Measuring & Optimizing Findability + GMV

in eCommerce

CXP commerce experts

AGENDA

1. Getting the Basics right

2. A large-scale Measurement of Search Quality

3. A new Composite Model for eCommerce Search Sessions

4. Experiments & Results

CXP commerce experts

Measuring Search Quality

are the results served by
an e-commerce engine
for a given query
good or not?

CXP commerce experts

Getting the Basics right

1.Defining Quality

Is it perceived Relevance?
Is it Search Bounce rate?
Is it Search CTR?
Is it Search CR?
Is it GMV contribution?
Is it CLV?
… or a combination of all?

2.Measuring Quality

Explicit Feedback

Human Quality Judgments

Implicit Feedback

derived from various user activity signals as a proxy for Search Quality.

CXP commerce experts

Getting the Basics right

3.Measure correctly

Be aware of bots and crawlers

sometimes up to 60% of the searches are not explicitly requested by users

Correctly track search-redirects, search-campaings, etc.

from our experience only 7 out of 10 do this correctly

4.Be aware of Bias

Presentation-bias

Promotions-bias

Position-bias

MRR vs. Result-size-bias

CXP commerce experts

State-of-the-art Approaches

Explicit Feedback

Human Relevance Judgments

Let human experts label search results from an ordinal rating.

From there we can calculate NDCG, expected reciprocal rank and weighted information gain

almost impossible to scale

Implicit Feedback

User Engagement Metrics

We can use implicit feedback derived from various user activity signals.

