

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



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



**Carnegie Mellon University**



# Ranking models are prevalent in AI applications

Recommended

**BEFORE YOU BUY**

Days Gone - Before You Buy

gameranx

48K views • 2 hours ago

**Fencing, explained**

Vox

1.3M views • 2 years ago

**MIRRORLESS vs DSLR**

Why mirrorless cameras are taking over

Engadget

638K views • 1 week ago

**ODST**

The Tragedy of Halo 3: ODST | Game Maker's Toolkit

Game Maker's Toolkit

183K views • 2 days ago

**FOLDING PHONE**

HUAWEI MATE X HANDS-ON

MrMobile [Michael Fisher]

2.2M views • 2 months ago

**WHITE DWARFS**

The Last Light Before Eternal Darkness - White Dwarfs &...

Kurzgesagt - In a Nutshell

8M views • 1 year ago

**EXTRAORDINARY RITUALS**

EXTRAORDINARY RITUALS

4:14

**A bullet with a purpose**

A bullet with a purpose

2:15

Retrieval  
Augmented  
Generation (RAG)  
Future of LLMs

All baby formula

Similac Advance Infant Formula with Iron, Baby Formula Powder, 30.8-oz...

1.93 Pound (Pack of 1)

Similac 360 Total Care Infant Formula, Has 5 HMO Prebiotics, Our Closest...

2.25 Pound (Pack of 3)

Best Seller

Similac 360 Total Care Infant Formula, Has 5 HMO Prebiotics, Our Closest...

2 Fl Oz (Pack of 12)

★★★★★ 5,108

10K+ bought in past month

\$15<sup>78</sup> (\$1.32/Count)

\$14.99 with Subscribe & Save discount

Extra 40% off when you subscribe

SNAP EBT eligible

prime One-Day FREE delivery Tomorrow, Aug 7

Or FREE delivery Overnight 4 AM - 8 AM on \$25 of qualifying items

Add to cart



# Ranking models need to consider multiple objectives

All ▾ baby formula

1

OVER 48% MORE!

MADE 17 MONTHS

NO PALM OIL

NO STEAROL

Similac

ADVANCE

Complete nutrition for baby's first year

BRAIN NUTRITION

EYE HEALTH

GROWTH DEVELOPMENT

Opti-RO

INFANT FORMULA WITH IRON

NET WT. 30.8 OZ (870g)

Similac Advance Infant Formula with Iron, Baby Formula Powder, 30.8-oz...

1.93 Pound (Pack of 1)

★★★★★ 1,917

10K+ bought in past month

\$35<sup>98</sup> (\$1.17/Ounce)

List: \$38.09

\$34.18 with Subscribe & Save discount

Extra 40% off when you subscribe

SNAP EBT eligible

✓prime One-Day FREE delivery Tomorrow, Aug 7

Add to cart

1

OVER 70% MORE!

MADE 17 MONTHS

Similac

360 Total Care

Our closest formula to breast milk

3 Similac 360 Total Care

Similac 360 Total Care Infant Formula, Has 5 HMO Prebiotics, Our Closest...

2.25 Pound (Pack of 3)

★★★★★ 356

6K+ bought in past month

\$175<sup>25</sup> (\$1.62/Ounce)

\$166.49 with Subscribe & Save discount

Extra 40% off when you subscribe

✓prime One-Day FREE delivery Tomorrow, Aug 7

Add to cart

1

OVER 70% MORE!

MADE 17 MONTHS

Similac

360 Total Care

Our closest formula to breast milk

3 Similac 360 Total Care

Similac 360 Total Care Infant Formula, Has 5 HMO Prebiotics, Our Closest...

2 Fl Oz (Pack of 12)

★★★★★ 5,108

10K+ bought in past month

\$15<sup>78</sup> (\$1.32/Count)

