


News Feed

Is Ranking Good?

██████████ · SUNDAY, MAY 6, 2018 · READING TIME: 18 MINUTES 

Probably.

Even asking this question feels slightly blasphemous at Facebook. So many experiments and product launches demonstrate the value of ranking that it's value is often taken as an article of faith. When I proposed a long-term ranking holdout to a colleague not long ago, we had a serious discussion about whether it would be ethical to deprive users of something so valuable for so long.

Redacted for Congress

And yet, after two years studying and working on News Feed, I find myself less confident of the answer than I'd like to be, and maybe less confident than I ever was before. As I prepare to move on to new challenges, I thought it would be worth collecting my thoughts on this in one place, as much to help myself sort through them as for anyone else's benefit.

Acknowledgments:

[REDACTED]
[REDACTED] all provided valuable feedback on earlier drafts. The original impetus for writing this was a long conversation with [REDACTED]. Thank you all.

The case for ranking

To be clear, I'm using the word "ranking" here to mean an ordering of content by **importance or relevance**. Any list of content is inherently ordered by *something*, but I'm distinguishing between these sorts of predictive orderings and other approaches like purely chronological or even alphabetical orders. So, for example, in this nomenclature, entries in a typical dictionary are *ordered* (alphabetically) but not *ranked*. Also, this note largely concerns what we generally call the Facebook "Blue App," although some of the thinking likely applies to ranking in other contexts.

The experimental and experiential data in support of ranking is extensive, and nearly universal. When we switch a random set of users to a pure chronologically News Feed, their usage and engagement immediately drops (see [this recent study](#), for example). When Instagram launched feed ranking, virtually all metrics improved (see [this retrospective on the first year's effects](#)). Whenever we've tried to compare ranked and unranked feeds, ranked feeds just seem better. And this same pattern repeats when we rank Messenger contacts (see [some recent wins here](#)) or the Stories tray (see [these slides](#)). **Ranking items in order of importance or relevance just seems better than chronological, alphabetical, or other more arbitrary orderings.**

Ranking has a long and distinguished history. **Newspapers are ranked, with the most important stories on page one.** (It's difficult to imagine a useful newspaper presented any other way.) **Search results are ranked.** (Imagine a version of Google search where results were displayed chronologically, and how useless this would be.) In most contexts where we're presented a large quantity of information, we tend to order the information by importance (or relevance, which I use here synonymously). **Ranking is such an obvious solution in these cases that we often don't even notice we're doing it:** these days we tend to watch TV in a "ranked" order, starting with what we feel like watching and not whatever happened to come out most recently.

The trouble with ranking

Given the data presented above, many people would consider the question of ranking effectively closed. Part of the challenge in discussing ranking objectively is that the benefits are discrete and measurable, while the costs are often indirect and diffuse and not easily observable in experiments. But I think it is worth enumerating these costs even if we can't necessarily quantify them.

Ranking changes the semantics of friending

In a world without ranking, the consequences of adding a friend on Facebook are very clear: you will now see this friend's posts in your feed. If you don't want to see their posts, you need to unfriend (or at least unfollow) them. The cost of maintaining a connection to that friend are potentially high: you have to put up with their posts, no matter how frequent or boring or even offensive they are. (The first person who explained this to me was [REDACTED])

With a ranked feed, the consequences of adding a friend on Facebook are a lot more murky. You may see their posts, and you may not. You may see all of the posts, or none of them, or anything in between. If you add a friend you don't *really* like, the ranking algorithm should figure this out and stop showing you their posts.

So in this ranked world, the cost of maintaining a connection to a distant friend are not only low; they are roughly zero. With a perfect algorithm, the cost would truly be zero. The only posts from this friend we would ever see are those we care about. This could happen once a day, once a week, once a year, or even just once, period. The better the ranking algorithm, the lower the cost of “bad” friendships. As a result, there is no longer any incentive to clean up your friend graph – friends you don’t care about become irrelevant and inconsequential, without you having to remove them.

