as part of news or blog content. However, a manual inspection of 100 random news and blog mentions in first bursts yielded 1 official press release. Future work might more systematically characterize content per platform and its relationship to information retention.

Differences in platforms Our findings indeed highlighted differences in information retention across platforms. Blogs performed significantly better than any other platform, with a median information retention score twice the size of the second best-performing platform, news. Some of the patterns we saw were partly explained by the typical length of content in the platforms: platforms with longer mentions intuitively have higher information retention. For example, mentions on news and blogs are longer than posts on Twitter and Facebook and therefore have higher retention scores. However, differences in platforms do not perfectly correlate with content length: blogs, for example, had a median score twice that of news despite the fact that the median length of news mentions is slightly larger than blogs. Moreover, we also saw that bursts being on certain platforms had little effect on the information retention of adjacent bursts and long-term information retention.

However, we also note that the tendency towards more information loss over time was substantially less so for sequences starting in Facebook and Twitter, which were relatively flat over time. A closer look at earlier and later bursts on Twitter did not yield substantive differences in terms of mention length or score. The flatter trajectory of Facebook and Twitter bring support to the explanation that lower information retention later on in sequences are driven by platform differences; that is, that the lower-fidelity nature of social media content may overall drag down information retention over time as social media comes to dominate conversation. Future work should test how content between platforms may shape one another, such as testing how introducing higher fidelity content on one platform might affect the qualities of content on another.

Encouraging multi-platform discussions The differences in platforms, particularly with the flatter trajectories of Facebook and Twitter, emphasize that platforms matter. For example, we found that sequences with more platforms tended to have higher information retention scores, at all sequence sizes. While these patterns reveal no causality, they do highlight that cross-platform online discussions of science may be important for richer, more contextualized discussions of research articles. One possible explanation may be that when multiple platforms are active at the same time, there is a greater diversity of actors, content, media coverage, and information sources for people to interact with, which could improve the collective ability of individuals to parse important information about scientific findings. Encouraging multi-platform discussions of science may therefore promote more accurate and reliable engagement with scientific information. For sequences that start on social media, for example, encouraging multi-platform discussions in later sequences may instead improve information retention. On the other hand, being on multiple platforms is a marker of successful spread; articles mentioned on multiple venues may have been written more appealingly and clearly, making it easier to distill a core message and improving information retention online overall. These point to at least two potential strategies for improving information retention: first, encouraging active and recurrent dissemination by researchers across outlets beyond social media; and second, promoting the sharing of research findings on *multiple* platforms *concurrently* to diversify the collective voices engaged in online discussion of the work.

Limitations and future work

An important limitation of our work is that information retention as a construct is difficult to quantify, subject to human sense-making and judgement. However, as our measure is simple, it is also much more scrutable than complex approaches—important when interpreting results given the subjective and potentially biased nature of assessing information retention. Moreover, our measure is able to capture a reasonable estimate of how *key* information is retained which is essential for large-scale analyses as reflected in our validation survey. Our measure performs much better than a random baseline, and does extremely well when human experts also agree in their assessments. Moreover, results were consistent when changing the keyphrase extraction method. Regardless, future work might develop more refined measures for information retention, such as also capturing synonyms of keyphrases.

Further, we focus on research papers that are among the most mentioned online in two years, which is nonrepresentative of "typical" papers. Regardless, our sample covers a broad set of research topics, a wide range of mention counts, and an important test set for developing frameworks to understand information retention in diffusion, as they represent, by definition, the papers most likely to circulate widely and have impact. Future work might investigate "typical" trajectories in sharing science online (similar to patterns of structural virality in **?**) or, alternatively, patterns of information retention in other types of "bursty" online content (**???**).

We also note limitations of the dataset and framework. Our data consisted of direct DOI mentions of research articles, which excludes comments or other related content that may be part of the discussion of an article. This has implications in particular with respect to Twitter "threads" used to get around character limits[6]. However, in our dataset the median character length of Twitter mentions was 107, far below the character limit of tweets, suggesting that mentions are typically not threads. Indeed, examining a random sample of 100 tweets in our data yielded only 3 that were in threads of any sort. Moreover, our method is consistent over time and across platforms (e.g., we do not collect replies or similar threads of Facebook, news, or blogs content).