\$14.99 with Subscribe & Save discount

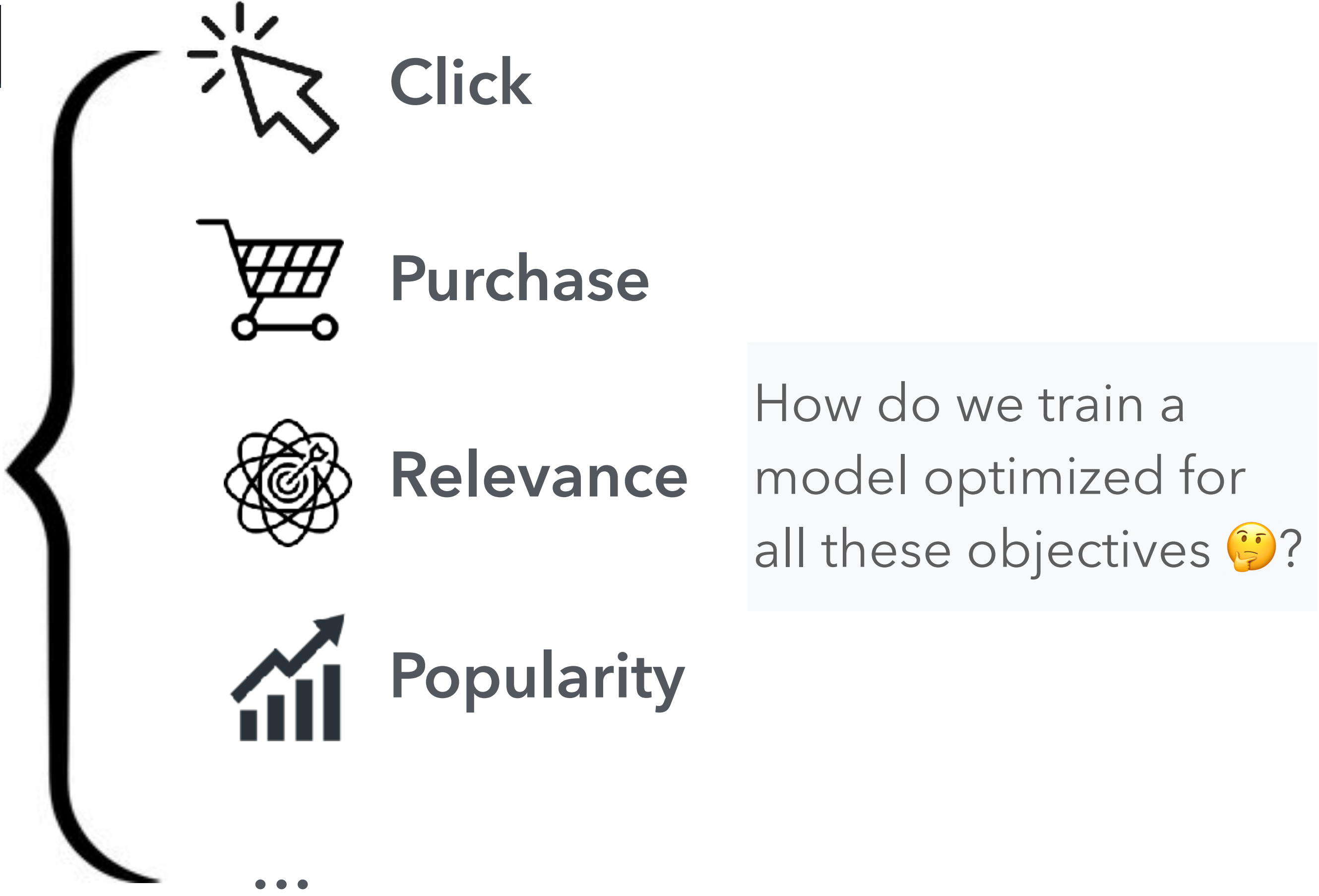
Extra 40% off when you subscribe

SNAP EBT eligible

✓prime One-Day FREE delivery Tomorrow, Aug 7

Or FREE delivery Overnight 4 AM - 8 AM on \$25 of qualifying items

Add to cart

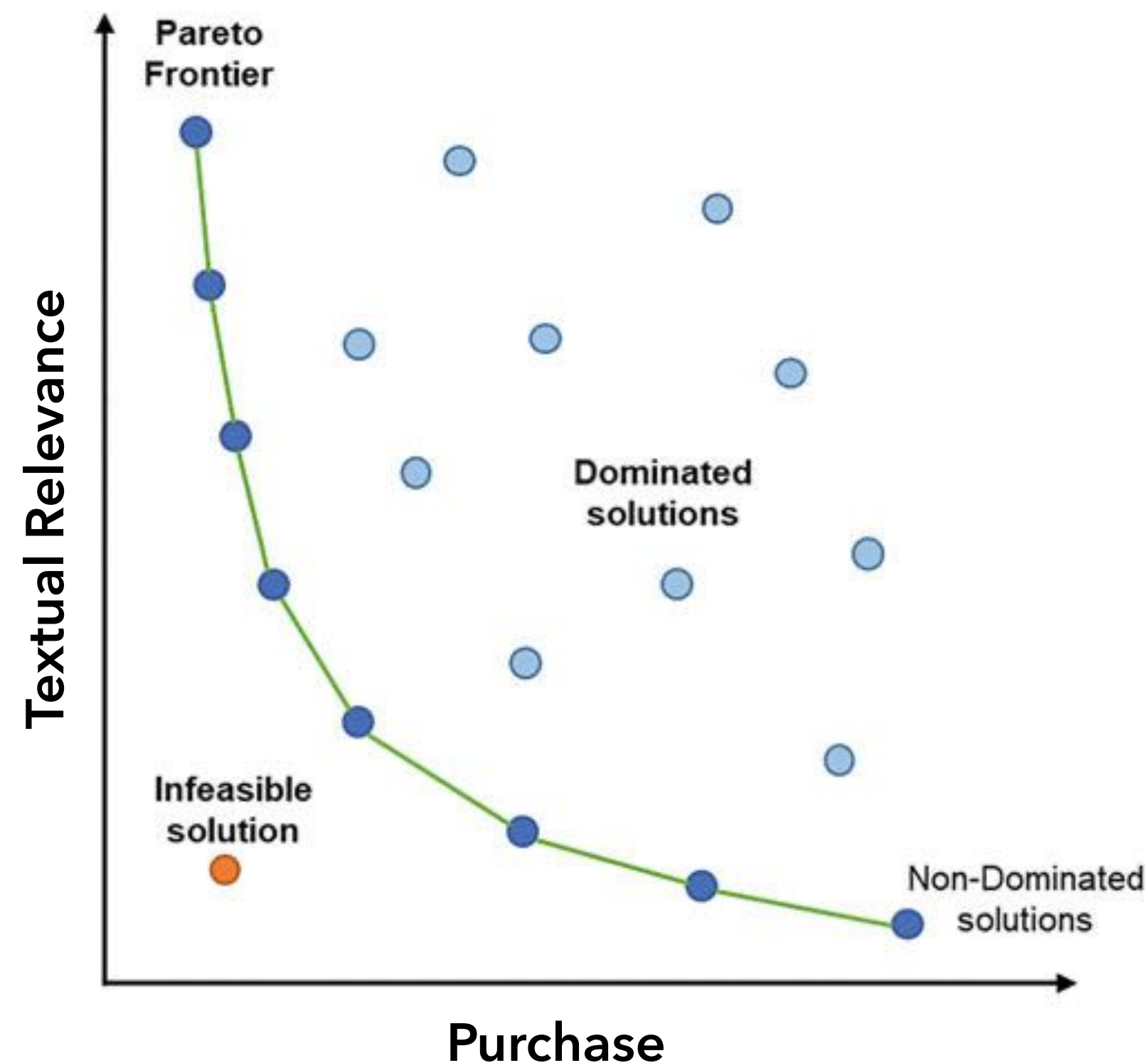


# The Reality of Multi-objective Optimization

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**!

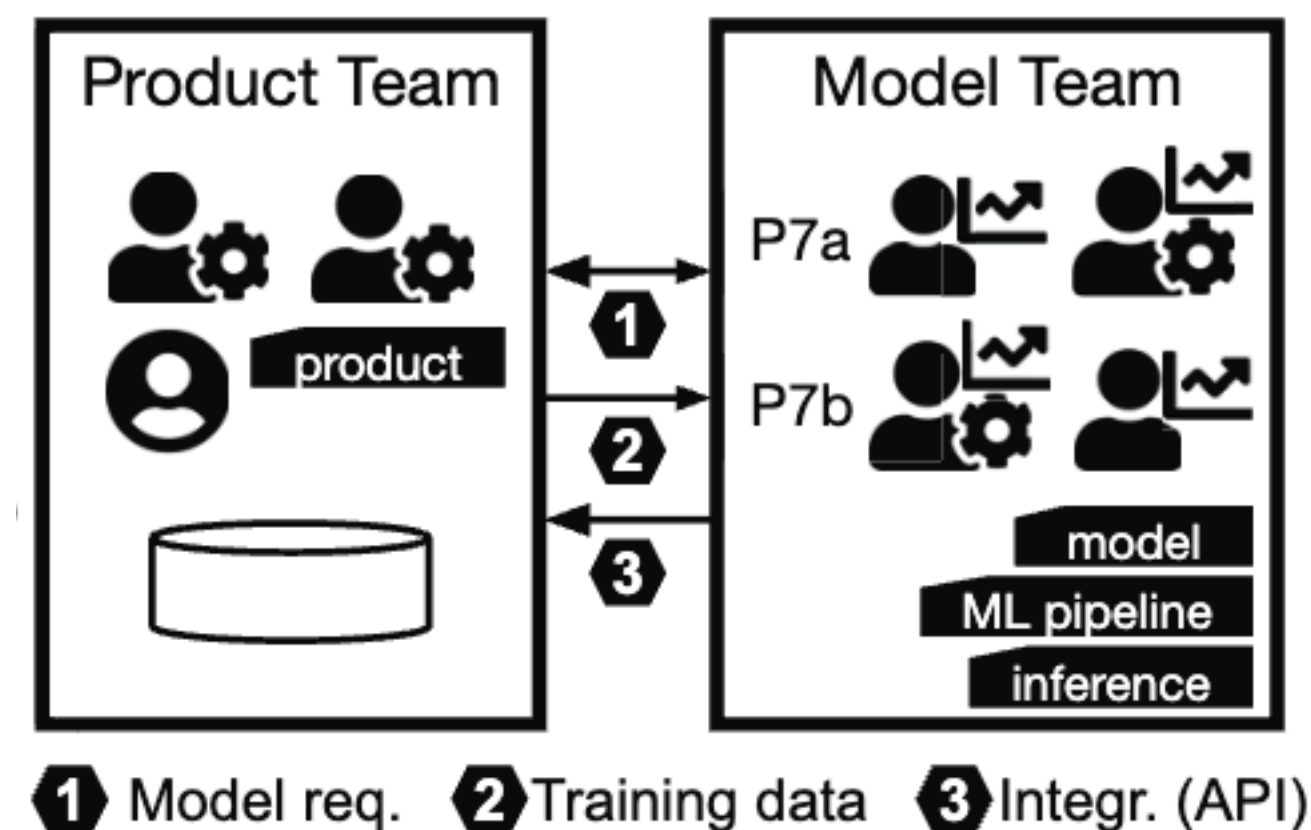
*“a wicked problem is a problem that is difficult or impossible to solve because of incomplete, contradictory, and changing requirements that are often difficult to recognize.”*

