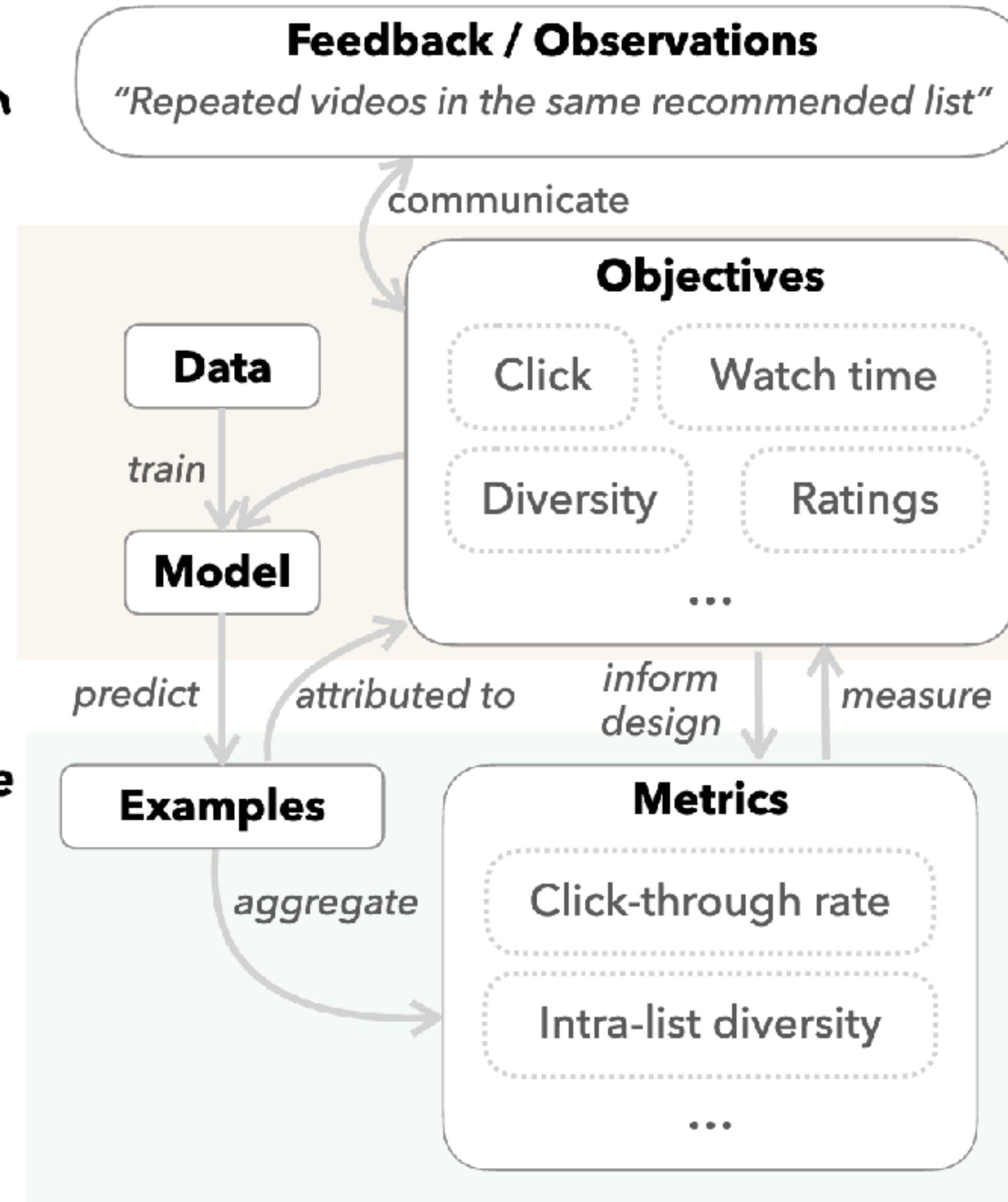


**Feedback / Observations**  
*"Repeated videos in the same recommended list"*



**Design**

**Evaluate**



**Objectives** should be surfaced as the main object for stakeholders to **navigate** through the design space, **communicate** their findings, and **negotiate** over trade-offs.

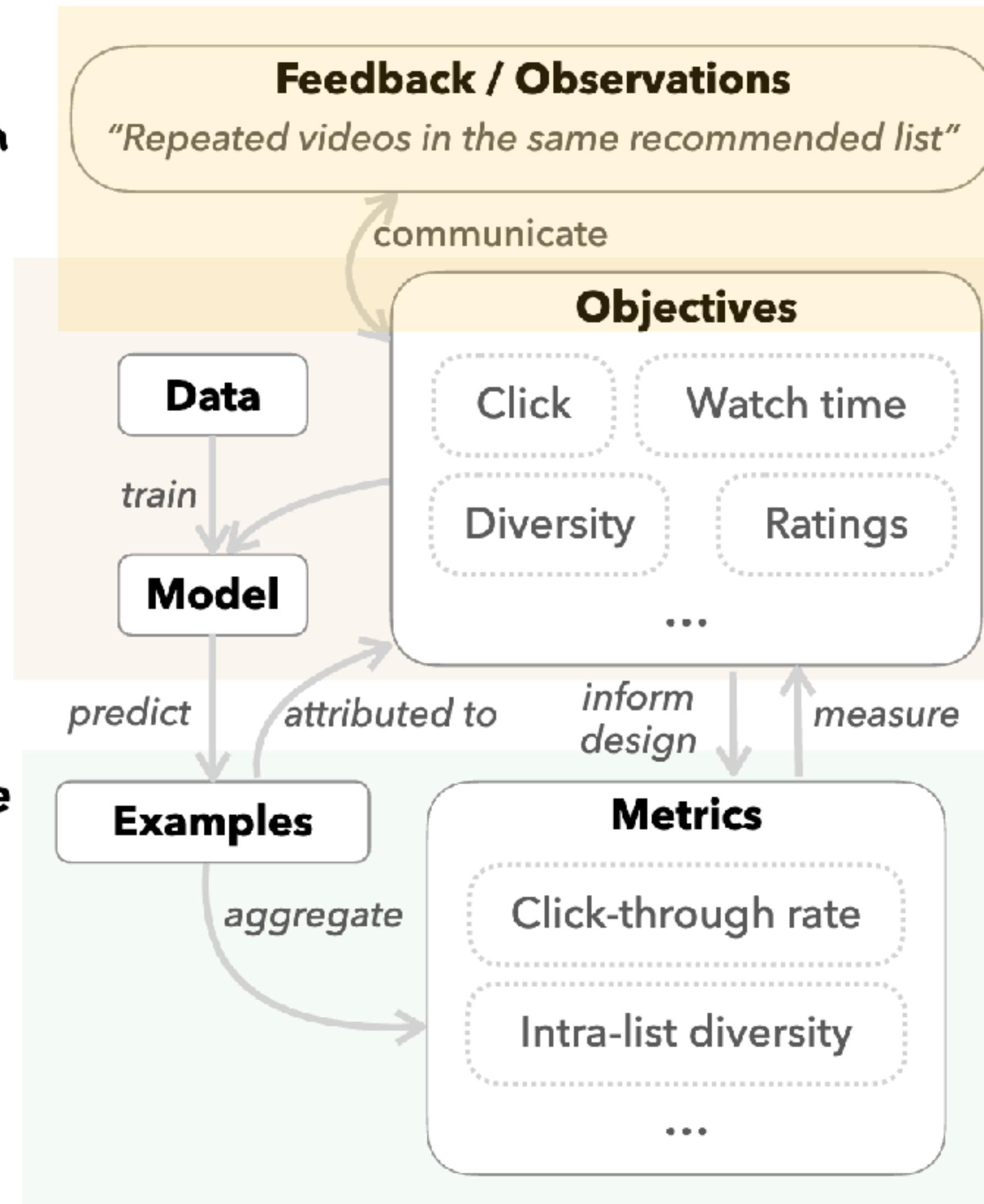
This helps provides a **shared language** and **appropriate guidance** for practitioners.