CTR, MRR…

noisy

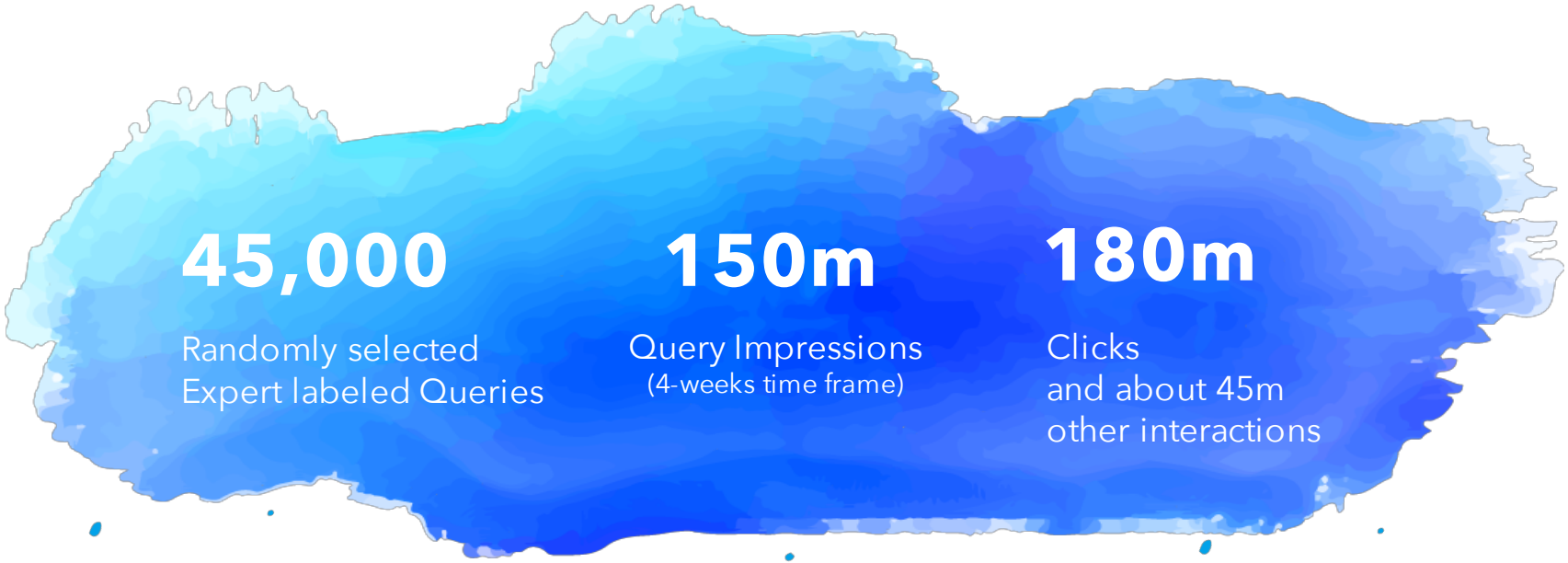
CXP commerce experts

Validation

a large-scale Measurement of Search Quality in eCommerce

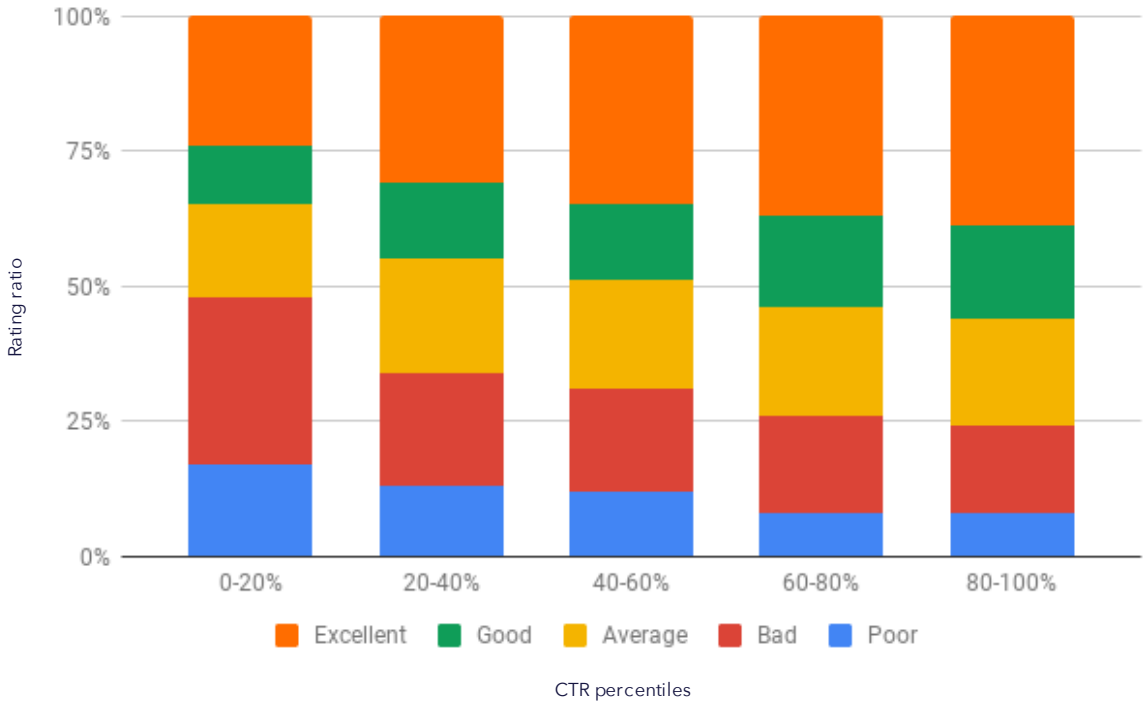
CXP commerce experts