# 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!

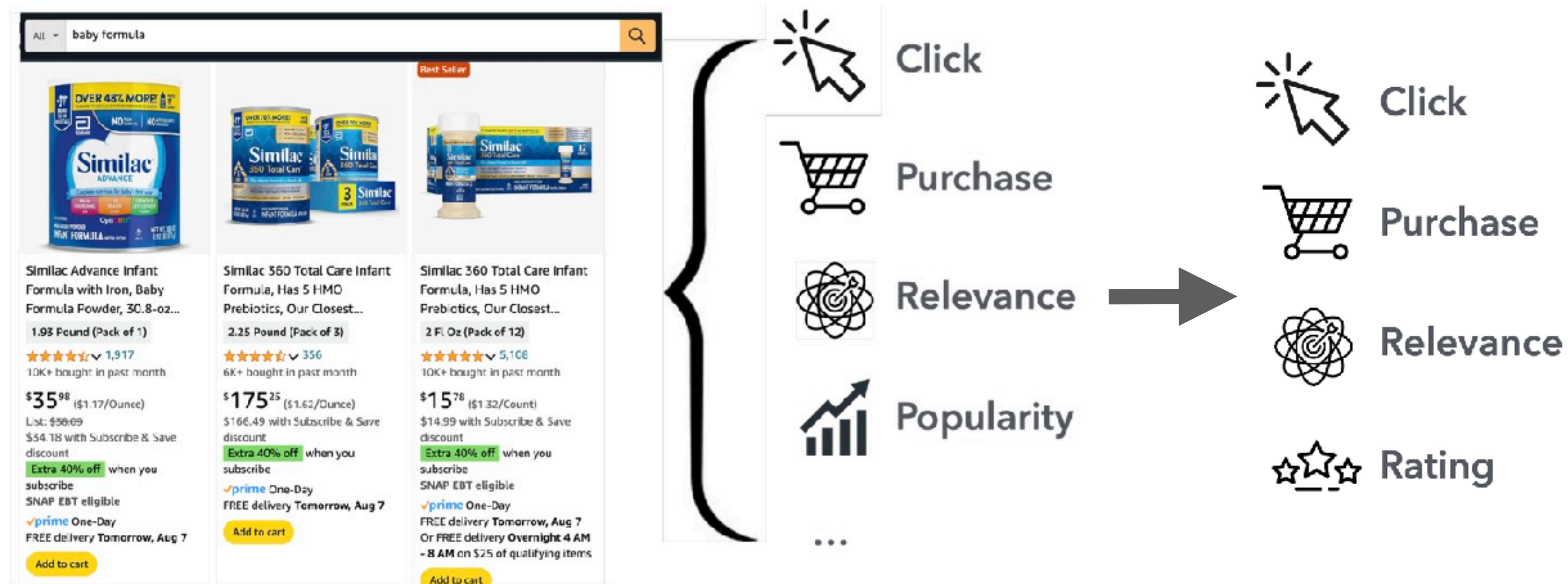
How to provide a shared language for practitioners to  
**communicate**  and **collaborate** .

1 Model req. 2 Training data 3 Integr. (API)

*Struggle to provide actionable feedback & incorporate feedback!*

Nahar, Nadia, et al. "Collaboration challenges in building ml-enabled systems: Communication, documentation, engineering, and process." Proceedings of the 44th international conference on software engineering. 2022.

# Observation #2: Design & Evaluation Challenges



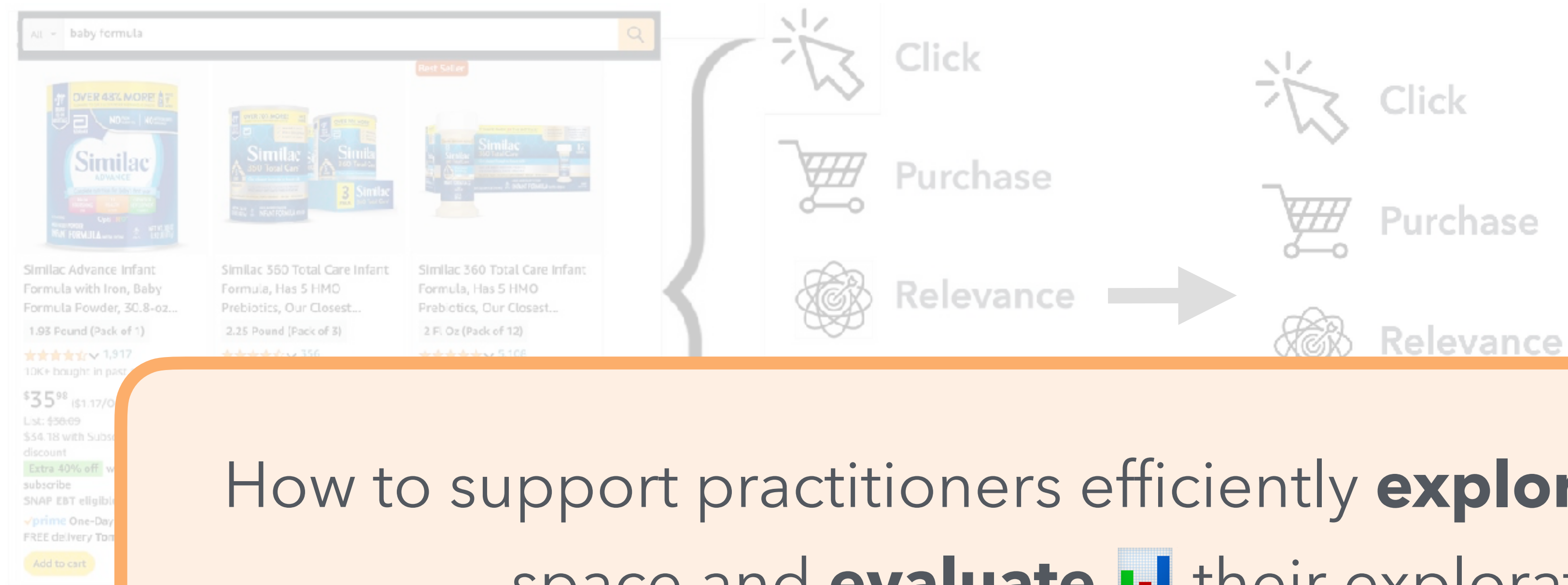
***Huge design space to iterate over***

Many different potential objectives, and different 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

# Observation #2: Design & Evaluation Challenges



How to support practitioners efficiently **explore** 🔍 the design space and **evaluate** 📊 their explorations?

*Hug*  
*iterate over*

ways to combine them into the model.

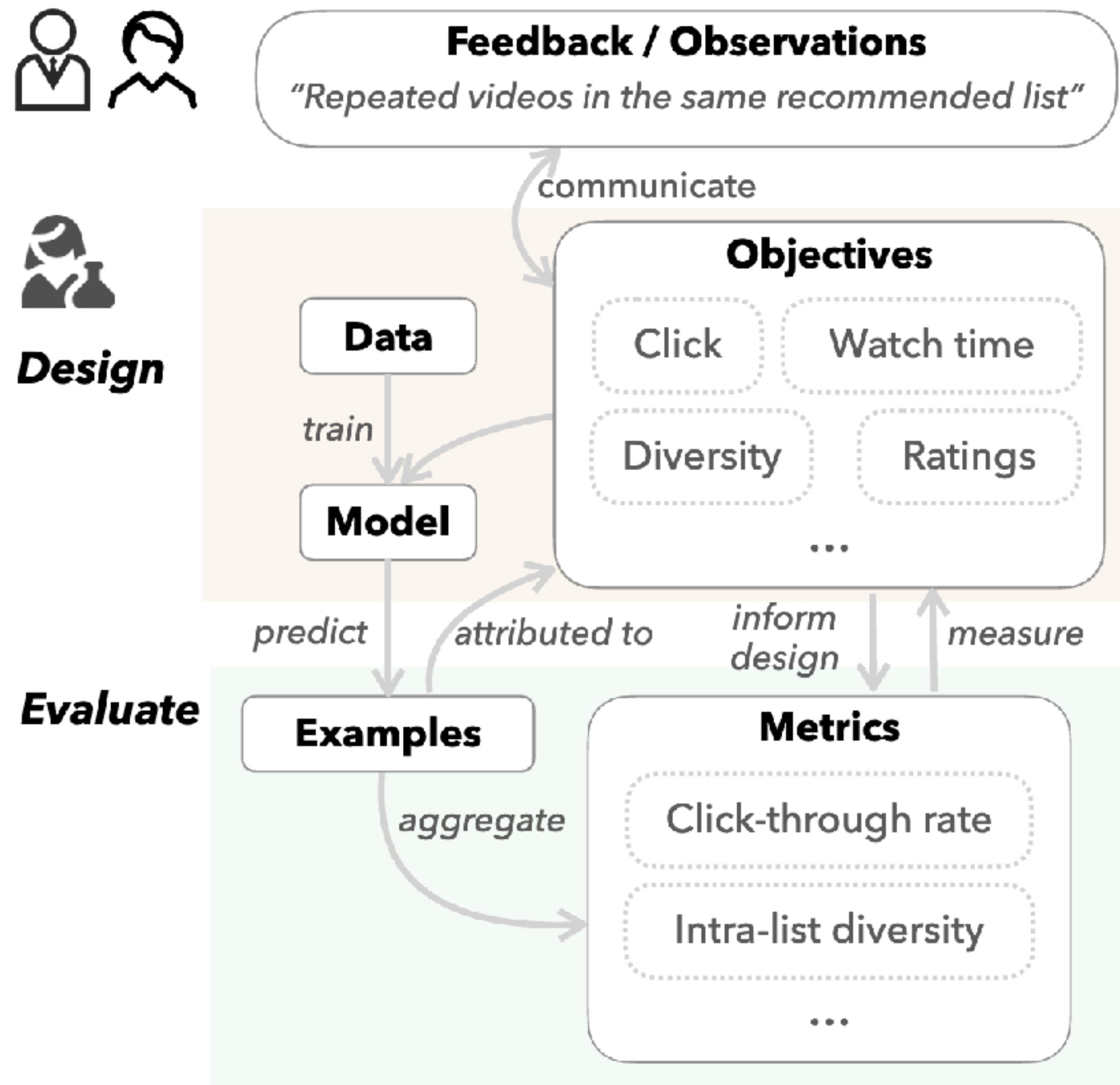
*Lots of evaluation  
information to track*

