

Message

**From:** Alexander Grushetsky [Redacted]@google.com]  
**Sent:** 1/3/2019 7:19:27 AM  
**To:** Matthew Gray [Redacted]@google.com]  
**CC:** Eric Lehman [Redacted]@google.com]; Pandu Navak [Redacted]@google.com]; Bruno Possas [Redacted]@google.com]; Paul Haahr [Redacted]@google.com]; David Bailey [Redacted]@google.com]; Cody Kwok (Chung) [Redacted]@google.com]; Abhinandan Das [Redacted]@google.com]; Berthier Ribeiro-Neto [Redacted]@google.com]; Dimitra Papachristou [Redacted]@google.com]; Hyung Kim [Redacted]@google.com]; Danny Sullivan [Redacted]@google.com]  
**Subject:** Re: thought for the holidays

On Wed, Jan 2, 2019 at 8:25 AM Matthew Gray [Redacted]@google.com> wrote:

On Wed, Dec 26, 2018 at 4:48 PM Eric Lehman [Redacted]@google.com> wrote:  
Hi,

I'd like to offer a thought for contemplation over the break:

**Within the near future, a deep ML system will clearly outperform Google's 20-year accumulation of relevance algorithms for web search.**

Here, I'm just talking about relevance; that is, determining whether a document and query are talking about the same thing. There is a lot more to web ranking for which ML seems much less appropriate. But I think basic relevance is the major task in web ranking and probably "objective" enough to go after pretty effectively with ML.

None of us can see the future, but my bet is that this is nearly certain to be true within 5 years and could be true even within 6 months. One problem after another that is similar in flavor to web ranking has fallen, and there is little reason to think that web ranking is somehow exceptional. Indeed, this holiday thought stems from recent advances in web answers, where deep ML (in the form of BERT) abruptly subsumed essentially all preceding work.

I agree strongly and think the 6 month estimate is much closer than the 5 year estimate. I think in WebAnswers it introduces as massive new set of problems and opportunities that I'm really excited about but are definitely triggering a fearful reaction among many. I'll be sending out my beginning of year "vision letter" to the WebAnswers team this week in which I'm hoping to partially address this and am happy to share.

For the web answers team, the tidal wave of deep ML that arrived in the last few weeks was a complete shock. With this warning, we should not allow ourselves to be caught off-guard again; rather, we should start thinking through the implications now. And \*now\* is really the time, because in the new year I expect a lot of web ranking engineers to reflect on BERT and start thinking along these same lines.

One consideration is that such a deep ML system could well be developed outside of Google-- at Microsoft, Baidu, Yandex, Amazon, Apple, or even a startup. My impression is that the Translate team experienced this. Deep ML reset the translation game; past advantages were sort of wiped out. Fortunately, Google's huge investment in deep ML largely paid off, and we excelled in this new game. Nevertheless, our new ML-based translator was still beaten on benchmarks by a small startup. The risk that Google could similarly be beaten in relevance by another company is highlighted by a startling conclusion from BERT: huge amounts of user

Ex. No.  
UPX0197

1:20-cv-03010-APM

Redacted

REDACTED FOR PUBLIC FILING

GOOG-DOJ-04925211



feedback can be largely replaced by unsupervised learning from raw text. That could have heavy implications for Google.

Relevance in web search may not fall quickly to deep ML, because we rely on memorization systems that are much larger than any current ML model and capture a ton of seemingly-crucial knowledge about language and the world. And there are lots of performance challenges, specialized considerations, etc. Still, my guess is that the advantages of our current approach will eventually crumble; ML is advancing very fast, and traditional techniques are not.

I don't know how other think about this. Maybe this prospect was already obvious to you. Or you might think this view of the future is just wrong. Personally, I'm inclined to think that this future is near-inevitable, but-- despite that-- I hadn't taken the next step of thinking through implications. Some questions to ponder might include:

- **Can we take actions now so that this transition is something that we drive rather than something to which we fall victim?** Personally, I don't want the perception in a few years to be, "Those old school web ranking types just got steamrolled and somehow never saw it comin'..." Could we instead, say, formulate some 2019 collaboration goal with research to beat our best existing **Redacted** prediction with a deep model?

I'm going to come back to something I've been saying a bunch in the past 8 weeks, which is I think we need to deconstruct the various reasons why we often prefer our own "artisanal" ranking systems over an essentially end-to-end deep system (like Ads has) and how we might bring those attributes to deep ML models instead of just thinking about how we bring the quality benefits of deep ML models into our current "understandable" models.

In particular I really think we care about a handful of things:

- **How can we make narrowly focused changes that are both predictable and don't introduce a lot of churn?** **Redacted**

# Redacted

- **How can we anticipate errors and both failure and success modes?** **Redacted**

# Redacted

- **How can we work on the system at scale?** **Redacted**

# Redacted

- **How do we avoid becoming overly-reliant on metrics directly?** **Redacted**

# Redacted

# Redacted

- How might we discuss this possible future with people working on web ranking without crushing morale? Redacted

# Redacted

# Redacted



# Redacted

And we already know, one signal could be more powerful than the whole big system on a given metric. For example, I'm pretty sure that NavBoost alone was / is more positive on clicks (and likely even on precision / utility metrics) by itself than the rest of ranking (BTW, engineers outside of Navboost team used to be also not happy about the power of Navboost, and the fact it was "stealing wins")

# Redacted

I think I heard that the Translate team decided to go "all in" on large-scale ML some years ago, which seems wise in retrospect. I'm skeptical of such an extreme step around relevance today, because we might sacrifice significant gains by traditional means between now and the time a deep-ML approach really takes over-- which I think is probably at least a couple years out. Yet hearing the BERT wake-up call and not adjusting our plans seems unwise as well.

Anyway, this is on my mind during this quieter time, and I thought I'd share.

/Eric