# Key idea: Objectives-centric Design & Evaluation



**Design**

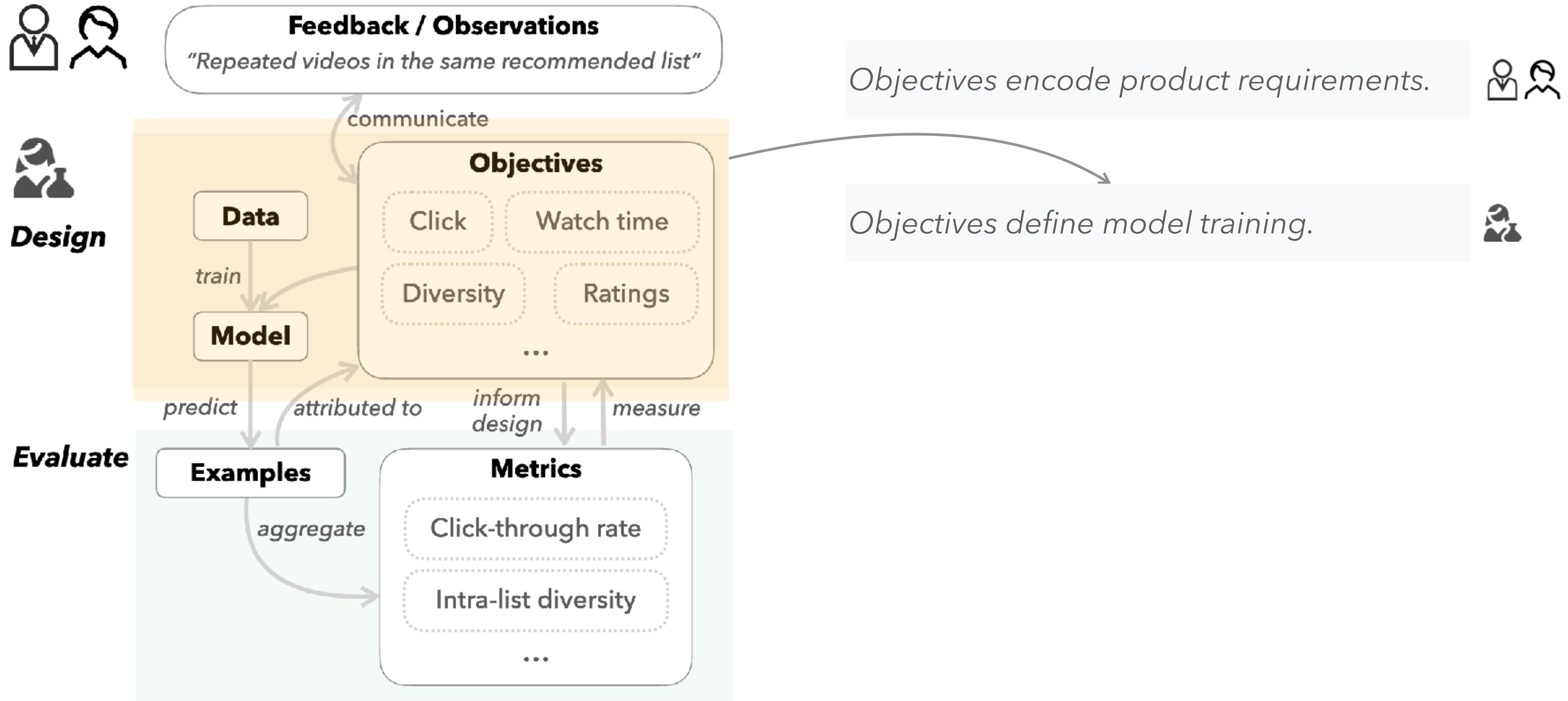
**Evaluate**



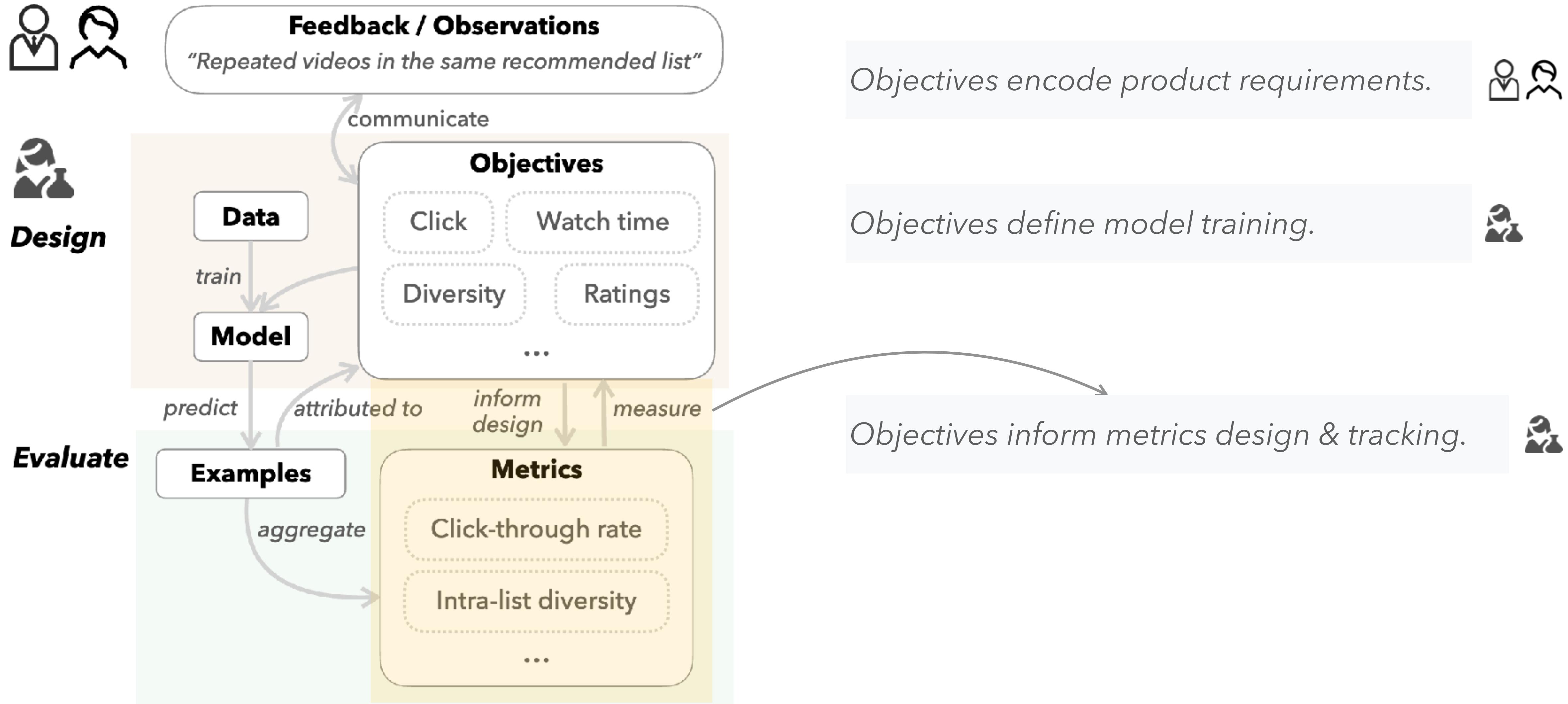
Objectives encode product requirements.