Take a look at your friend list on Facebook some time. If you’re anything like me, you’ll be shocked at some of the names. I certainly have “friends” I couldn’t pick out of a lineup and have no recollection of ever friending. And I have many friends I can’t remember seeing a post from pretty much ever. Maybe some of those people never post, but presumably many of them do, and our ranking engine has been quietly doing its job, sticking those posts far enough down in my feed that they never bother me.

For me at least, the contrast with Instagram is striking. I have many fewer friends there. Because Instagram only recently launched a ranked feed, most of my connections were made in an unranked world, where following someone meant seeing every picture they posted. I was (and, to some extent, still am) quite cautious about who I followed. If someone posted too often, or I wasn’t *really* interested in them, I dropped them pretty quickly. Also, Instagram effectively has two popular surfaces (including Explore), while Facebook only has one. This gives users another way of expressing their intent when using the app – when they want “interesting stuff from people I don’t really know,” they can go to a special surface just for that. (██████████ explained this to me.) Facebook has to mix such content into the main feed, attempting to intuit when users might want to see it.

Given a sufficiently good ranking algorithm, the logical thing to do is to friend everybody. This may seem like hyperbole, but it's not. Even someone in some faraway country who speaks only languages I don't understand might snap a photo I'd enjoy seeing, and the perfect ranking algorithm would find that and surface it for me. (And even if they post only text, we have translation systems that could help me understand it.) User research suggests that people often don't realize their feed is ranked, so the incentives here operate largely unconsciously, but the incentives are still there.

But in such a world, why maintain friend lists at all? If the perfect friend list is one that's all-inclusive, why not just save people the trouble and making everyone friends with everyone else by default? By **reducing the cost of friending close to zero, ranking changes the semantics of friending from "I care about you" to "I might conceivably care about something you share someday"** – which is likely true of just about anybody. Friend lists are still useful to the ranking algorithm, as a hint about who you might be interested in, but if people friend too liberally, the list ceases to be a very good hint.

Ranking favors consumption at the expense of production

That last rhetorical question has a clear answer, of course, and that answer revolves around privacy.

While I might be fascinated by that hypothetical snapshot posted by a faraway stranger, that photographer may not *want* me to see their photo. Maybe it's a picture of their house or their kid or some private moment in their life. Maybe it reveals something about their personal identity – their beliefs or customs or sexuality – that they aren't comfortable sharing with the world. Maybe it's an embarrassing moment they don't mind sharing with their family or friends, but wouldn't want their boss to see.

While ranking reduces the cost of friending to near zero from a consumption perspective, it does not necessarily do so from a production perspective. Put another way, **with a well-ranked feed, broad (or even universal) friending adds value to consumption, while adding risk to production.** You get to see that rare gem posted by a distant friend, but distant friends may also see that embarrassing photo that in retrospect maybe you shouldn't have posted at all.

Again, as far as we can measure in experiments, the net result is positive. In some cases, the wide audience potential of having many friends may inspire some people to post more. But **it is hard to believe that the loose social graph encouraged by ranking isn't also suppressing some personal sharing.** (One theory for why sharing through Stories is doing so well is that it doesn't run the risk of appearing "accidentally" in some distant friend's feed.) We certainly hear anecdotally that people post less because they fear too many people seeing their content, but this effect is hard to measure empirically due to all the confounding factors driving friend counts.

Ranking can promote problematic content

Insofar as problematic content is often more engaging than unproblematic content, ranking-by-engagement runs the risk of favoring the problematic.

Imagine a media environment where blatant lies are rare, but very interesting. (I would argue this is roughly the current situation.) Now imagine that a friend of yours in such an environment shares a lie. At first even with a ranked feed, it might get somewhat lost in the sea of content, but eventually someone sees it, and because it's interesting, they like it or comment on it and share it on to their own friend network. Now the ranking algorithm starts to figure out how interesting this lie is (if it didn't work that out from the post content

Chats

and repeats. The post gets liked and shared. Now imagine the same situation, but without ranking. Someone still happens on the post eventually, either because they refreshed their feed right after it was shared or because they scrolled down far enough to find it, and they too like it or comment on it and share it on. But there's no acceleration. Their engagement with the lie doesn't boost its distribution, and the reshared post just gets filed away in people's feeds, no more or less likely to be read than anything else. It might spread, but without ranking, highly-engaging lies have no great advantage over more staid truths.