Our - Are we doing it right? - study @ search|hub.io



CXP commerce experts

Search Result Ratings vs CTR percentile buckets

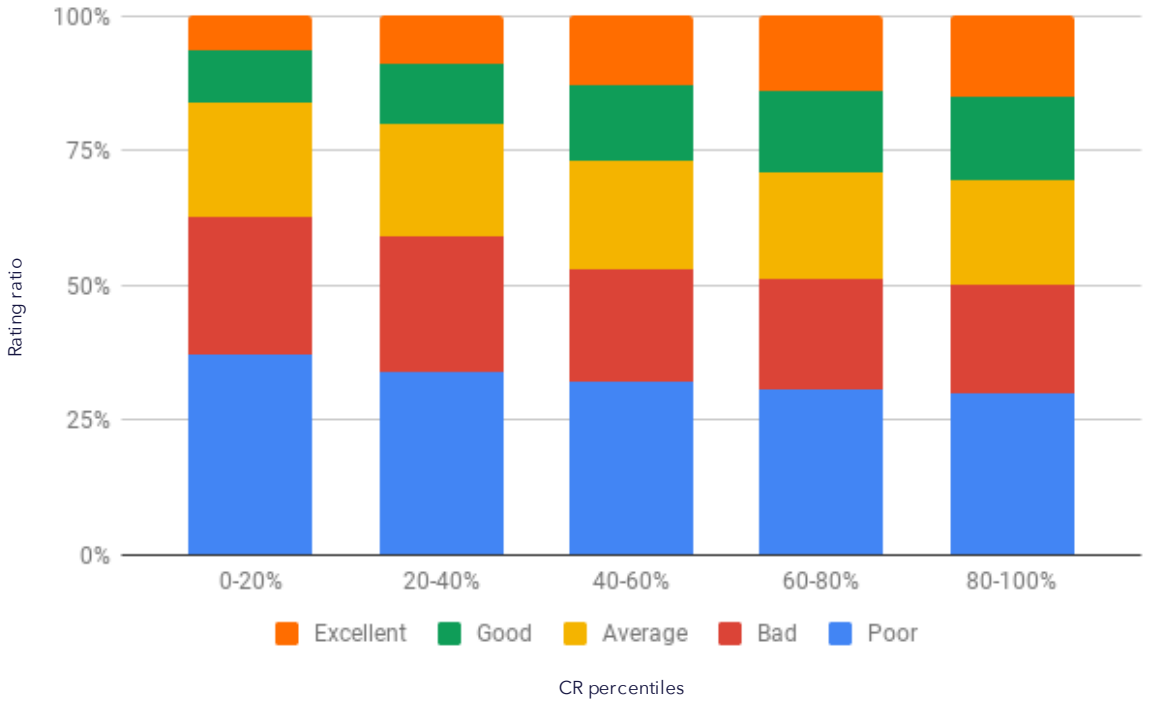


Not really what we where expecting to see?