- 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

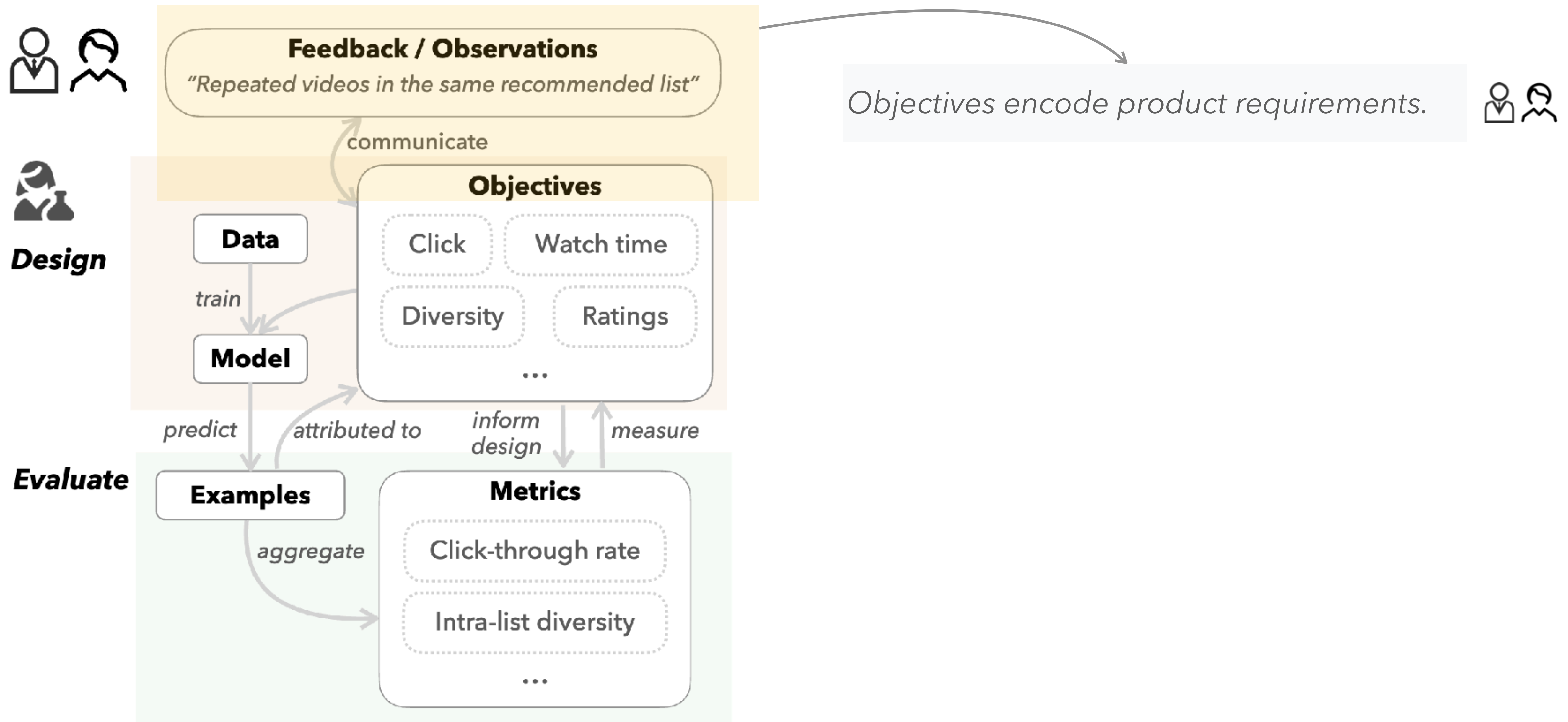


**Objectives** should be surfaced as the main object for stakeholders to **navigate** through the design space, **communicate** their findings, and **negotiate** over trade-offs.

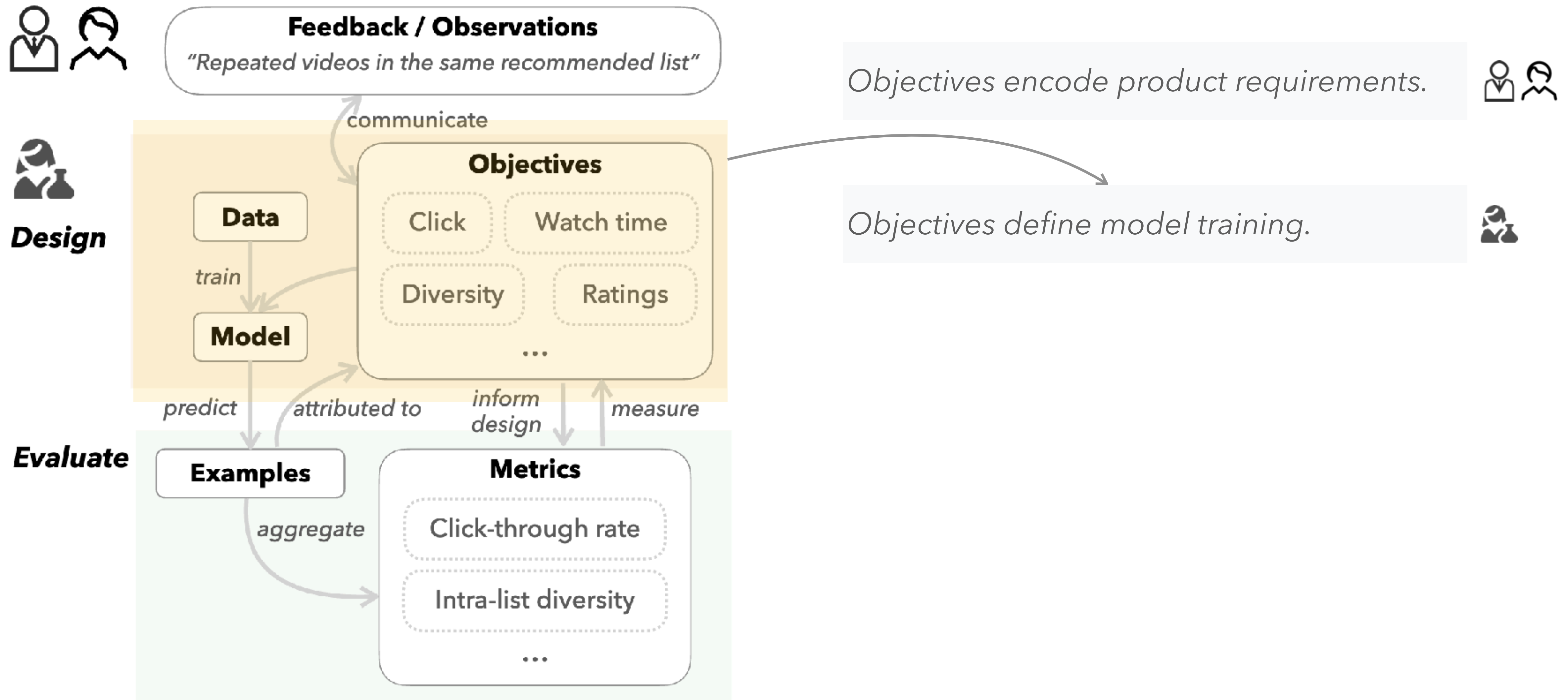
This helps provides **a shared language** and **appropriate guidance** for practitioners.