In this way, an unranked feed can provide a sort of herd immunity against misinformation. The misinformation might still spread, but that spread won't be amplified, and may well die out before it gets too far. (By the same token, an unranked feed might also slow the spread of true information.)

Of course, there are other defenses against misinformation and other problematic content, and we are deploying these today. But without ranking, there wouldn't be the same level the urgency to these efforts we see today.

Also, the fact that ranking encourages us to have very broad friend graphs makes this situation a bit worse still. Imagine you have hundreds of friends who rarely post anything you really want to see. (Again, this is probably your actual situation.) Now one of them reshares this incredibly fascinating and engaging story they also got from a distant friend – which happens to be a hoax. Because this post is so much more interesting than most of what they post, suddenly it jumps to the top of your feed, rather than languishing far below where you'd ever typically scroll. **Not only does ranking allow bad content to spread faster – it allows bad content to spread farther due to the costless accumulation of friends.**

Ranking changes our relationship to users

If a friend sends you an offensive text, you don't blame Apple or AT&T; you blame your friend. If an uncle inundates you with email hoaxes, you don't blame Gmail or Outlook; you blame your uncle.

But **when that misinformation described above spreads on Facebook, people blame Facebook, and not the friends or pages that shared it.**

And sadly this isn't an entirely irrational reaction. With a ranked feed, Facebook decides what content spreads and what content doesn't. We try to do this with neutral principles that don't favor certain viewpoints, but those principles have consequences. As soon as ranking is introduced, we are effectively curating people's feeds. We don't exercise nearly the editorial control of media sites like The News York Times or Wall Street Journal, but we're not Gmail anymore either.

This has tremendous consequences for our relationship to our users that I worry we have not fully absorbed. **A ranked feed is a curated feed, and we are the curators.**

This puts us in a very different relationship to our readers than that of an email provider. We are not a pass-through platform for users to share thoughts. We are aggregators, and this implies a very different set of responsibilities. We are taking these more seriously today than ever before, but I still worry that we haven't come to terms with ranking as the root cause.

Ranking implicitly devalues content

Why don't we rank email? Many of us get far too much and often miss important messages, and yet it is still presented in chronological order. Gmail's Priority Inbox clearly has some sort of ML scoring function for messages, but even they use it only for classifying into more- and less-important, with messages in each of those buckets listed chronologically. Why not go all the way and rank the messages by importance?

I think the answer is that with ranked email, we'd miss even more.

The working assumption with email is that you can't miss anything important (even though many of us do), which is why most of us still dutifully walk through our inboxes message-by-message, multiple times a day. The underlying assumption of a ranked feed is that almost everything is optional, that you only have time for the best of the best. I think this is important enough to be worth repeating: the **underlying assumption of a ranked feed is that almost everything is optional.**

But is everything actually optional? I've been in many meetings where executives at the highest levels of Facebook ask us why they don't see all their closest friends' posts in their feeds. This is not a hard question to answer. You don't see all your closest friends' content in your feed because we have other content we think you'd like better, and we consider your friends' content to be optional.

Chats

Redacted for Congress

When these people ask us this question, though, they already know the answer. They just don't like it – because those posts don't feel optional to them. At a deep level, **ranking devalues content, by which I mean commodifies it into an endless stream of optional stories we can read or not depending on whether anything better comes along.** And that feels wrong, not all the time, but much of the time.

Ranking requires private data collection

This may be an obvious point, but **a chronological feed requires no maintenance of private data or usage history beyond posts and lists of friends.** We don't need to know what you've liked or commented on in the past in order to present an unranked feed to users. We don't need to know almost anything about our users at all

Of course, we might still need such data for advertising, and so some would argue that we might as well use it for feed ranking. But *not* using such private data for feed ranking would clarify the user tradeoffs and allow for a stronger “opt out” option, since users who opted out of ads targeting would not need to have this data stored for them at all. (Today, we still need to collect this data even for users who opt out of ad targeting because we need it for ranking.)

Can Ranking be Fixed?