only 53% of the highly clicked SERPs have Ratings >= 4

CXP commerce experts

Search Result Ratings vs CR percentile buckets



Oh no –
it's getting worse

only 50% of the highly converting SERPs have Ratings >= 3

CXP commerce experts

Query = bicycle

CXP commerce experts

Query = bicycle

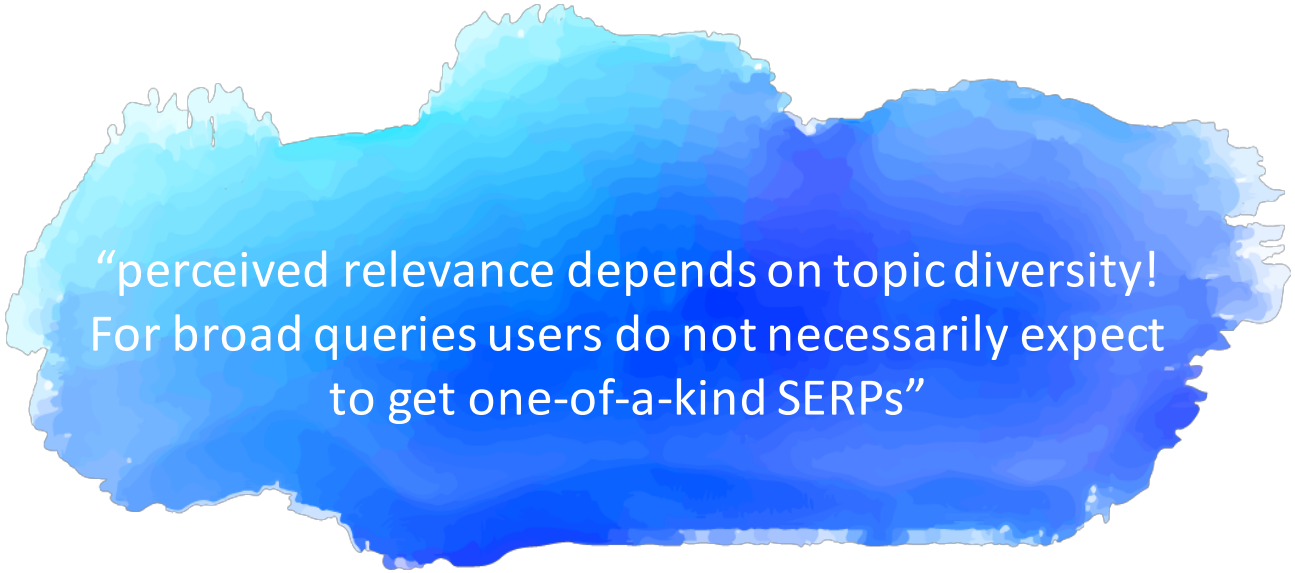
Expert Rating - 5

+21% Clicks

+17% GMV

Expert Rating - 2

CXP commerce experts



CXP commerce experts

Query = women shoes

CXP commerce experts

Query = women shoes

Expert Rating - 5

Expert Rating - 5

CXP commerce experts

CXP commerce experts

CXP commerce experts

A New Approach

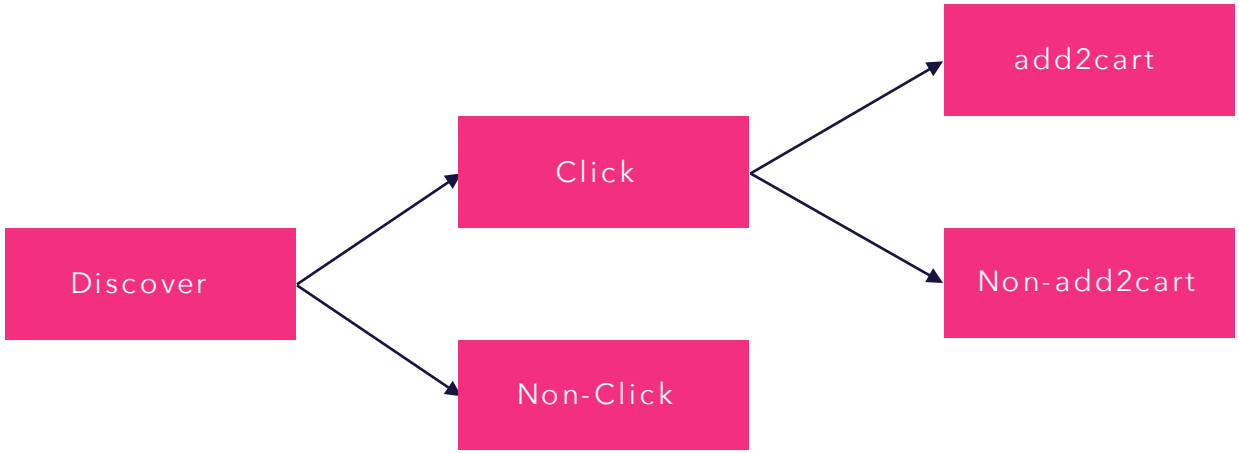
Composite Model for Measuring Search Quality in eCommerce

CXP commerce experts

What do we want to optimize?

Our Goal is to maximise the expected SERP interaction probability and GMV contribution. Where eCommerce search consists of two different stages.

Picking a candidate (click) and deciding to purchase (add2cart)

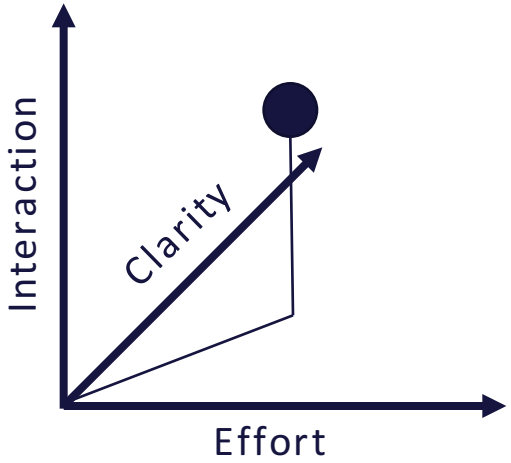


$$P(\Phi = 1 \mid i, q) = P(\Omega = 1 \mid i, q) P(\Phi = 1 \mid \Omega = 1, i, q)$$

CXP commerce experts

Optimizing the entire search shopping journey

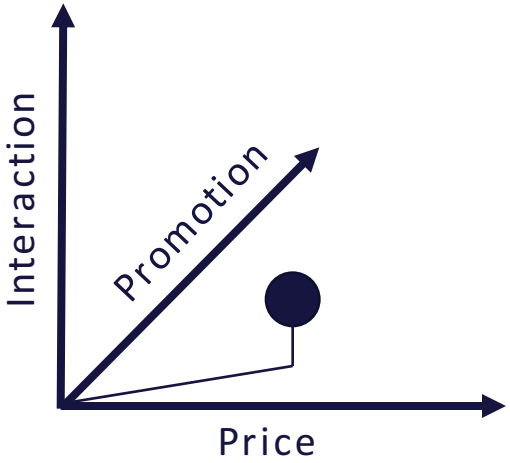
Findability $f_c()$



Click Probability

+

Sellability $f_s()$



Cart Probability

CXP commerce experts

Findability: a straight forward Model

Intuitively Findability is a measure for the ease with which information can be found. However the accurate you can specify what you are searching for the easier it might be.

$$f_C = f(clarity,\ effort,\ Impressions,\ldots)$$

clarity – a measure of how specific or broad a query is – Query Intent Entropy

effort – a measure of the effort to navigate through the search-result in order to find specific products