# Key idea: Objectives-centric Design & Evaluation

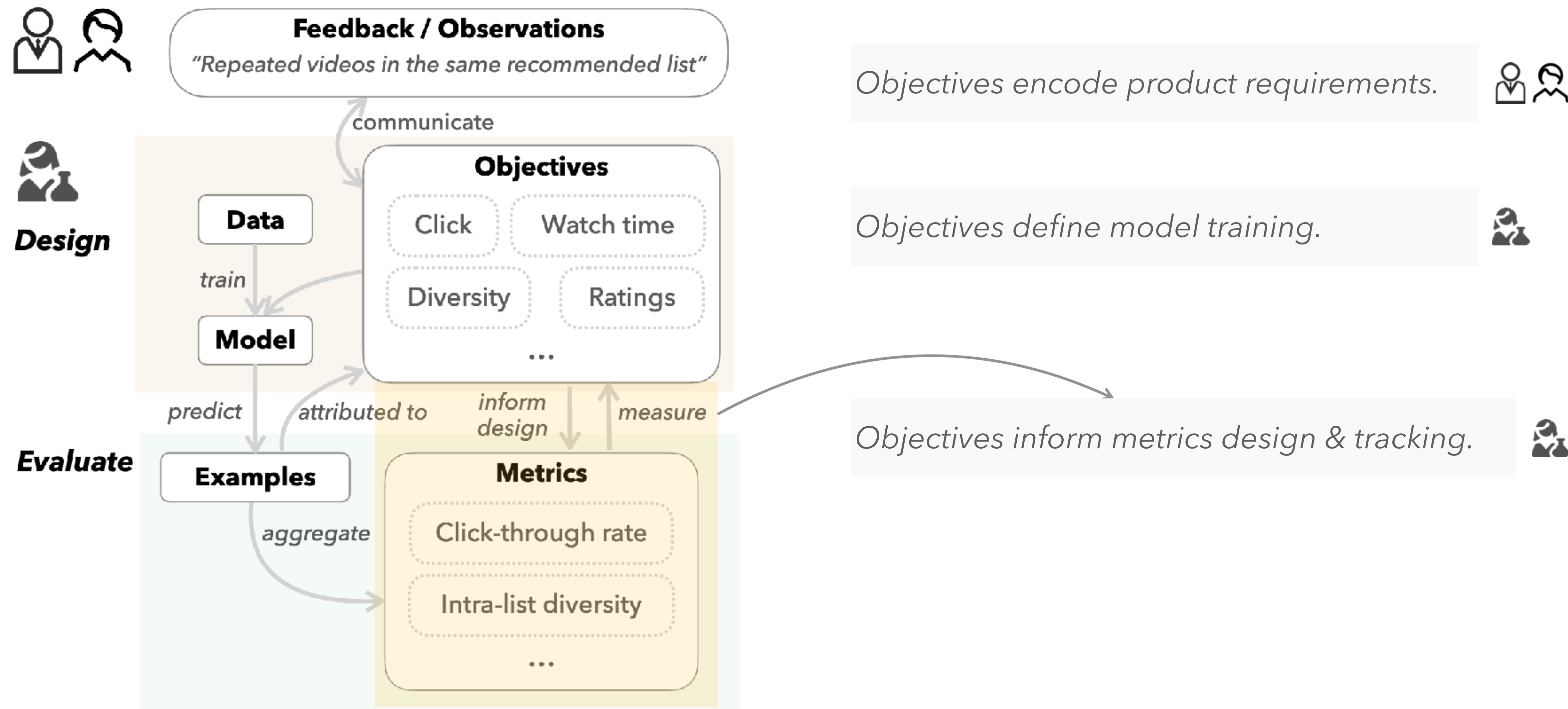


# Key idea: Objectives-centric Design & Evaluation

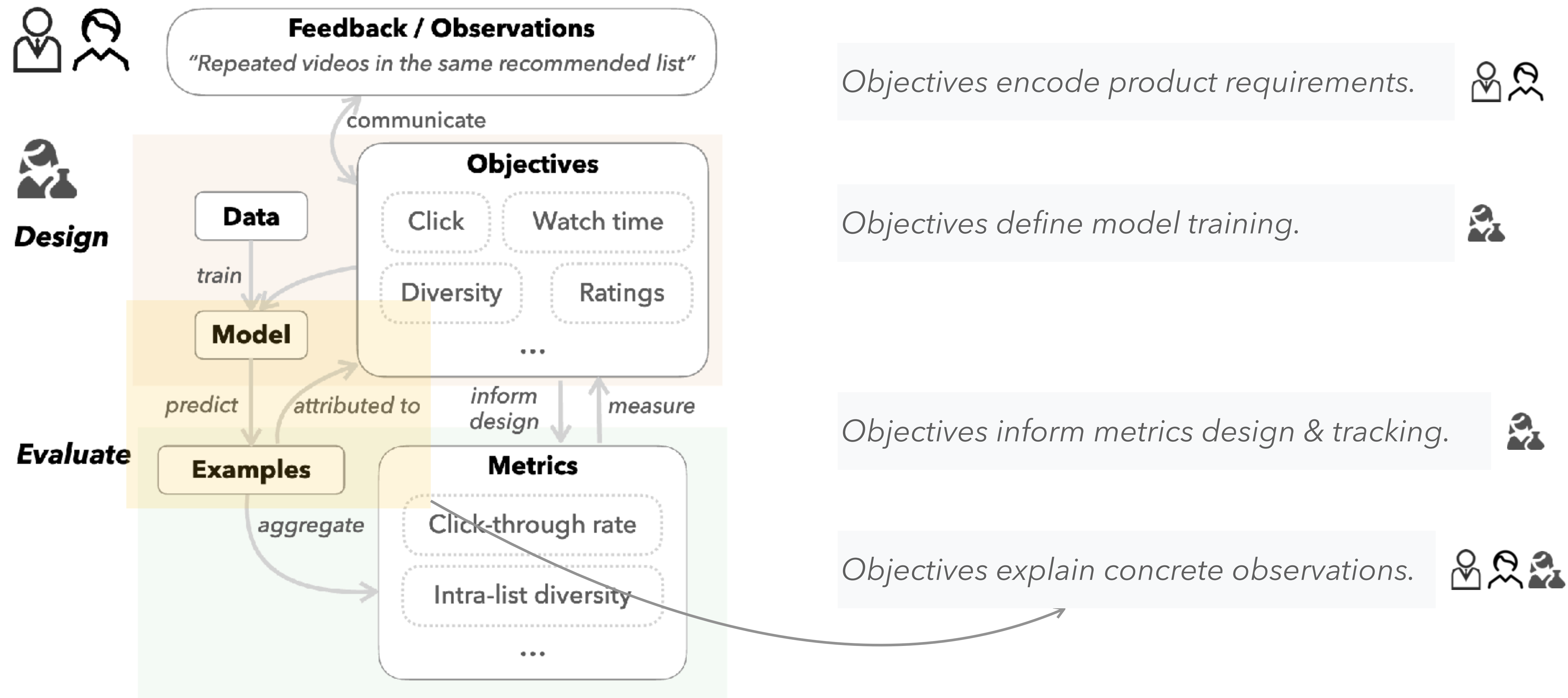




# Key idea: Objectives-centric Design & Evaluation



# Key idea: Objectives-centric Design & Evaluation





# 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** 📊

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

*Objectives encode product requirements.*



*Objectives define model training.*



*Objectives inform metrics design & tracking.*



*Objectives explain concrete observations.*





# Orbit System Walkthrough

Models

T1 x T2 x

Model AModel 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

<div>ndcg_click_prob</div> <div>-2.06%</div>	1.0000 (15,559)	0.9794 (15,559)
<div>highly_rated_density</div> <div>+18.24%</div>	0.6530 (15,559)	0.7721 (15,559)
<div>ndcg_purchase_prob</div> <div>-6.19%</div>	0.9921 (15,559)	0.9308 (15,559)
<div>exact_density</div> <div>+30.70%</div>	0.4953 (15,559)	0.6474 (15,559)

