Dark Social: We Have the Whole History of the Web Wrong

By Alexis Madrigal

Here's a pocket history of the web, according to many people. In the early days, the web was just pages of information linked to each other. Then along came web crawlers that helped you find what you wanted among all that information. Some time around 2003 or maybe 2004, the social web really kicked into gear, and thereafter the web's users began to connect with each other more and more often. Hence Web 2.0, Wikipedia, MySpace, Facebook, Twitter, etc. I'm not strawmanning here. This is the dominant history of the web as seen, for example, in this Wikipedia entry on the 'Social Web.'

But it's never felt quite right to me. For one, I spent most of the 90s as a teenager in rural Washington and my web was highly, highly social. We had instant messenger and chat rooms and ICQ and USENET forums and email. My whole Internet life involved sharing links with local and Internet friends. How was I supposed to believe that somehow Friendster and Facebook created a social web out of what was previously a lonely journey in cyberspace when I knew that this has not been my experience? True, my web social life used tools that ran parallel to, not on, the web, but it existed nonetheless.

To be honest, this was a very difficult thing to measure. One dirty secret of web analytics is that the information we get is limited. If you want to see how someone came to your site, it's usually pretty easy. When you follow a link from Facebook to The Atlantic, a little piece of metadata hitches a ride that tells our servers, "Yo, I'm here from Facebook.com." We can then aggregate those numbers and say, "Whoa, a million people came here from Facebook last month," or whatever.

> ## tl;dr version
>
> 1. The sharing you see on sites like Facebook and Twitter is the tip of the 'social' iceberg. We are impressed by its scale because it's easy to measure.
>
> 2. But most sharing is done via **dark social** means like email and IM that are difficult to measure.
>
> 3. According to new data on many media sites, 69% of social referrals came from dark social. 20% came from Facebook.
>
> 4. Facebook and Twitter do shift the paradigm from private sharing to public publishing. They structure, archive, and monetize your publications.

There are circumstances, however, when there is no referrer data. You show up at our doorstep and we have no idea how you got here. The main situations in which this happens are email programs, instant messages, some mobile applications*, and whenever someone is moving from a secure site ("https://mail.google.com/blahblahblah") to a non-secure site (http://www.theatlantic.com).

This means that this vast trove of social traffic is essentially invisible to most analytics programs. I call it **DARK SOCIAL**. It shows up variously in programs as "direct" or "typed/bookmarked" traffic, which implies to many site owners that you actually have a bookmark or typed in www.theatlantic.com into your browser. But that's not actually what's happening a lot of the time. Most of the time, someone Gchatted someone a link, or it came in on a big email distribution list, or your dad sent it to you.

Nonetheless, the idea that "social networks" and "social media" sites created a social web is pervasive. Everyone behaves as if the traffic your stories receive from the social networks (Facebook, Reddit, Twitter, StumbleUpon) is the same as all of your social traffic. I began to wonder if I was wrong. Or at least that what I had experienced was a niche phenomenon and most people's web time was not filled with Gchatted and emailed links. I began to think that perhaps Facebook and Twitter has dramatically expanded the volume of -- at the very least -- linksharing that takes place.

Everyone else had data to back them up. I had my experience as a teenage nerd in the 1990s. I was not about to shake social media marketing firms with my tales of ICQ friends and the analogy of dark social to dark energy. ("You can't see it, dude, but it's what keeps the universe expanding. No dark social, no Internet universe, man! Just a big crunch.")

And then one day, we had a meeting with the real-time web analytics firm, Chartbeat. Like many media nerds, I love Chartbeat. It lets you know exactly what's happening with your stories, most especially where your readers are coming from. Recently, they made an accounting change that they showed to us. They took visitors who showed up without referrer data and split them into two categories. The first was people who were going to a homepage (theatlantic.com) or a subject landing page (theatlantic.com/politics). The second were people going to any other page, that is to say, all of our articles. These people, they figured, were following some sort of link because no one actually types "http://www.theatlantic.com/technology/archive/2012/10/atlast-the-gargantuan-telescope-designed-to-find-life-on-other-planets/263409/." They started counting these people as what they call direct social.

The second I saw this measure, my heart actually leapt (yes, I am that much of a data nerd). This was it! They'd found a way to quantify dark social, even if they'd given it a lamer name!

On the first day I saw it, this is how big of an impact dark social was having on The Atlantic.