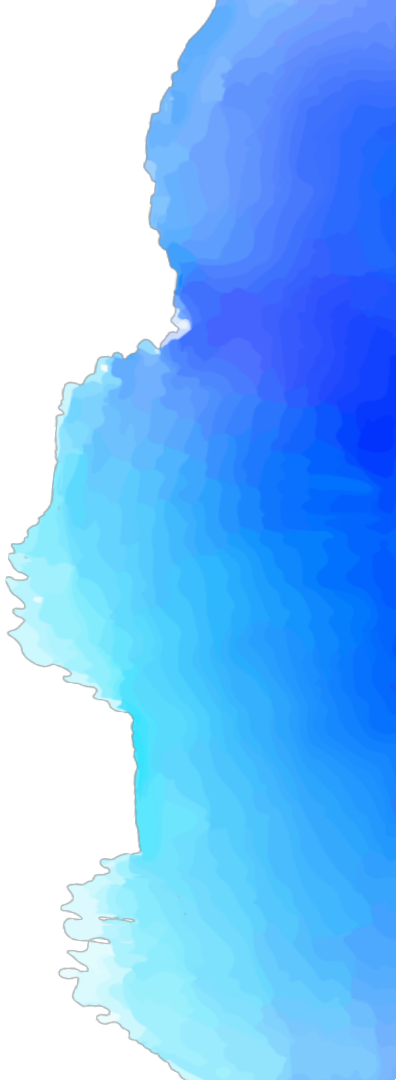
CXP commerce experts

Sellability: a straight forward Model

Intuitively Sellability can be seen as a binary measure.
The selected item is added to the basket or not.

$$f_s = f(price, promotion, add\text{-}2\text{-}basket, \ldots)$$

a measure of the relative price-drop for a specific product



CXP commerce experts

Optimization function

We model Findability as a LTR-Problem and directly optimize NDCG
While Sellability is modeled as a binary classification problem

- Revenue Contribution $= \sum_{s \in S} \sum_{i \in i^S} Price(i) P\left(\Phi = 1 \mid i^S, q^S\right)$

Price of item i

Probability of an add-2-cart

CXP commerce experts

Experiment

Composite Model for Measureing Search Quality in eCommerce

CXP commerce experts

Experiments

Evaluation Metrics

- Ranking Metric: NDCG
- Revenue Metric : Revenue/query@k

Baseline Models

Click
- RankNet
- RankBoost
- LambdaRank
- LambdaMART

Purchase
- SVM
- Logistic Regression
- Random Forest

Both
- Our tuned composite Model (CCM)

CXP commerce experts

Findability - Features

Activity Time

- Time to first Click
- Time to first Refinement
- Time to first add to Cart
- Dwell time of the query

Positional

- Position of first product clicked
- Positions seen but not clicked
- Top-k Click rate

Activity aggregates

- Number of clicks
- Number of cart adds
- Number of filters applied
- Number of sorting changes
- Number of impressions
- Click Success
- Cart Success

CXP commerce experts

Findability - Features

Query specifics

- Query Length by chars
- Query Length by words
- Contains specifiers
- Contains modifiers
- Contains range specifiers
- Contains units

Query Meta Data

- Query Intent Category**
- Query type (Intent diversity)**
- Query Intent-Score**
- Query Intent refinement Similarity**
- Query / Result Intent Similarity**
- Query Intent Frequency**
- Query Frequency
- Suggested Query / Recommended Query
- Number of results

**search|hub specific Signals

CXP commerce experts

Experimental Results: NDCG

Type	Method	Click NDCG@12			Purchase NDCG@12			Revenue NDCG@12		
		Train	Validation	Test	Train	Validation	Test	Train	Validation	Test
Click	RankNet	0,1691	0,1675	0,1336	0,1622	0,1669	0,1626	0,1641	0,1649	0,1315
	RankBoost	0,1858	0,1715	0,1285	0,1856	0,1715	0,1667	0,1858	0,1715	0,1273
	LambdaRank	0,1643	0,1637	0,1319	0,1628	0,1660	0,1624	0,1663	0,1667	0,1325
	LambdaMART	0,2867	0,1724	**0,1370**	0,2867	0,1724	0,1666	0,2867	0,1724	0,1329
Purchase	SVM	0,1731	0,1719	0,1296	0,1776	0,1701	0,1705	0,1762	0,1699	0,1280
	Logistic Regression	0,1919	0,1687	0,1272	0,1919	0,1687	0,1729	0,1919	0,1687	0,1292
	Random Forrest	0,3064	0,1632	0,1323	0,3035	0,2236	0,1744	0,3033	0,1634	0,1335
Both	LambdaMART + RF	0,2661	0,2325	0,1313	0,2800	0,2260	0,1637	0,2661	0,2322	0,1292
	CCM	0,1741	0,1533	0,1340	0,2678	0,1815	**0,1776**	0,2007	0,1676	**0,1478**

