

Hi! I know this is a challenging time, and so I appreciate this chance to get to talk to you. This is a big group and VC is hard. Please don't hesitate to unmute and jump in with comments or questions.

Ex. No.
UPX0219
1:20-cv-03010-APM

REDACTED FOR PUBLIC FILING & ABRIDGED

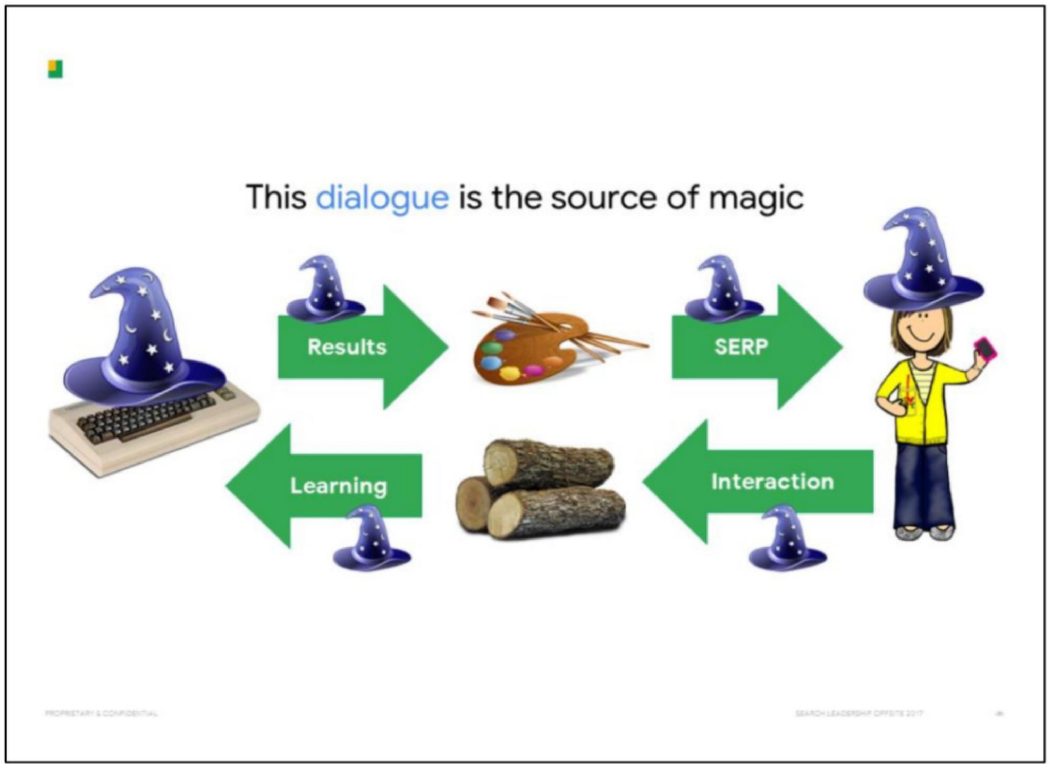
Redacted

GOOG-DOJ-16512397

ranking and search work at a high level-- from my perspective-- and the critical role that logging plays in that story.

Behind all this, I'm selling something. And that something is the idea that even though ranking can seem like a needy, special-case customer, effective logging for ranking is central to Search, to Google as a whole, and beyond.

REDACTED FOR PUBLIC FILING & ABRIDGED



This two-way dialogue is the source of Google's magic.

With every query, we give a some knowledge, and get a little back. Then we give some more, and get a little more back.
These bits add up. After a few hundred billion rounds, we start lookin' pretty smart!

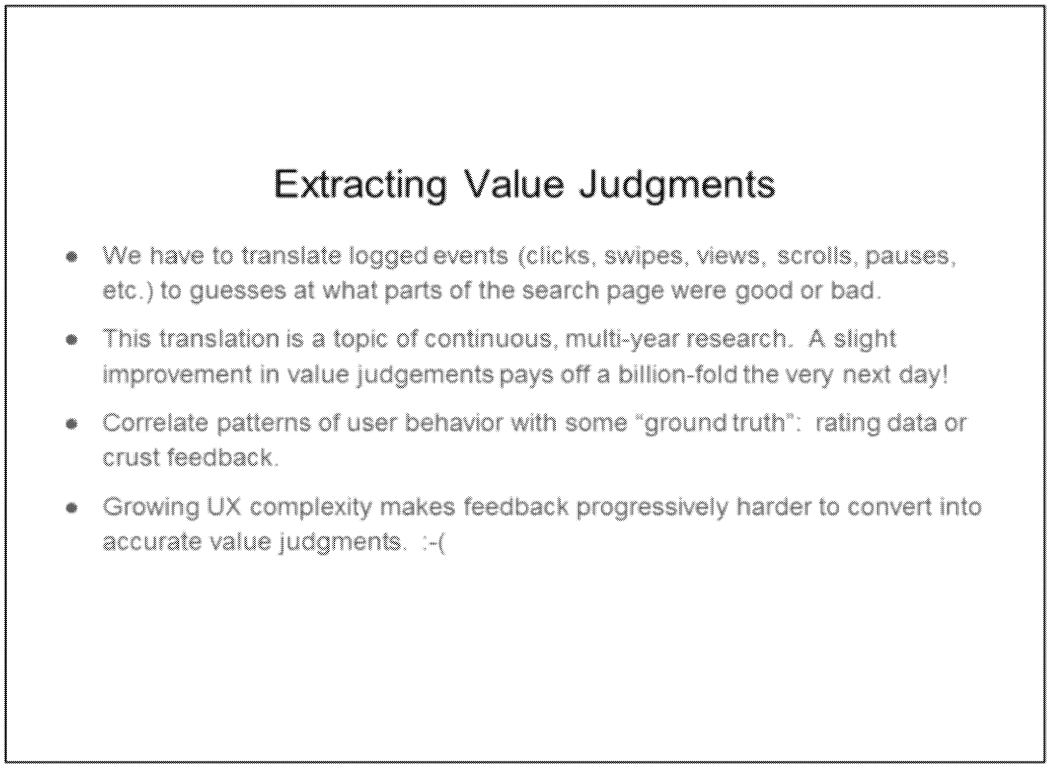
Search is a bit like a potluck, where every person brings one dish of food to share. This a great, big spread of food that everyone can enjoy. But it only works because everyone contributes a little bit.