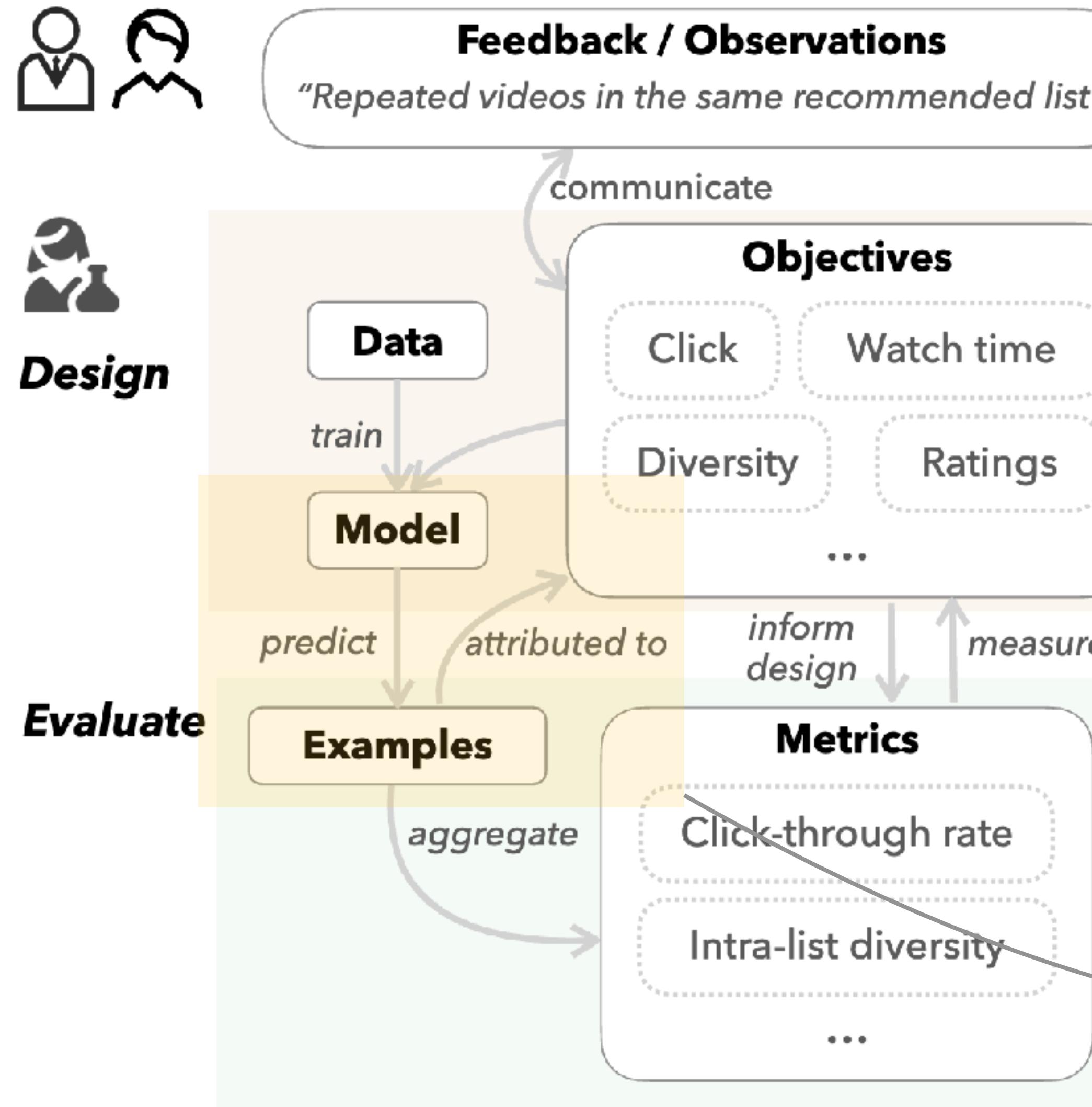
# Key idea: Objectives-centric Design & Evaluation



# Key idea: Objectives-centric Design & Evaluation



# Key idea: Objectives-centric Design & Evaluation



Objectives encode product requirements.