+10.7%

better than the best single model

CXP commerce experts

Experimental Results: Revenue/query@k

Type	Method	Rev@1	Rev@2	Rev@3	Rev@4	Rev@5	Rev@6	Rev@7	Rev@8	Rev@9	Rev@10	Rev@11	Rev@12
Click	RankNet	4,16 €	4,36 €	4,55 €	4,57 €	4,71 €	4,86 €	4,85 €	4,96 €	5,08 €	5,16 €	5,17 €	5,20 €
	RankBoost	**4,25 €**	4,36 €	4,36 €	4,43 €	4,62 €	4,81 €	4,86 €	4,98 €	5,11 €	5,18 €	5,25 €	5,28 €
	LambdaRank	4,07 €	4,29 €	4,41 €	4,52 €	4,72 €	4,88 €	5,04 €	5,05 €	5,27 €	5,38 €	5,40 €	5,44 €
	LambdaMART	4,15 €	4,22 €	4,40 €	4,74 €	4,94 €	5,17 €	5,35 €	5,49 €	5,25 €	5,37 €	5,41 €	5,46 €
Purchase	SVM	4,10 €	4,22 €	4,43 €	4,44 €	4,60 €	4,80 €	4,97 €	5,12 €	5,25 €	5,37 €	5,40 €	5,43 €
	Logistic Regression	3,99 €	4,32 €	4,32 €	4,36 €	4,41 €	4,47 €	4,59 €	4,62 €	4,75 €	4,75 €	4,78 €	4,81 €
	Random Forrest	4,20 €	4,48 €	4,52 €	4,67 €	4,82 €	4,96 €	5,12 €	5,26 €	5,38 €	5,51 €	5,57 €	5,62 €
Both	LambdaMART + RF	4,11 €	4,19 €	4,39 €	4,72 €	4,86 €	5,03 €	5,18 €	5,21 €	5,33 €	5,44 €	5,48 €	5,51 €
	CCM	4,19 €	**4,57 €**	**4,73 €**	**5,10 €**	**5,25 €**	**5,45 €**	**5,61 €**	**5,77 €**	**5,96 €**	**6,09 €**	**6,17 €**	**6,24 €**

+11.0%

better than the best single model

CXP commerce experts

Summary

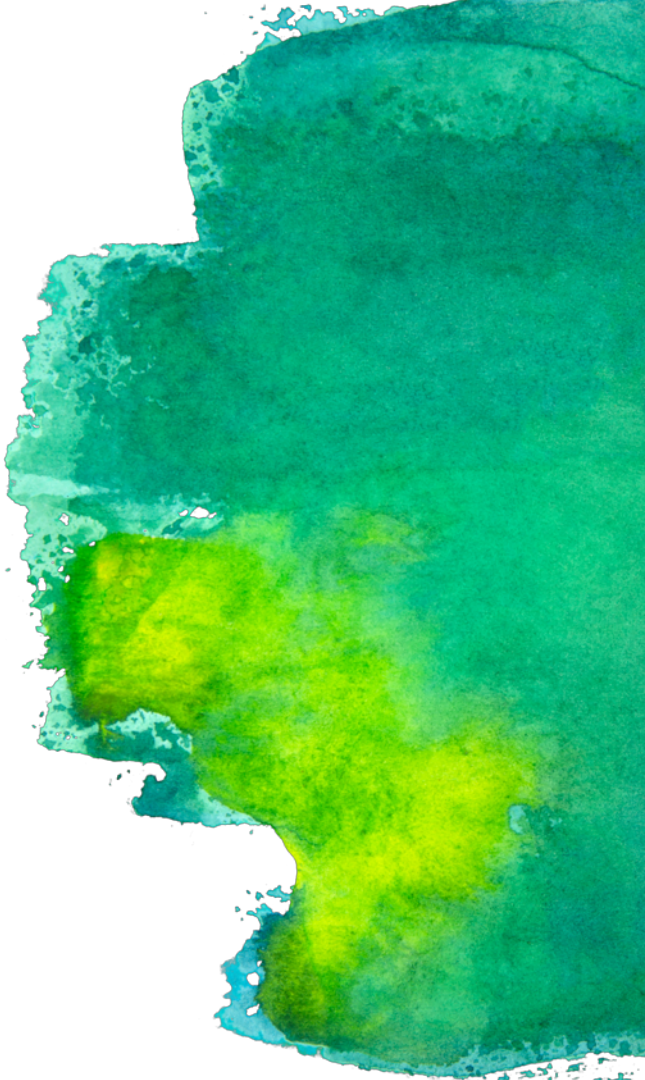
Keep your Tracking clean and handle bias