In a similar way, search is powered by a huge mass of knowledge. But it isn't something we create.
Rather, everyone who comes to search contributes a little bit of knowledge to the system from which everyone can benefit.

REDACTED FOR PUBLIC FILING & ABRIDGED

Redacted

GOOG-DOJ-16512413



The logs do not contain explicit value judgments-- this was a good search result, this was a bad one.
So we have to some how translate the user behaviors that are logged into value judgments.
And this translation is really tricky, a problem that people have worked on pretty steadily for more than 15 years.
People work on it because value judgements are the foundation of Google search.
If we can squeeze a fraction of a bit more meaning out of a session, then we get like a billion times that the very next day.
The basic game is that you start with a small amount of "ground truth" data that says this thing on the search page is good, this is bad, this is better than that.
Then you look at all the associated user behaviors, and say, "Ah, this is what a user does with a good thing! This is what a user does with a bad thing! This is how a user shows preference!"
Of course, people are different and erratic. So all we get is statistical correlations, nothing really reliable.
For example:

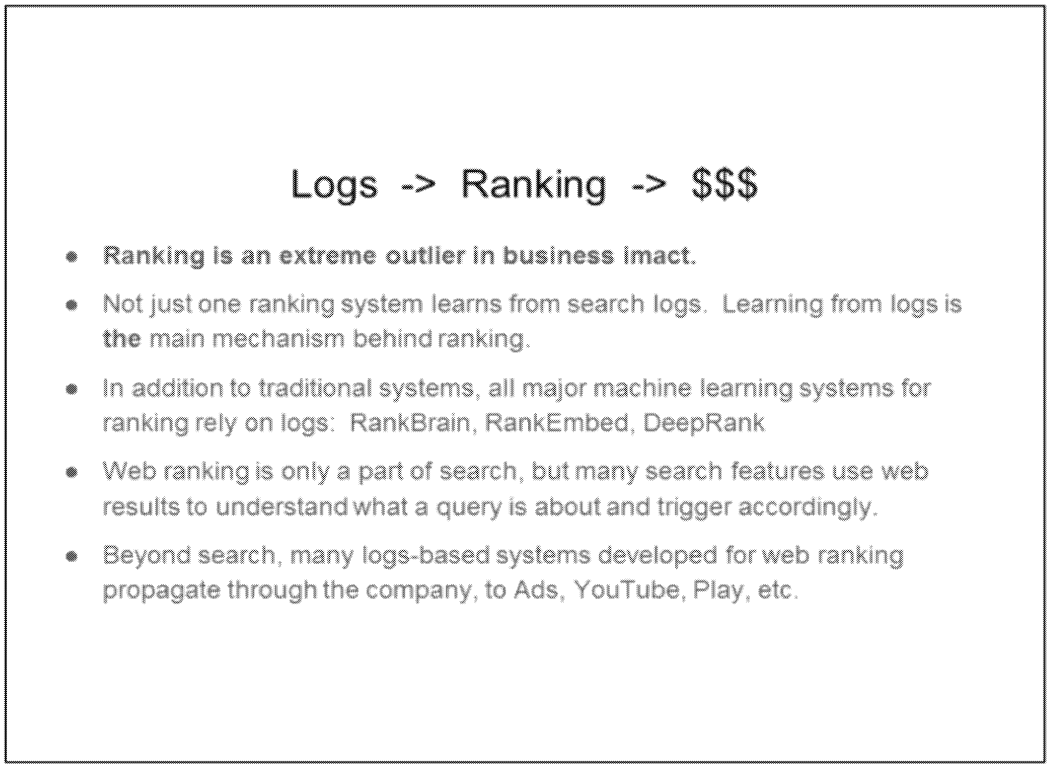
Confidential

- If someone clicks on three search results, which one is bad? Well, likely ALL of them, because it is probably a hard query if they clicked 3 results. Challenge is to figure out which one is most promising.

REDACTED FOR PUBLIC FILING & ABRIDGED

Redacted

GOOG-DOJ-16512417



...and here comes the part I warned you about. I'm selling something. I'm selling the idea of the logs term keeping the needs of the ranking team in mind. Pretty please with sugar on top.
But the basic reason is that the ranking team is really weird in one more way, and that is business impact.
As I mentioned, not one system, but a great many within ranking are built on logs.
This isn't just traditional systems, like the one I showed you earlier, but also the most cutting-edge machine learning systems, many of which we've announced externally-- RankBrain, RankEmbed, and DeepRank.
Web ranking is only a part of search, but many search features use web results to interpret the query and trigger accordingly.
So supporting ranking supports search as a whole.
But even beyong this, technologies developed in search spread out across the company to Ads, YouTube, Play, and elsewhere.
So-- I'm not in finance-- but grossly speaking, I think a huge amount of Google business is tied to the use of logs in ranking.

REDACTED FOR PUBLIC FILING & ABRIDGED

Redacted