Objectives define model training.



Objectives inform metrics design & tracking.



Objectives explain concrete observations.



# Objectives can be used as boundary objects

**Better communication**   
**and collaboration** 

*"a boundary object is information, such as specimens, field notes, and maps, used in different ways by different communities for collaborative work through scales"*

Objectives encode product requirements.



Objectives define model training.



Objectives inform metrics design & tracking.

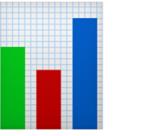


Objectives explain concrete observations.



- Leigh Star, Susan (2010-09-01). "This is Not a Boundary Object: Reflections on the Origin of a Concept". *Science, Technology, & Human Values*. 35 (5).
- PAIR Symposium 2020, Ed Chi: A twist on loss functions as boundary objects

# Objectives can help navigate design space and forage evaluation information

**More efficient exploration**  and  
**more thorough evaluation** 

*Objectives encode product requirements.*



*Objectives define model training.*



- What designs have I explored?
- What should I explore next?
- What metric(s) should I monitor?
- Why is this model design problematic?

*Objectives inform metrics design & tracking.*



*Objectives explain concrete observations.*



Orbit 

# System Walkthrough

## Models

T1 x T2 x

Model A Model B

T1 T2

Metric

ndcg\_click\_prob

Slices ⓘ

All instances	1.0000 (15,559)	0.9794 (15,559)
30 quart coolers	1.0000 (15)	0.9912 (15)
numerical	1.0000 (3,194)	0.9766 (3,194)

Metrics ⓘ

ndcg_click_prob	1.0000 (15,559)	0.9794 (15,559)	-2.06%
highly Rated Density	0.6530 (15,559)	0.7721 (15,559)	+18.24%
ndcg_purchase_prob	0.9921 (15,559)	0.9308 (15,559)	-6.19%
exact_density	0.4953 (15,559)	0.6474 (15,559)	+30.70%

Metadata ⓘ

keywords

NOT Search Aa ab \* SET

match '2t baby girl toys' x

keywords (match '2t baby girl toys') x clear all

click purchase highly rated click relevant click

T1 T2

Combined objectives: click \* 1 +

Combined objectives: click \* 3 +

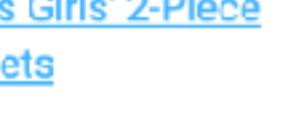
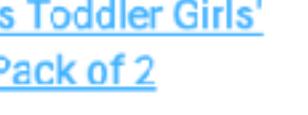
relevant click \* 1 +

highly rated click \* 1 +

Type a new model name here. '⌘ + ↵' to submit.

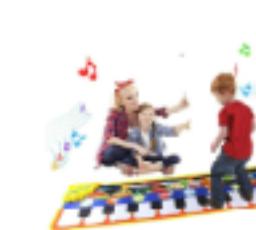
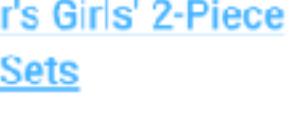
2t baby girl toys

RANKING:

1.  → ↑2 
2.  → ↑2 
3.  → ↑2 
4.  → ↑2 

2t baby girl toys

RANKING:

1.  ↑8 → 
2.  ↑11 → 
3.  ↓2 → 
4.  → 

Query Groups Per Page 1 1 - 1 of 1

## Models

T1 x T2 x

Model A Model B

T1 T2

Metric

ndcg\_click\_prob

Slices ⓘ

All instances	1.0000 (15,559)	0.9794 (15,559)
30 quart coolers	1.0000 (15)	0.9912 (15)
numerical	1.0000 (3,194)	0.9766 (3,194)

Metrics ⓘ

ndcg_click_prob	1.0000 (15,559)	0.9794 (15,559)	-2.06%
highly Rated Density	0.6530 (15,559)	0.7721 (15,559)	+18.24%
ndcg_purchase_prob	0.9921 (15,559)	0.9308 (15,559)	-6.19%
exact_density	0.4953 (15,559)	0.6474 (15,559)	+30.70%

Metadata ⓘ

keywords

NOT Search Aa ab \* SET

match '2t baby girl toys' x

keywords (match '2t baby girl toys') x clear all

click : purchase : highly rated click : relevant click :

# 1. Objectives Overview

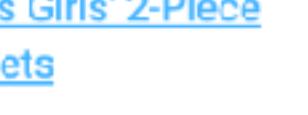
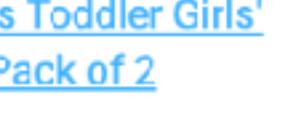
T1 Combined objectives: click \* 1

T2 Combined objectives: click \* 3 + relevant click \* 1 highly rated click \* 1

Type a new model name here. '⌘ + ↵' to submit.

2t baby girl toys

RANKING:

1.  → ↑2 
2.  → ↑2 
3.  → ↑2 
4.  → ↑2 

2t baby girl toys

RANKING:

1.  ↑8 → 
2.  ↑11 → 
3.  ↓2 → 
4.  → 

Query Groups Per Page 1 1 - 1 of 1

## Models

T1 x T2 x

Model A Model B

T1 T2

Metric

ndcg\_click\_prob

Slices ⓘ

All instances	1.0000 (15,559)	0.9794 (15,559)
30 quart coolers	1.0000 (15)	0.9912 (15)
numerical	1.0000 (3,194)	0.9766 (3,194)

Metrics ⓘ

ndcg_click_prob	1.0000 (15,559)	0.9794 (15,559)	-2.06%
highly Rated Density	0.6530 (15,559)	0.7721 (15,559)	+18.24%
ndcg_purchase_prob	0.9921 (15,559)	0.9308 (15,559)	-6.19%
exact_density	0.4953 (15,559)	0.6474 (15,559)	+30.70%

Metadata ⓘ

keywords

NOT Search Aa ab \* SET

match '2t baby girl toys' x

keywords (match '2t baby girl toys') x clear all

click : purchase : highly rated click : relevant click :

Drag the objective to apply it.

click\_probability

click \* 1

T2

Combined objectives: click \* 3 + relevant\_click \* 1 + highly\_rated\_click \* 1

Type a new model name here. '⌘ + ↵' to submit.