Dark social — 56.5%
Facebook — 21.6%
Twitter — 11.2%
Reddit
YComb
Digg
Stumble
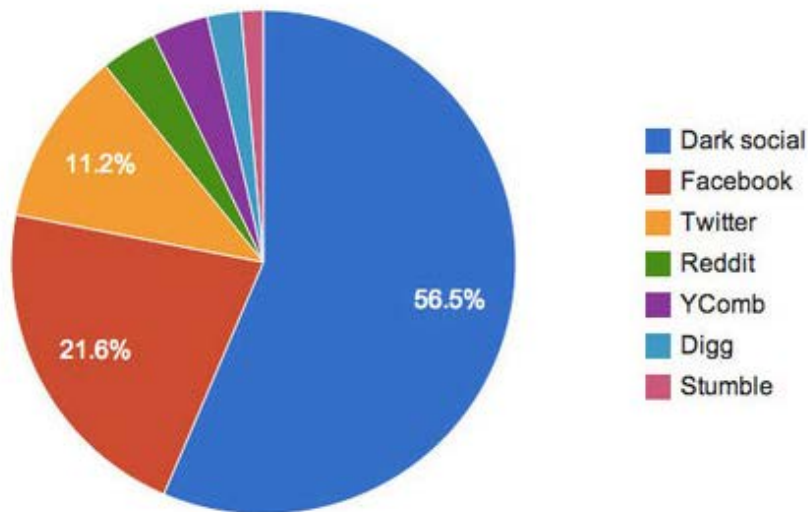
Just look at that graph. On the one hand, you have all the social networks that you know. They're about 43.5 percent of our social traffic. On the other, you have this previously unmeasured darknet that's delivering 56.5 percent of people to individual stories. This is not a niche phenomenon! It's more than 2.5x Facebook's impact on the site.

Day after day, this continues to be true, though the individual numbers vary a lot, say, during a Reddit spike or if one of our stories gets sent out on a very big email list or what have you. Day after day, though, dark social is nearly always our top referral source.

Perhaps, though, it was only The Atlantic for whatever reason. We do really well in the social world, so maybe we were outliers. So, I went back to Chartbeat and asked them to run aggregate numbers across their media sites.

Get this. Dark social is even more important across this broader set of sites. Almost 69 percent of social referrals were dark! Facebook came in second at 20 percent. Twitter was down at 6 percent.

All in all, direct/dark social was 17.5 percent of total referrals; only search at 21.5 percent drove more visitors to this basket of sites. (FWIW, at The Atlantic, social referrers far outstrip search. I'd guess the same is true at all the more magaziney sites.)

There are a couple of really interesting ramifications of this data. First, on the operational side, if you think optimizing your Facebook page and Tweets is "optimizing for social," you're only halfway (or maybe 30 percent) correct. The only real way to optimize for social spread is *in the nature of the content itself.* There's no way to game email or people's instant messages. There's no power users you can contact. There's no algorithms to understand. This is pure social, uncut.

Second, the social sites that arrived in the 2000s did not create the social web, but they did structure it. This is really, really significant. In large part, they made sharing on the Internet an act of publishing (!), with all the attendant changes that come with that switch. Publishing social interactions makes them more visible, searchable, and adds a lot of metadata to your simple link or photo post. There are some great things about this, but social networks also give a novel, permanent identity to your online persona. Your taste can be monetized, by you or (much more likely) the service

itself.

Third, I think there are some philosophical changes that we should consider in light of this new data. While it's true that sharing came to the web's technical infrastructure in the 2000s, the behaviors that we're now all familiar with on the large social networks was present long before they existed, and persists despite Facebook's eight years on the web. The history of the web, as we generally conceive it, needs to consider technologies that were outside the technical envelope of "webness." People layered communication technologies easily and built functioning social networks with most of the capabilities of the web 2.0 sites in semi-private and without the structure of the current sites.

If what I'm saying is true, then the tradeoffs we make on social networks is not the one that we're told we're making. We're not giving our personal data in exchange for the ability to share links with friends. Massive numbers of people -- a larger set than exists on any social network -- already do that outside the social networks. Rather, we're exchanging our personal data in exchange for the ability to publish and archive a record of our sharing. That may be a transaction you want to make, but it might not be the one you've been told you made.

* Chartbeat datawiz Josh Schwartz said it was unlikely that the mobile referral data was throwing off our numbers here. "Only about four percent of total traffic is on mobile at all, so, at least as a percentage of total referrals, app referrals must be a tiny percentage," Schwartz wrote to me in an email. "To put some more context there, only 0.3 percent of total traffic has the Facebook mobile site as a referrer and less than 0.1 percent has the Facebook mobile app."

This article available online at:

http://www.theatlantic.com/technology/archive/2012/10/dark-social-we-have-the-whole-history-of-the-web-wrong/263523/