Query types really matter

- generic vs. precise
- informational vs. inspirational

The Discovery & Buying Process is a complex Journey

Do not oversimplify the problem by using Explicit Feedback for SERP relevance only

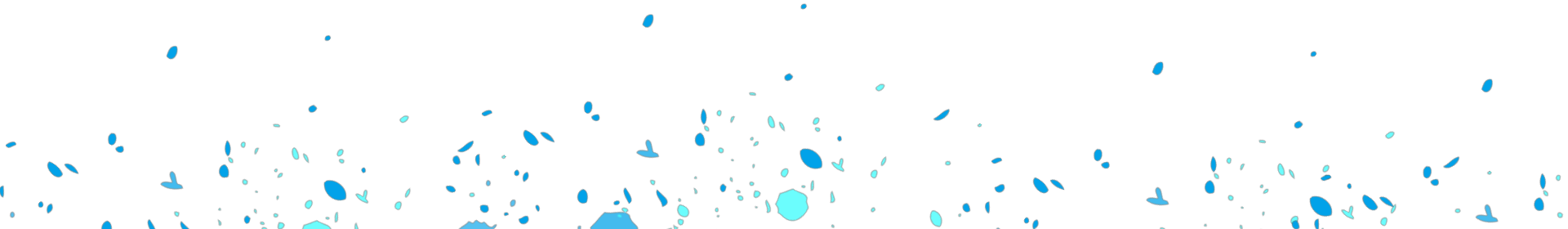


Thanks!

Any questions?

You can find me at:

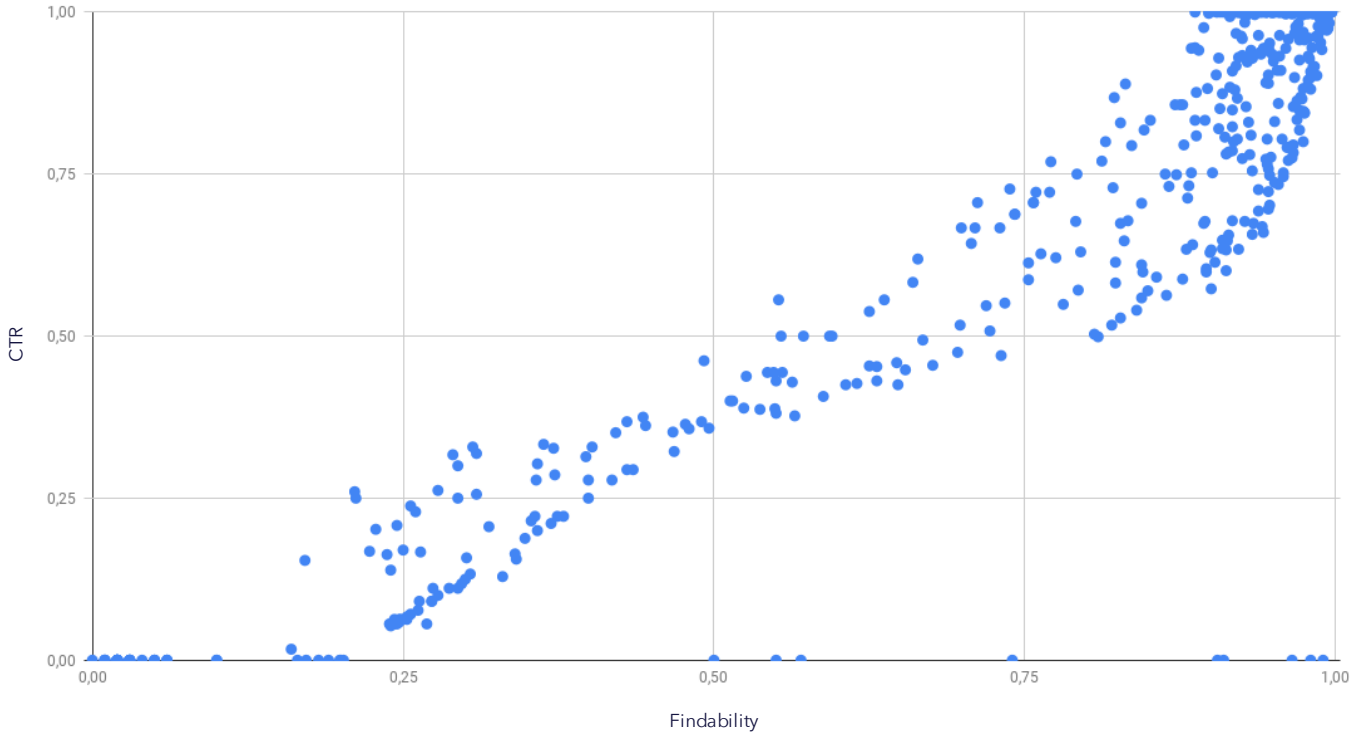
@Andy_wagner1980
andreas.wagner@commerce-experts.com