# 1. Objectives Overview

2t baby girl toys

RANKING:

1. → ↑2 [Disney Toddler Boys' Toy Story Short Sleeve T-Shirt](#)
2. → ↑2 [Simple Joys by Carter's Girls' 2-Piece Assorted Rashguard Sets](#)
3. → ↑2 [Simple Joys by Carter's Toddler Girls' Pull-On Fleece Pants, Pack of 2](#)
4. → ↑2 [Simple Joys by Carter's Girls' 2-Piece Assorted Rashguard Sets](#)

2t baby girl toys

RANKING:

1. ↑8 → [ToyVelt Toy Piano for Toddler Girls - Cute Piano for Kids with Built-in Microphone & Music Modes - Best Birthday Gifts for 3 4 5 Year Old Girls - Educational Keyboard Musical Instrument Toys](#)
2. ↑11 → [Toys for 1-6 Year Old Girls Boys Toddlers Infant Kids, Gifts for 1-3 Year Old Boys Girls Piano Music Dance Mat with 19 Keys Piano Mat, 8 Musical Instruments Build-in Speaker & Recording Function](#)
3. ↓2 → [Disney Toddler Boys' Toy Story Short Sleeve T-Shirt](#)
4. → ↑2 [Simple Joys by Carter's Girls' 2-Piece Assorted Rashguard Sets](#)

Query Groups Per Page 1 1 - 1 of 1

## Models

T1 x T2 x

Model A Model B

T1 T2

Metric ndcg\_click\_prob

Slices ⓘ

All instances	1.0000 (15,559)	0.9794 (15,559)
30 quart coolers	1.0000 (15)	0.9912 (15)
numerical	1.0000 (3,194)	0.9766 (3,194)

Metrics ⓘ

ndcg_click_prob	1.0000 (15,559)	0.9794 (15,559)	-2.06%
highly Rated Density	0.6530 (15,559)	0.7721 (15,559)	+18.24%
ndcg_purchase_prob	0.9921 (15,559)	0.9308 (15,559)	-6.19%
exact_density	0.4953 (15,559)	0.6474 (15,559)	+30.70%

Metadata ⓘ

keywords

NOT Search Aa ab \* SET

match '2t baby girl toys' x

keywords (match '2t baby girl toys') x clear all

click : purchase : highly rated click : relevant click :

# 1. Objectives Overview

T1 Combined objectives: click \* 1

T2 Combined objectives: click \* 3 + relevant click \* 1 + highly rated click \* 1

Type a new model name here. '⌘ + ↵' to submit.

2t baby girl toys

RANKING:

- Disney Toddler Boys' Toy Story Short Sleeve T-Shirt → ↑2
- Simple Joys by Carter's Girls' 2-Piece Assorted Rashguard Sets → ↑2
- Simple Joys by Carter's Toddler Girls' Pull-On Fleece Pants, Pack of 2 → ↑2
- Simple Joys by Carter's Girls' 2-Piece Assorted Rashguard Sets

2t baby girl toys

RANKING:

- ToyVelt Toy Piano for Toddler Girls - Cute Piano for Kids with Built-in Microphone & Music Modes - Best Birthday Gifts for 3 4 5 Year Old Girls - Educational Keyboard Musical Instrument Toys ↑8 →
- Toys for 1-6 Year Old Girls Boys Toddlers Infant Kids, Gifts for 1-3 Year Old Boys Girls Piano Music Dance Mat with 19 Keys Piano Mat, 8 Musical Instruments Build-in Speaker & Recording Function ↑11 →
- Disney Toddler Boys' Toy Story Short Sleeve T-Shirt ↓2 →
- Simple Joys by Carter's Girls' 2-Piece Assorted Rashguard Sets

Query Groups Per Page 1 - 1 of 1

## Models

T1 x T2 x

Model A Model B

T1 T2

Metric

ndcg\_click\_prob

Slices ⓘ

All instances	1.0000 (15,559)	0.9794 (15,559)
30 quart coolers	1.0000 (15)	0.9912 (15)
numerical	1.0000 (3,194)	0.9766 (3,194)

Metrics ⓘ

ndcg_click_prob	1.0000 (15,559)	0.9794 (15,559)	-2.06%
highly Rated Density	0.6530 (15,559)	0.7721 (15,559)	+18.24%
ndcg_purchase_prob	0.9921 (15,559)	0.9308 (15,559)	-6.19%
exact_density	0.4953 (15,559)	0.6474 (15,559)	+30.70%

Metadata ⓘ

keywords

NOT Search Aa ab \* SET

match '2t baby girl toys' x

keywords (match '2t baby girl toys') x clear all

click purchase highly rated click relevant click

T1 T2

Combined objectives: click \* 1 +

Combined objectives: click \* 3 +

relevant click \* 1 +

highly rated click \* 1 +

Type a new model name here. '⌘ + ↵' to submit.

2t baby girl toys

RANKING:

- Disney Toddler Boys' Toy Story Short Sleeve T-Shirt → ↑2
- Simple Joys by Carter's Girls' 2-Piece Assorted Rashguard Sets → ↑2
- Simple Joys by Carter's Toddler Girls' Pull-On Fleece Pants, Pack of 2 → ↑2
- Simple Joys by Carter's Girls' 2-Piece Assorted Rashguard Sets

2t baby girl toys

RANKING:

- ToyVelt Toy Piano for Toddler Girls - Cute Piano for Kids with Built-in Microphone & Music Modes - Best Birthday Gifts for 3 4 5 Year Old Girls - Educational Keyboard Musical Instrument Toys ↑8 →
- Toys for 1-6 Year Old Girls Boys Toddlers Infant Kids, Gifts for 1-3 Year Old Boys Girls Piano Music Dance Mat with 19 Keys Piano Mat, 8 Musical Instruments Build-in Speaker & Recording Function ↑11 →
- Disney Toddler Boys' Toy Story Short Sleeve T-Shirt ↓2 →
- Simple Joys by Carter's Girls' 2-Piece Assorted Rashguard Sets

Query Groups Per Page 1 1 - 1 of 1

## 2. Qualitative Evaluation

## Models

T1 x T2 x

Model A Model B

T1 T2

Metric

ndcg\_click\_prob

Slices ⓘ

All instances	1.0000 (15,559)	0.9794 (15,559)
30 quart coolers	1.0000 (15)	0.9912 (15)
numerical	1.0000 (3,194)	0.9766 (3,194)

Metrics ⓘ

ndcg_click_prob	1.0000 (15,559)	0.9794 (15,559)	-2.06%
highly Rated Density	0.6530 (15,559)	0.7721 (15,559)	+18.24%
ndcg_purchase_prob	0.9921 (15,559)	0.9308 (15,559)	-6.19%
exact_density	0.4953 (15,559)	0.6474 (15,559)	+30.70%

Metadata ⓘ

keywords

NOT Search Aa ab \* SET

match '2t baby girl toys' x

keywords (match '2t baby girl toys') x clear all

click purchase highly rated click relevant click

T1 Combined objectives: click \* 1

T2 Combined objectives: click \* 3 + relevant click \* 1 highly rated click \* 1

Type a new model name here. '⌘ + ↵' to submit.