Facebook suffers from the curse of a great product, such that almost everything we try that's radically different turns out worse. But it's important to remember that this is to be expected. We've spent a decade perfecting the ranked News Feed, investing person-centuries of time in getting the details right. It would be very weird if a few people working for a few months or a half could make measurably something better – and yet that is often the bar we hold ourselves to when trying alternatives. We've also trained two billion users (at least) that News Feed works this way. Old habits are hard to break, even if we do build something significantly better.

And the benefits of ranking are undeniable, and probably deserve more space than I've given them. **If we abruptly stopped ranking News Feed tomorrow, the results would be disastrous for the company by most metrics we care about.** So is there a way to keep the undeniable benefits of ranking without incurring all the costs?

Stories and other hybrids

It's possible that we could find a hybrid approach. Perhaps content from your true friends should be presented in full, always, without any real intermediation from Facebook. Ranking could then be reserved for the next tiers of content, from people you may be vaguely interested in or want to follow more loosely, as well as from professional publishers (ie, Pages). Or perhaps the distinction is not strictly by content author, but also content type – maybe it's only truly personal posts from your true friends that deserve unranked treatment, while the links and reshares from them should go into the broader hopper with everything else, to be surfaced only when they're particularly interesting.

If you step back, this hybrid model is a lot like where we are with Stories. Intimate, personal posts from our friends are presented to us verbatim, unranked, while everything else appears below in a ranked order. Of course, Stories makes lots of other changes to our sharing model, too, perhaps most notably trading ephemerality for permanence. It's still unclear which of these changes are essential to the success of that format and which, if any, are not.

Multiple feeds and partial ranking

There are other approaches. Something analogous to Gmail model might have multiple feeds, each unranked in itself, but with posts sorted between feeds based on importance – a close friend feed, perhaps, or a priority feed, followed by feeds for less important content. We've tried things like this before on Facebook and failed, but it's always difficult to know if the failure is due to deep-seated habits or a truly worse product experience. Instagram's Explore has been much more successful, suggesting this approach can theoretically work.

One can also think of recency or freshness as another signal that gets weighted along with everything else. A pure chronological feed (what I'm calling unranked) would put all the weight on recency and nothing else. But of course, any weight is possible. Perhaps a feed with a very large weight on recency could result in a sort of partial ranking, that looked largely chronological but with other very interesting stories popping up out of order. The weighting of recency could in some way correlate with closeness, such that you got a chronological feed of content from your closest friends, but with a sprinkling of content from other sources mixed in occasionally, or provided after all close friend content had been consumed. There are proposals to show unconnected recommendations at the “end of feed,” and this could be similar to that: first you'd see a chronological feed of close friend content, then a ranked feed from other sources, then finally some unconnected recommendations if you still wanted more. [REDACTED] tells me Twitter does something similar to this.)

As a side point, aggressive ranking can actually penalize recency. Imagine two equally great posts, one from a second ago and one from a few hours ago. The super recent one hasn't had time to prove its greatness through engagement, so will likely get ranked below the slightly older post. Placing a higher value on recency might level the playing field here. (This point was also explained top me by [REDACTED])

Redacted for Congress

Could we ease out of ranking?

When we give users an unranked News Feed, their engagement lags that of new users with ranked Feeds, but not by nearly as much as for established users. (The results are about half as large.) This suggests that at least to some extent, **users have adapted their behavior to ranking, and that new users might simply use the product differently without ranking.** They might still use the product somewhat less, but maybe not as much less as we think. (It's also possible that new users simply haven't had time to build up big friend lists, and so don't yet experience the full benefits of ranking.) Another approach would be to gradually rank less aggressively, giving people more time to adjust their friending behavior.

Can we just rank better?

Some of the down sides of ranking could be seen as simple bugs. All engagement is not equally useful as a signal of quality – perhaps some of the integrity issues we see around ranking could be solved with “cleaner” signals. Can we design ranking algorithms that are robust to problematic engagement? Can we find better ways to distinguish quality engagement from insidious engagement? There is already on-going work on News Feed to uncover better signals, but right now this is treated as a separate problem from the core ranking algorithm design. Perhaps it shouldn't be.