Metadata

keywords

NOT

Search

Aa \_ab\_\*

SET

match '2t baby girl toys'

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

click purchase highly\_rated\_click relevant\_click


T1

Combined objectives: click \* 1

2t baby girl toys

RANKING:

1.




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

→

↑2

2.




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

→

↑2

3.




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

→

↑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

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

2t baby girl toys

RANKING:

1.




[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

→

2.




[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

→

3.




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

↓2

→

4.



[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

T1

Model B

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'

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

click

purchase

highly\_rated\_click

relevant\_click

1. Objectives Overview


T1

Combined objectives: click \* 1

2t baby girl toys

RANKING:

1.




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

→

↑2

2.




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

→

↑2

3.




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

→

↑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

RANKING:

1.




[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

→

2.




[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

→

3.




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

↓2

→

4.



[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 AModel B

T1

T2

Metric

ndcg\_click\_prob

Slices

All instances1.00000.9794(15,559)(15,559)

30 quart coolers1.00000.9912(15)(15)

numerical1.00000.9766(3,194)(3,194)

Metrics

ndcg\_click\_prob1.00000.9794-2.06%(15,559)(15,559)

highly\_rated\_density0.65300.7721+18.24%(15,559)(15,559)

ndcg\_purchase\_prob0.99210.9308-6.19%(15,559)(15,559)

exact\_density0.49530.6474+30.70%(15,559)(15,559)

Metadata

keywords

NOTSearchAa\_ab.\*SET

match '2t baby girl toys'

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

click purchase highly\_rated\_click relevant\_click

1. Objectives Overview

Drag the objective to apply it.

click\_probabilityclick \* 1

T2

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

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

2t baby girl toys

RANKING:

1.

Disney Toddler Boys' Toy Story Short Sleeve T-Shirt

→ ↑2

2.

Simple Joys by Carter's Girls' 2-Piece Assorted Rashguard Sets

→ ↑2

3.

Simple Joys by Carter's Toddler Girls' Pull-On Fleece Pants, Pack of 2

→ ↑2

4.

Simple Joys by Carter's Girls' 2-Piece Assorted Rashguard Sets

2t baby girl toys

RANKING:

1.

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 →

2.

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 →

3.

Disney Toddler Boys' Toy Story Short Sleeve T-Shirt

↓2 →

4.

Simple Joys by Carter's Girls' 2-Piece Assorted Rashguard Sets

Query Groups Per Page11 - 1 of 1

Models

T1 x T2 x

Model AModel 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\_prob1.00000.9794-2.06%(15,559)(15,559)

highly\_rated\_density0.65300.7721+18.24%(15,559)(15,559)

ndcg\_purchase\_prob0.99210.9308-6.19%(15,559)(15,559)

exact\_density0.49530.6474+30.70%(15,559)(15,559)

Metadata

keywords

NOT

Search

Aa\_ab.\*SET

match '2t baby girl toys'

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

Disney Toddler Boys' Toy Story Short Sleeve T-Shirt

→ ↑2

2.

Simple Joys by Carter's Girls' 2-Piece Assorted Rashguard Sets

→ ↑2

3.

Simple Joys by Carter's Toddler Girls' Pull-On Fleece Pants, Pack of 2

→ ↑2

4.

Simple Joys by Carter's Girls' 2-Piece Assorted Rashguard Sets

2t baby girl toys

RANKING:

1.

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 →

2.

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 →

3.

Disney Toddler Boys' Toy Story Short Sleeve T-Shirt

↓2 →

4.

Simple Joys by Carter's Girls' 2-Piece Assorted Rashguard Sets

Query Groups Per Page11 - 1 of 1



Models

T1 x T2 x

Model A

T1

Model B

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'

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.

2t baby girl toys

RANKING:

1.

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

→ ↑2

2.

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

→ ↑2

3.

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

→ ↑2

4.

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

2t baby girl toys

RANKING:

1.

[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 →

2.

[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 →

3.

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

↓2 →

4.

[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 × T2 ×

Model A

T1

Model B

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 ⓘ

<div>ndcg_click_prob</div> <div>-2.06%</div>	1.0000 (15,559)	0.9794 (15,559)
<div>highly_rated_density</div> <div>+18.24%</div>	0.6530 (15,559)	0.7721 (15,559)
<div>ndcg_purchase_prob</div> <div>-6.19%</div>	0.9921 (15,559)	0.9308 (15,559)
<div>exact_density</div> <div>+30.70%</div>	0.4953 (15,559)	0.6474 (15,559)

Metadata ⓘ

keywords

NOT

Search

Aa \_ab . \*

SET

match '2t baby girl toys' ×

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.

2t baby girl toys


RANKING:

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


→ ↑2

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

→ ↑2


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

→ ↑2


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

2t baby girl toys


RANKING:

1.  [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 →

2.  [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 →

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

↓2 →

4.  [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%

### 3. Quantitative Evaluation

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 x

T2 Combined objectives: click \* 3 x + relevant\_click \* 1 x + highly\_rated\_click \* 1 x

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


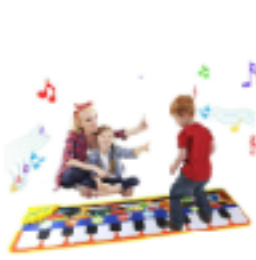


2t baby girl toys

RANKING:

1.		<a href="#">Disney Toddler Boys' Toy Story Short Sleeve T-Shirt</a>
2.		<a href="#">Simple Joys by Carter's Girls' 2-Piece Assorted Rashguard Sets</a>
3.		<a href="#">Simple Joys by Carter's Toddler Girls' Pull-On Fleece Pants, Pack of 2</a>
4.		<a href="#">Simple Joys by Carter's Girls' 2-Piece Assorted Rashguard Sets</a>

2t baby girl toys

RANKING:

1.		<a href="#">ToyVelt Toy Piano for Toddler Girls – Cute Piano for Kids with Built-in Microphone &amp; Music Modes - Best Birthday Gifts for 3 4 5 Year Old Girls – Educational Keyboard Musical Instrument Toys</a>
2.		<a href="#">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 &amp; Recording Function</a>
3.		<a href="#">Disney Toddler Boys' Toy Story Short Sleeve T-Shirt</a>
4.		<a href="#">Simple Joys by Carter's Girls' 2-Piece Assorted Rashguard Sets</a>

Query Groups Per Page 1 1 - 1 of 1



Models

T1 × T2 ×

Model A

T1

Model B

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 ⓘ

<div>ndcg_click_prob</div> <div>-2.06%</div>	1.0000 (15,559)	0.9794 (15,559)
<div>highly_rated_density</div> <div>+18.24%</div>	0.6530 (15,559)	0.7721 (15,559)
<div>ndcg_purchase_prob</div> <div>-6.19%</div>	0.9921 (15,559)	0.9308 (15,559)
<div>exact_density</div> <div>+30.70%</div>	0.4953 (15,559)	0.6474 (15,559)

Metadata ⓘ

keywords

NOT

Search

Aa \_ab .\*

SET

match '2t baby girl toys'

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.