2t baby girl toys

RANKING:

- Disney Toddler Boys' Toy Story Short Sleeve T-Shirt → ↑2
- Simple Joys by Carter's Girls' 2-Piece Assorted Rashguard Sets → ↑2
- Simple Joys by Carter's Toddler Girls' Pull-On Fleece Pants, Pack of 2 → ↑2
- Simple Joys by Carter's Girls' 2-Piece Assorted Rashguard Sets

2t baby girl toys

RANKING:

- ToyVelt Toy Piano for Toddler Girls - Cute Piano for Kids with Built-in Microphone & Music Modes - Best Birthday Gifts for 3 4 5 Year Old Girls - Educational Keyboard Musical Instrument Toys ↑8 →
- Toys for 1-6 Year Old Girls Boys Toddlers Infant Kids, Gifts for 1-3 Year Old Boys Girls Piano Music Dance Mat with 19 Keys Piano Mat, 8 Musical Instruments Build-in Speaker & Recording Function ↑11 →
- Disney Toddler Boys' Toy Story Short Sleeve T-Shirt ↓2 →
- Simple Joys by Carter's Girls' 2-Piece Assorted Rashguard Sets

Query Groups Per Page 1 1 - 1 of 1

## 2. Qualitative Evaluation

## Models

T1 X T2 X

Model A

Model B

T1

T2

Metric

ndcg\_click\_prob

Slices ⓘ



All instances

1.0000  
(15,559)

0.9794  
(15,559)

30 quart coolers

1.0000  
(15)

0.9912  
(15)

numerical

1.0000  
(3,194)

0.9766  
(3,194)

Metrics ⓘ



ndcg\_click\_prob  
-2.06%

1.0000  
(15,559)

0.9794  
(15,559)

highly Rated Density  
+18.24%

0.6530  
(15,559)

0.7721  
(15,559)

ndcg\_purchase\_prob  
-6.19%

0.9921  
(15,559)

0.9308  
(15,559)

exact\_density  
+30.70%

0.4953  
(15,559)

0.6474  
(15,559)

Metadata ⓘ

### 3. Quantitative Evaluation

keywords

match '2t baby girl toys'

NOT

Search

Aa ab ab \*

SET

keywords ( match '2t baby girl toys' ) X

clear all



click :

purchase :

highly\_rated\_click :

relevant\_click :



T1

Combined objectives: click \* 1

2t baby girl toys



→ ↑2



[Disney Toddler Boys' Toy Story Short Sleeve T-Shirt](#)

2.



→ ↑2



[Simple Joys by Carter's Girls' 2-Piece Assorted Rashguard Sets](#)

3.



→ ↑2



[Simple Joys by Carter's Toddler Girls' Pull-On Fleece Pants, Pack of 2](#)

4.



[Simple Joys by Carter's Girls' 2-Piece Assorted Rashguard Sets](#)

T2

Combined objectives: click \* 3

relevant\_click \* 1

highly\_rated\_click \* 1

2t baby girl toys

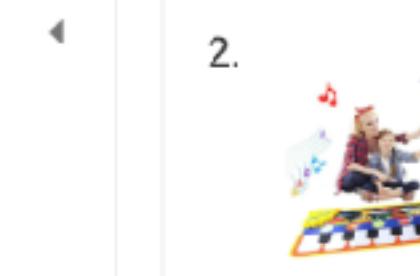


↑8 →



[ToyVelt Toy Piano for Toddler Girls - Cute Piano for Kids with Built-in Microphone & Music Modes - Best Birthday Gifts for 3 4 5 Year Old Girls - Educational Keyboard Musical Instrument Toys](#)

2.



↑11 →

[Toys for 1-6 Year Old Girls Boys Toddlers Infant Kids, Gifts for 1-3 Year Old Boys Girls Piano Music Dance Mat with 19 Keys Piano Mat, 8 Musical Instruments Build-in Speaker & Recording Function](#)

3.



↓2 →

[Disney Toddler Boys' Toy Story Short Sleeve T-Shirt](#)

4.



[Simple Joys by Carter's Girls' 2-Piece Assorted Rashguard Sets](#)

Type a new model name here. '⌘ + ↵' to submit.

Query Groups Per Page

1

1 - 1 of 1



## Models

T1 x T2 x

Model A Model B

T1 T2

Metric ndcg\_click\_prob

Slices ⓘ

All instances	1.0000 (15,559)	0.9794 (15,559)
30 quart coolers	1.0000 (15)	0.9912 (15)
numerical	1.0000 (3,194)	0.9766 (3,194)

Metrics ⓘ

ndcg_click_prob	1.0000 (15,559)	0.9794 (15,559)	-2.06%
highly Rated Density	0.6530 (15,559)	0.7721 (15,559)	+18.24%
ndcg_purchase_prob	0.9921 (15,559)	0.9308 (15,559)	-6.19%
exact_density	0.4953 (15,559)	0.6474 (15,559)	+30.70%

Metadata ⓘ

keywords

NOT Search Aa ab \* SET

match '2t baby girl toys' x

keywords (match '2t baby girl toys') x clear all

click purchase highly rated click relevant click

T1 T2

Combined objectives: click \* 1 +

Combined objectives: click \* 3 +

relevant click \* 1 +

highly rated click \* 1 +

Type a new model name here. '⌘ + ↵' to submit.