Conclusions

It's possible that the benefits of ranking are simply worth the costs. Clearly we should continue the work on News Feed Integrity and elsewhere on fixing whatever problems it causes, making ranking incrementally better with every launch. We have tried various things on Facebook like Explore Feed that might allow us to move away from our dependence on ranking, and these have not worked well. Perhaps content and user intention on Facebook are sufficiently heterogenous that ranking is truly essential.

And yet. We have given up a lot in the shift to ranking, and not always as a conscious and deliberate choice. Many choices we make in life without fully thinking them through turn out to be mistakes – or at least turn out worse than they might have had we considered the consequences. If a full cost-benefit analysis still favors a full ranking of News Feed and other surfaces, then we should certainly continue. But doing that analysis requires fully acknowledging these somewhat ephemeral costs and trying to find some way of balancing them against the benefits. **The only way to accomplish our mission is always to be willing to question our choices, and always be looking for ways to make better choices in the future.**



Like



Comment



Share



Save

Great note! Just to add to your examples, one example of when ranking is worse than chrono-sort, is when the ranking problem is inherently really hard. E.g. Gmail search results are ranked in reverse-chrono sort. With just the query as input, it is really hard for gmail to correctly guess which email you're looking for. With reverse-chrono, at least the user can scroll deterministically and get to the desired result. Ranking would make the experience way more confusing, given gmail will get the ranking wrong more often than not.



4

Like · Reply · 3y

Thanks for writing--this was very insightful for me



1

Like · Reply · 3y · Edited

I wonder how many people end up using See First as a way to signal who their actual real friends are for feed ranking purposes.



2

Like · Reply · 3y

Thanks for a very interesting and thought provoking read.

I think, however, that you have missed an important question. **What is the definition of 'good' in this context? I certainly do not know, but it's not actually clear to me that this claim is true:

> Part of the challenge in discussing ranking objectively is that the benefits are discrete and measurable, while the costs are often indirect and diffuse and not easily observable in experiments.

For this to be true, we need an objective measure of 'goodness,' which I'm unclear on. Is goodness even perfectly quantifiable?

- In your example of newspapers, for example, the goodness of ranking is determined by editorial judgement of what is most informative and educational to readers. And different editorial staffs may have different rankings of the same content.

- With feed, you state that 'with ranking, all usage and engagement metrics go up.' Does this mean that "goodness" for Facebook is engagement? Those metrics are certainly discrete and measurable, but do they measure 'good-ness?' If so, is goodness simply whatever each user finds most addictive?

Interested to hear your thoughts. Thanks again for the thought provoking read.



1

Like · Reply · 3y

Redacted for Congress

By "good," I mean that the benefits outweigh the costs. In the case of newspapers, I don't think the quality of the ranking of stories is in how well it follows the editorial opinions of the staff, which it effectively does by definition; the quality is in how well it informs readers. Good newspaper editors are good at ranking stories, and this makes those newspapers better for readers. In my note, I discuss several ways in which ranking makes the News Feed both better and worse. I would say it's good if the net effect is a better feed for our users.



1

Like · Reply · 3y

Yes I agree with that. I'm wondering what "better feed for our users" means. What defines "better"?



1

Like · Reply · 3y

That's a deep question. I don't think there's a short answer, which is part of what makes this all so hard. We don't have simple objective function we can maximize.



1

Like · Reply · 3y

Write a reply...



This is a fascinating note... thank you for sharing!

You write that, "with a perfect [feed-ranking] algorithm, the cost [of a friendship] would truly be zero." I totally agree with your reasoning that ranking changes the semantics of friending and generally makes people's mental models of "friendship" on Facebook looser. However, is it really true that the cost of a friendship would drop to approximately zero with perfect feed ranking?

I can think of a couple of reasons why that might not be the case:

- The first reason is actually related to your point that "ranking favors consumption at the expense of production": by agreeing to a Facebook friendship, I am taking on the risk that my content will be exposed to the person whom I've friended. It seems that this risk would remain even when content is perfectly ranked by interest to the viewer. Of course, we can think about another ranking problem: for a given producer, piece of content pair, rank that person's friends by how much the creator wants the friend to see the piece of content and cut off that list at exactly the right point. If we accomplished that task perfectly also, then the cost of adding a friend would drop even closer to zero; but another negative consequence of having loose friend lists would also disappear.