2t baby girl toys

RANKING:


1.



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

→ ↑2


2.



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

→ ↑2


3.



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

→ ↑2

4.



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

2t baby girl toys

RANKING:

1.



[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 →

2.



[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 →

3.



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

↓2 →

4.



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

Query Groups Per Page

1

1 - 1 of 1

⏪ ⏴ ⏵ ⏩



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

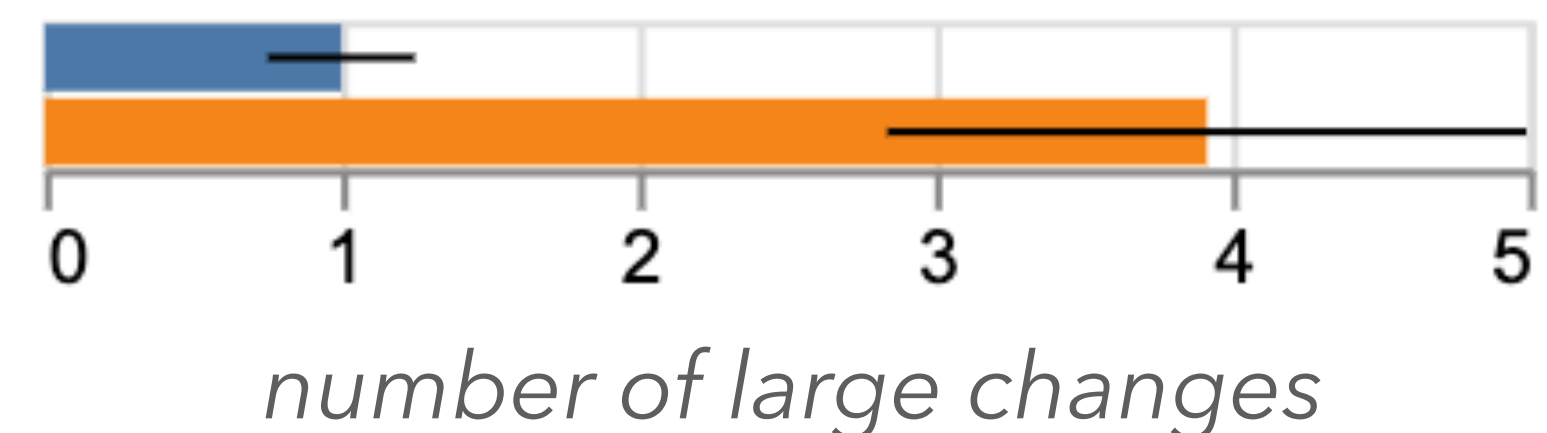
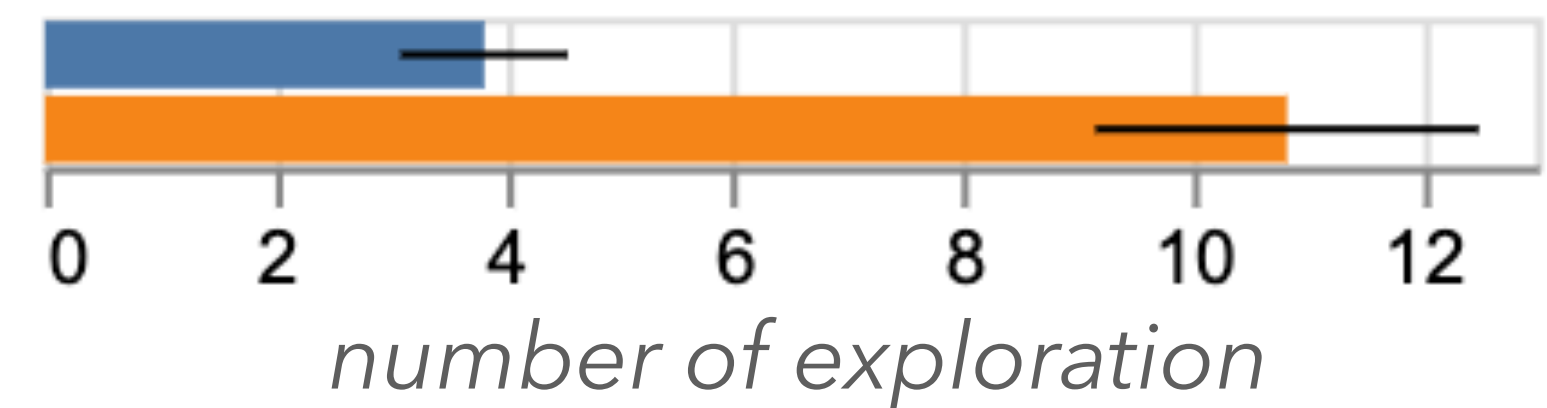
vs. Exploring objective design in computational notebooks, Orbit...

*Objective-centric interface design*



*Explore more designs (+183%)*

*Bigger changes in exploration (+292%)*





# Orbit helps users make more informed decisions

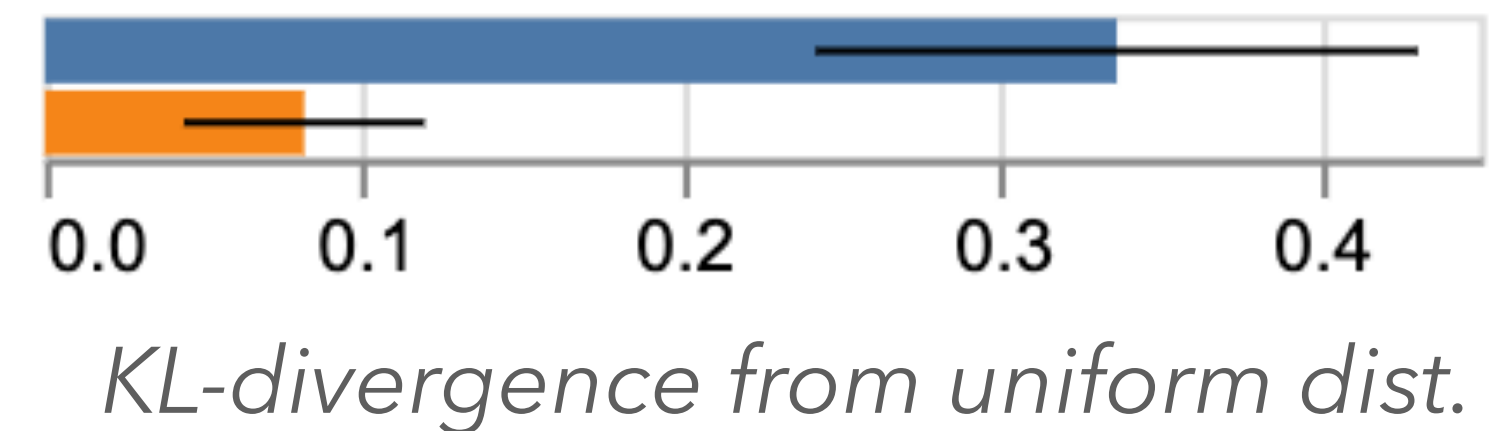
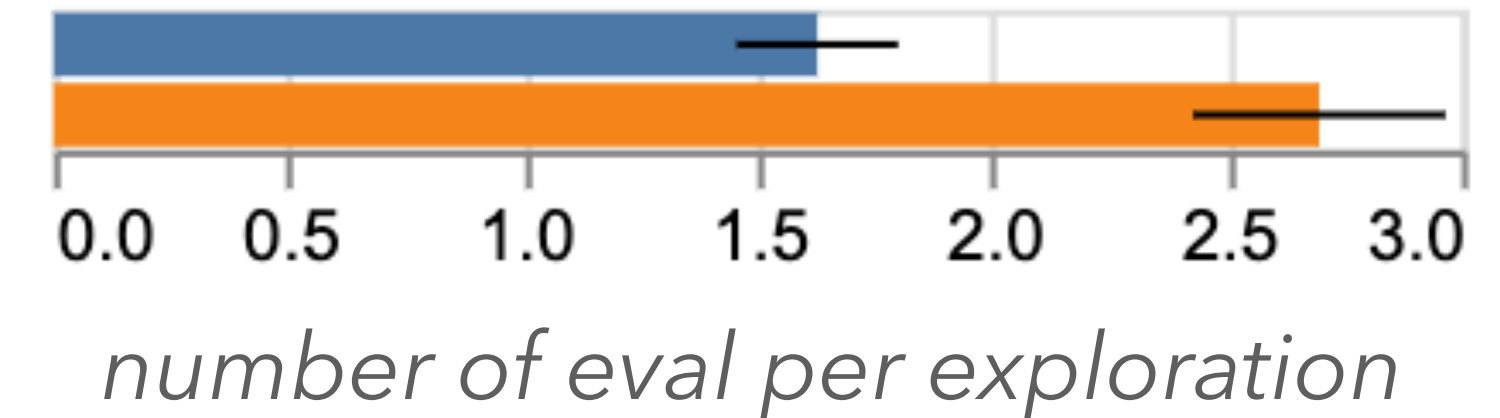
vs. Evaluating objective design in computational notebooks, Orbit...