## 4. Fine-grained Analysis

**2t baby girl toys**

RANKING:

- Disney Toddler Boys' Toy Story Short Sleeve T-Shirt → ↑2
- Simple Joys by Carter's Girls' 2-Piece Assorted Rashguard Sets → ↑2
- Simple Joys by Carter's Toddler Girls' Pull-On Fleece Pants, Pack of 2 → ↑2
- Simple Joys by Carter's Girls' 2-Piece Assorted Rashguard Sets

**2t baby girl toys**

RANKING:

- ToyVelt Toy Piano for Toddler Girls - Cute Piano for Kids with Built-in Microphone & Music Modes - Best Birthday Gifts for 3 4 5 Year Old Girls - Educational Keyboard Musical Instrument Toys ↑8 →
- Toys for 1-6 Year Old Girls Boys Toddlers Infant Kids, Gifts for 1-3 Year Old Boys Girls Piano Music Dance Mat with 19 Keys Piano Mat, 8 Musical Instruments Build-in Speaker & Recording Function ↑11 →
- Disney Toddler Boys' Toy Story Short Sleeve T-Shirt ↓2 →
- Simple Joys by Carter's Girls' 2-Piece Assorted Rashguard Sets

Query Groups Per Page 1 1 - 1 of 1

Models

T1 x T2 x

Model A Model B

T1 T2 Metric

ndcg\_click\_prob

Slices ⓘ

All instances 1.0000 (15,559) 0.9794 (15,559)

30 qu...

number

Metric

ndcg...

highly rated density +18.24% 0.6530 (15,559) 0.7721 (15,559)

ndcg purchase prob -6.19% 0.9921 (15,559) 0.9308 (15,559)

exact density +30.70% 0.4953 (15,559) 0.6474 (15,559)

Metadata ⓘ

keywords ( match '2t baby girl toys' ) clear all

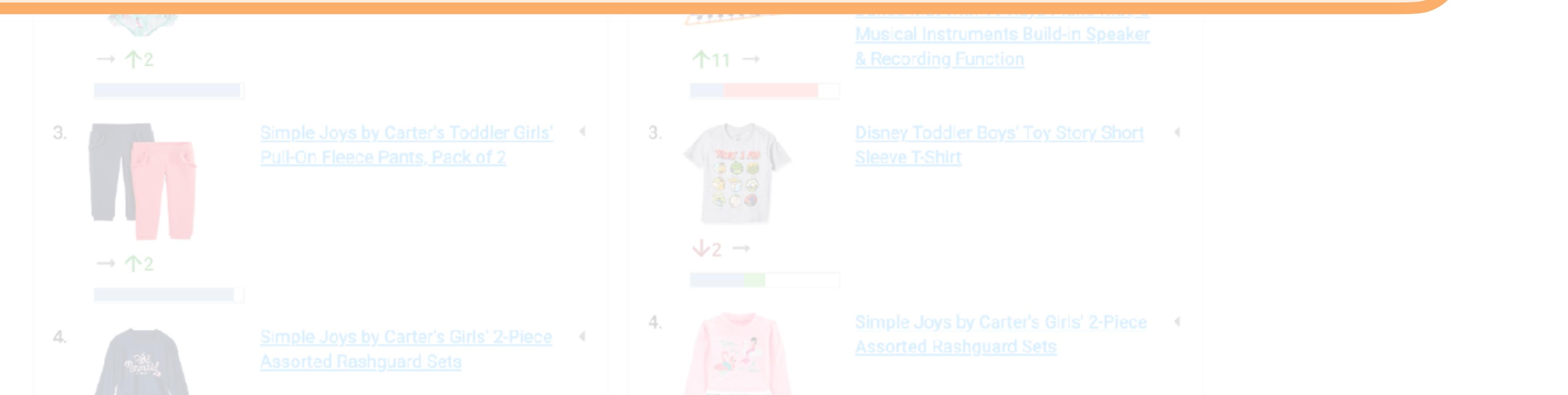
click purchase highly rated click relevant click

T1 Combined objectives: click \* 1

T2 Combined objectives: click \* 3  
relevant click \* 1  
highly rated click \* 1

Type a new model name here. '⌘ + ↵' to submit.

Orbit uses **objective-centered design** to support practitioners efficiently **explore** the design space and **evaluate** their explorations



# Evaluation

We conducted a within-subject controlled experiment (N=12) with industry practitioner to see whether Orbit helps users...

Explore the design space more efficiently

Make more informed decisions

Consider and communicate tradeoffs

# Orbit helps explore the design space more efficiently

vs. Exploring objective design in computational notebooks, Orbit...

**Objective-centric interface design**



**Explore more designs (+183%)**

**Bigger changes in exploration (+292%)**



number of exploration



number of large changes

# Orbit helps users make more informed decisions

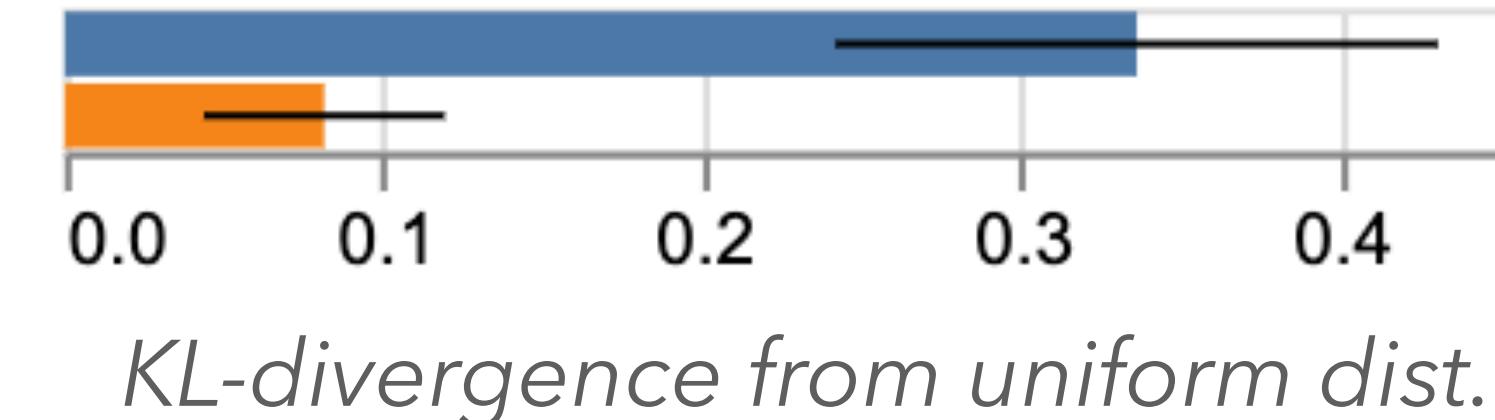
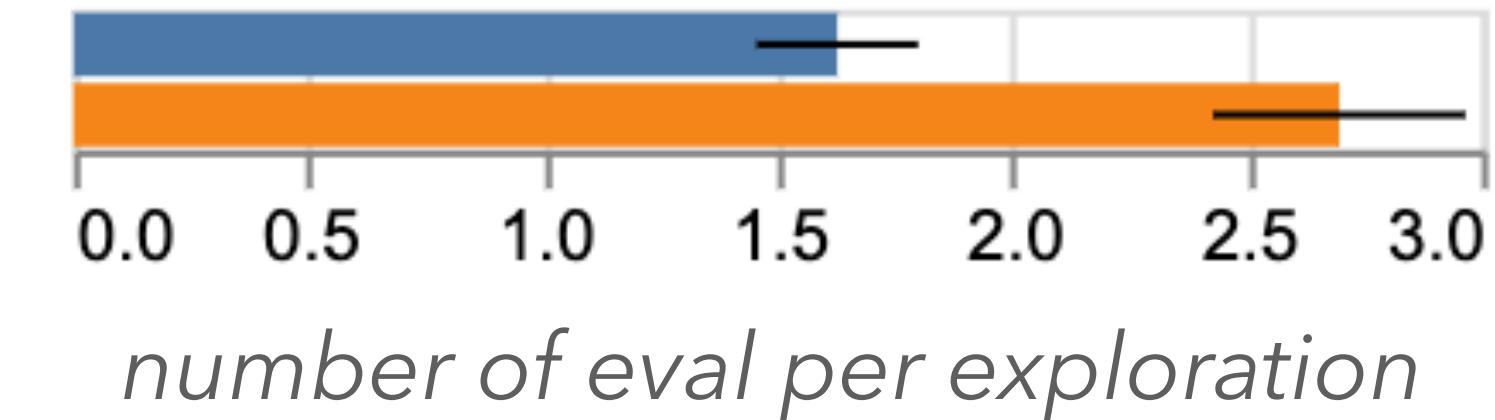
vs. Evaluating objective design in computational notebooks, Orbit...