- The second reason is that there can be costs / benefits to friending that don't have to do with content distribution at all. By making or accepting a friend request, I can signal a degree of closeness that might affect interactions in the offline world.

I've been working on metrics related to community-building and meaningfulness in Groups. The notion that Facebook friendship loses meaning in the limit of perfect ranking has come up in the context of that work, so this part of your note really intrigued me. Would love to hear your thoughts. Thanks!

Chats

Redacted for Congress

Like · Reply · 3y

[REDACTED] I should add that my first bullet point only makes sense if "perfect ranking for the viewer" means that we perfectly rank all content from that viewer's friends, not all content in the ecosystem (i.e., a universal-implicit-friending situation). In the universal-implicit-friending case, I take on no additional risk from making an explicit Facebook friend because that person could have seen my content irrespective of the existence of that friend tie.



Like · Reply · 3y

[REDACTED] Yes, you're right. It's just the "consumption cost" that goes to zero, but there could be (or are) other costs. In my experience, the consumption costs are the largest factors, but other users might feel differently.

Like · Reply · 3y

[REDACTED] I'd think for many users it would be the privacy cost of this random person now being able to see all the stuff you've shared to your friends that dominates rather than the consumption cost.



Like · Reply · 3y

Write a reply...



Cc Fidji Simo

Like · Reply · 2y

[REDACTED] For me the following point is very interesting and I didn't realize that before: "Not only does ranking allow bad content to spread faster – it allows bad content to spread farther due to the costless accumulation of friends."



Like · Reply · 2y

[REDACTED] Does unranked feed reduce the rate at which some engaging lie spreads or does it just free us of the responsibility? E.g WhatsApp seems to be pretty good at spreading lies – which is unranked.

Like · Reply · 2y

Redacted for Congress

Like · Reply · 2y

There's no experimental data on this that I know of, but ranking is designed to increase the distribution of engaging stories and misinformation is designed to be engaging. This is not to say that unranked platforms can't spread misinformation, which obviously happens routinely, as well.



1

Like · Reply · 2y

(TM) #sharebot Watch Core Ranking

Like · Reply · 2y

Thank you so much for writing this. I wanted to like every single conclusion you had in each section, it really resonated with me.

One thing I want to add, which speaks to your last point about "Can we just rank better?": We can only optimize ranking...
[See More](#)



2

Like · Reply · 2y

So many interesting points here, thank you! One thought re "Ranking implicitly devalues content" and the idea that ranking makes content optional: I'd posit instead that the size of our follow/friend lists makes content optional (let's ignore for a moment unconnected content, which is rarely shown in feed). Ranking, no ranking, I'll never see it all!

This feels to me like the flaw in the email analogy- I receive a relatively small number of non-spam emails per day and so I can read them all, no matter the order. This may not hold for everyone, but I think it's fairly universal that a very small % of the people in our email contact lists send us email on a given day (assuming reasonably good spam filters).

But given the size of most of our friend lists and the breadth of story types (this argument may fail if we remove edge-like stories and limit to actual posts which are rarer), the typical user will never have time to consume EVERY story related to anyone they've ever friended. It is wrong for me to imagine that in an unranked feed I'd avoid missing content from my close friends. In an unranked feed, the content I miss is effectively random (it's whatever happened to occur furthest from when I opened the app)- if I only have a few minutes to scroll and tons of recent stories from my hundreds of FB friends, I'll never have another chance to see any close friend content posted more than a few hours ago. At least in a ranked feed the content I miss is determined intentionally.

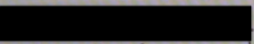
Like · Reply · 2y · Edited



1

Redacted for Congress

Like · Reply · 2y · Edited



You make a great point. I would turn this around, though: In a world without ranking, we would follow fewer people. Whether this is good or bad is certainly open to debate, but I think the inevitable outcome is that we would follow only people whose posts we truly wanted to see, rather than the very broad set of loose connections we tend to make now. Instagram used to be like this for me before ranking: I followed a small number of people, and dropped people if they posted too much and so got in the way of other posts I wanted to see.



1

Like · Reply · 2y



Write a reply...



Write a comment...



Chats

Redacted for Congress