*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*



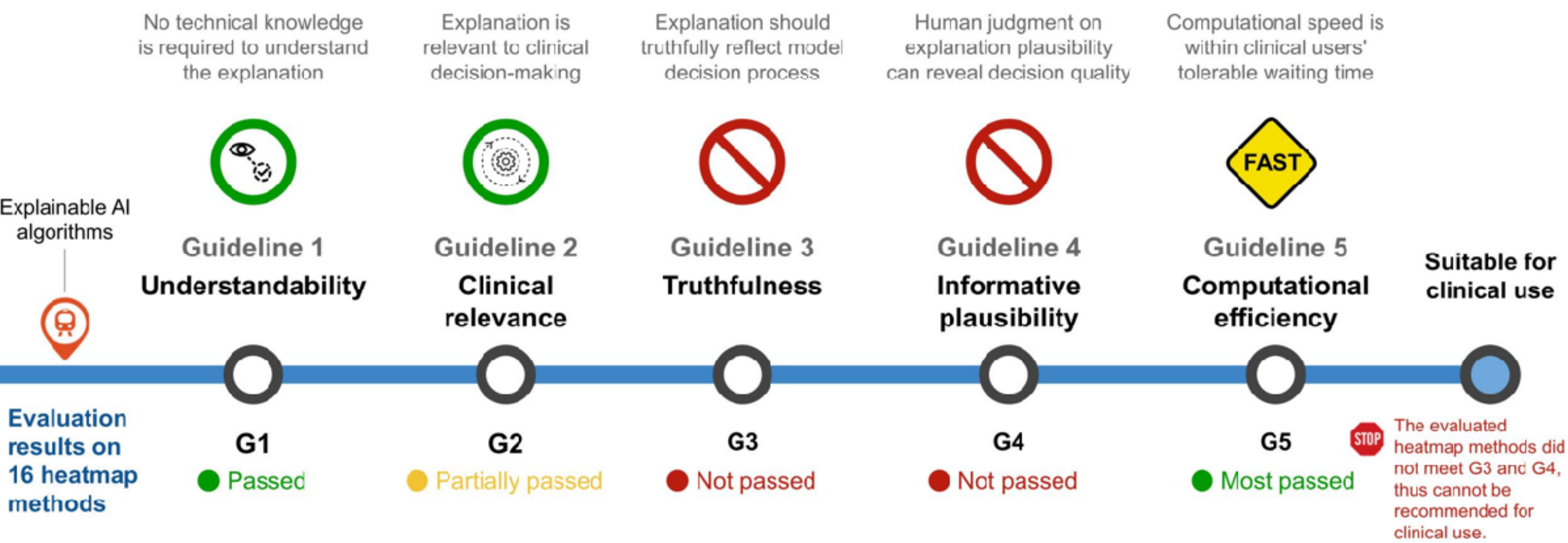
*More thorough thinking over  
trade-offs (+8.6%)*

*“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?*

## Clinical Explainable AI Guidelines



## 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.



# 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?

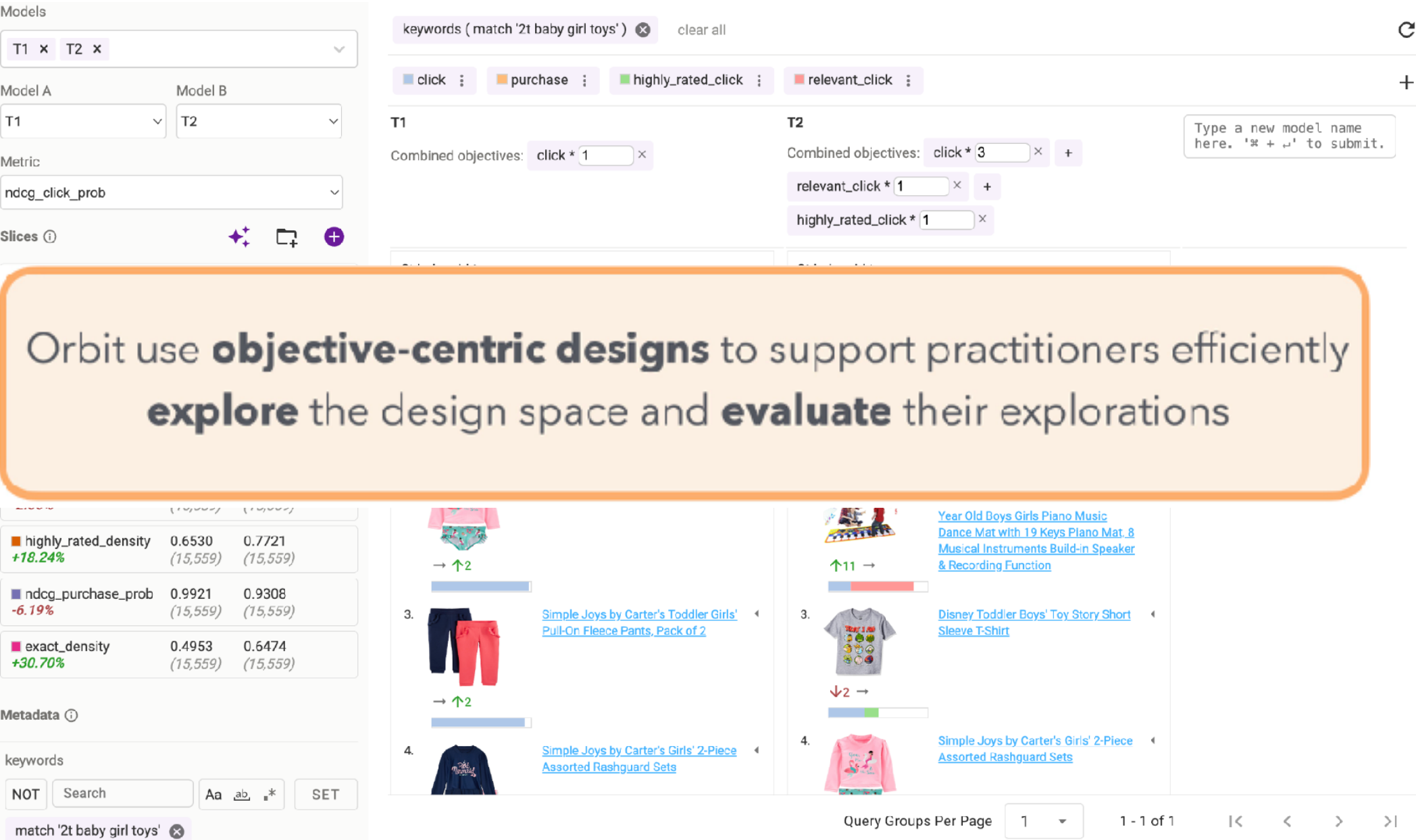
## Our user study shows Orbit is effective to

Explore the design space more efficiently

Make more informed decisions

Consider and communicate tradeoffs

## Key idea: Objectives-centric Design & Evaluation



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

Check out our paper!



amazon Carnegie Mellon University