**Gathered quan & qual eval results**



**More distinct evaluations (+66%)**

**More balanced evaluations (76%  
closer to uniform)**



# Orbit encourages users to communicate tradeoffs

vs. Objective design in computational notebooks, Orbit...

**Explicated objective design space**

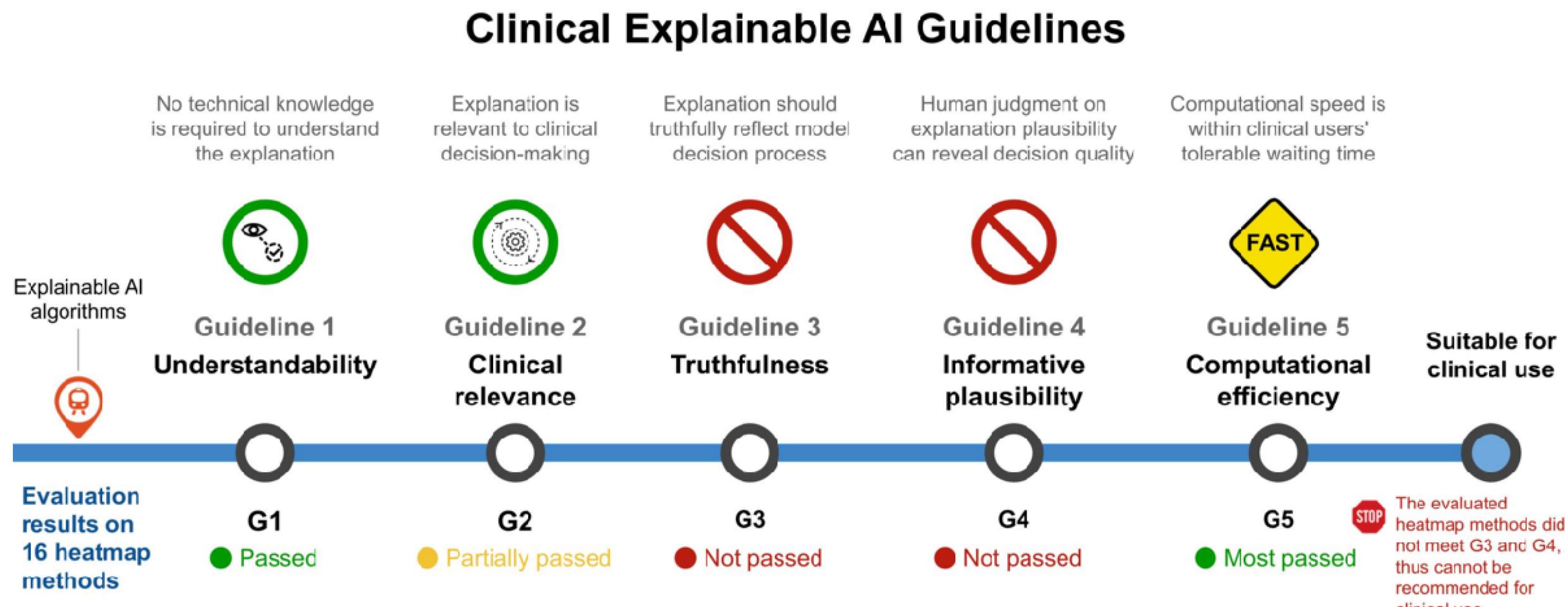


**More thorough thinking over  
trade-offs (+8.6%)**

"There exists **trade-off between objectives** such as popularity and exact...  
the weights/objectives suits for keywords with quantities may not perform good on the overall instances" (P5)

# Future Directions: Beyond Multi-objective Ranking

*Many ML problems are also multi-objectives – how do we help practitioners think about trade-offs and navigate thorough the design space?*



## LLM Criteria



### Simplicity



Does the example use only simple language by avoiding complex words/sentences, and presents situations or actions that a young child would be able to relate to?

### Faithfulness



The summary is devoid of factual errors, where a factual error is a statement that contradicts the source document, or is not directly stated, heavily implied, or logically entailed by the source document.

The evaluated heatmap methods did not meet G3 and G4, thus cannot be recommended for clinical use.

# Takeaways

## The Reality of Multi-objective Optimization

For multi-objective optimization, there is **no single "best" solution**.

Every solution needs to consider **trade-offs**.

With changing environment, there are **constantly new dimensions** (and trade-offs) to consider.

How to support model team and product team **communicate** 🗣 and **collaborate** 🤝?

How to support practitioners efficiently **explore** 🔎 the design space and **evaluate** 📈 their explorations?

Our user study shows Orbit is effective to

Explore the design space more efficiently

Make more informed decisions

Consider and communicate tradeoffs

## Key idea: Objectives-centric Design & Evaluation

Orbit uses **objective-centric designs** to support practitioners efficiently **explore** the design space and **evaluate** their explorations

Check out our paper!



amazon Carnegie Mellon University