Backup Slides

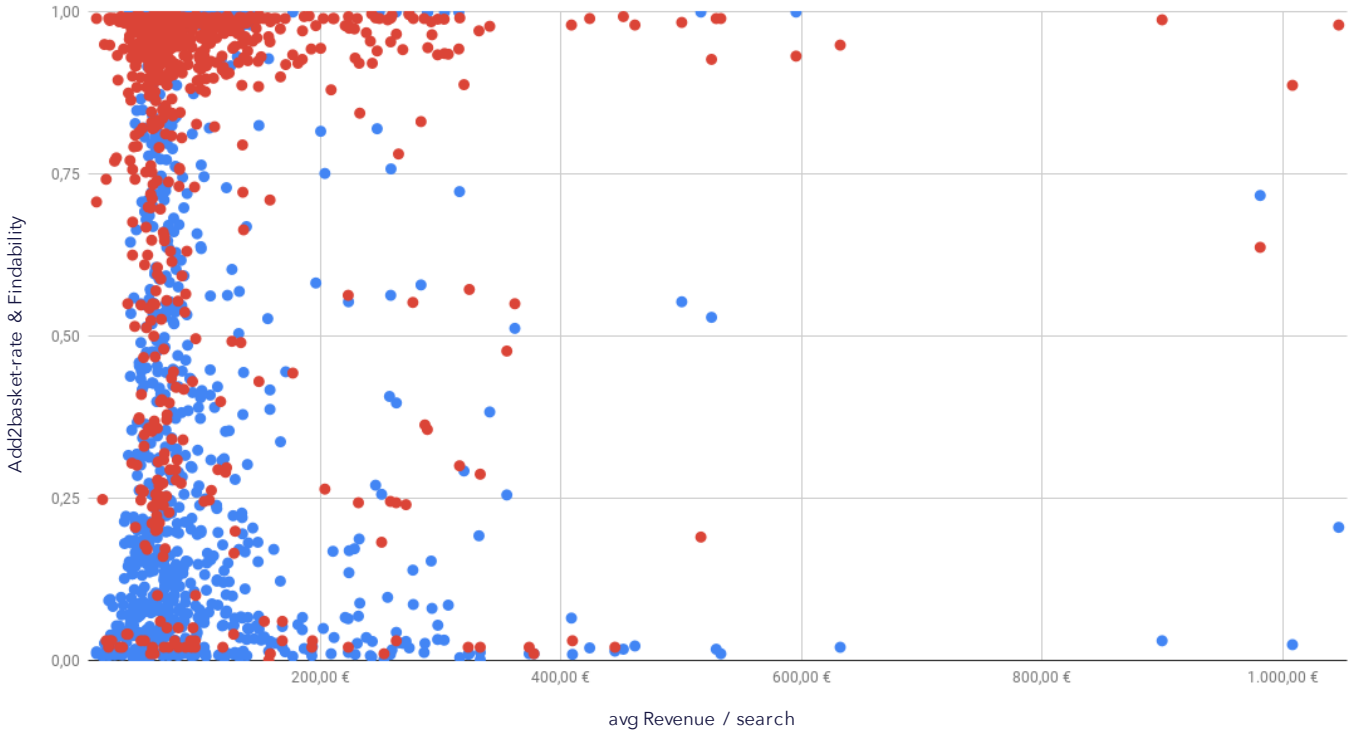
CXP commerce experts

Results – Findability as a Click Predictor



CXP commerce experts

Results – Findability as a add2Basket Predictor



CXP commerce experts

Results – Findability & Sellability as a add2Basket Predictor