Further, we use a simple keyword-based approach over other techniques such as topic modeling, which might cap-

[6]In 2020, the Twitter API added a "conversation ID" feature that enables researchers to pull all tweets that are in response to a given tweet; future work might extract threads summarizing research findings from these conversations.

ture more context around the diffusion process and help identify the *types* of discussions associated with high information retention. However, topic modeling at scale also poses challenges such as the need for manual inspection to label topics, especially with so many scientific fields. Such methods also tend to be noisy when applied to short documents like social media posts. For this study, we focus—as an initial step—on overall cross-platform information retention independent of the details of the surrounding context in which papers are mentioned. A natural next step for future work building on our contributions includes examining such context and its relationship with information retention.

Finally, we do not make empirical claims to causality. While we identify clear differences across platforms and trends across them, we cannot say, for example, that an increase in the number of platforms involved in a sequence *leads to* higher information retention. In part, this is because our sequences are not necessarily information cascades transferring information from one to another, but simply temporally ordered. However, this study does not set out to disentangle particular effects, such as the "telephone effect" (**?**), but instead provides an overview of patterns over time in a multi-platform information landscape. Exciting future research directions include parsing out the mechanisms by which platforms affect information retention. One potential such project includes a systematic analysis of the kinds of keywords found in mentions at various points in a sequence, which may not only articulate particular platform effects but also elucidate why the presence of multiple platforms is linked with higher information fidelity overall.

Ethics Statement

A serious consideration for the presentation of any quantitative measure is how it may incorporate biases or black-boxes of complex human concepts. As a difficult construct to quantify, information retention is a par excellence example of this, as noted in our discussion of the validation survey. In order to prevent our proposed quantification from becoming misapplied, we selected the most scrutable version of it, such that interpretation and limitations of what the measure suggests are clear. This simplicity also helps undermine unreasonable extrapolations in arguments about the *quality of information retention* of content when applied beyond this work. Additionally, research utilizing posts from sites containing content from individuals, such as social media and/or blog users, who may not be aware that their content is used for research also encounter important concerns about the integrity of users' privacy. In our work, we consider only public-facing content accessible by any arbitrary individual. In addition, we only use posts that were still available and not deleted by the users at the time of our study. Finally, our results describe user behavior only in aggregate and at a scale that leaves individual posts unidentifiable.

While minimizing potential risks, the expected benefits of our contributions to the understanding of information retention in the diffusion of scientific findings across platforms are substantial and foundational for a better appreciation of intentional and unintentional information distortion online. Our findings not only point to general patterns of information retention that might inform media strategies, such as deliberate dissemination across multiple channels, but also are generative in raising potential mechanisms for improving the fidelity of information in online discussions to be tested and examined for causality in future studies.

Conclusion

As scientific findings spread on the web, they are discussed across multiple platforms, shaping what information is retained. Accurately communicating science online has critical implications for policymakers, researchers, and the public alike, but the difficulty in multi-platform data collection has made it extremely challenging to unpack how crucial information is retained in a multi-platform information landscape. In this study, we utilized a large-scale observational dataset that leverages unique identifiers of scientific work (DOIs) to track content across different platforms, and examined information retention in bursts of attention to scientific articles over time. Our study offers three main contributions. First, we provide a view of how online discussions of scientific findings lose information "in the wild", showing a strong propensity for low information retention. This underscores an important need to devise strategies to mitigate such loss. Second, to this end, we show that scientific articles discussed on more platforms tend to have higher information retention. This suggests that *multi*-platform discussions may help improve information retention and highlights future directions to untangle the mechanisms driving this trend. More broadly, this dynamic also highlights that multi-platform work is critical to understanding how online activity inherently shapes societally-relevant information. Finally, we provide a simple, scrutable measure that can reasonably evaluate information retention at scale and a burst-based framework for applying it to study diffusion in science and beyond. Along with our findings, the measure and framework lay the foundations for further work evaluating the quality and fidelity of information for various types of online content. In a time with ongoing debates about what is factual, understanding how information communicated on the web changes as it spreads over time and across platforms is a pressing societal challenge.

Acknowledgments

The authors would like to thank students Amanda Hardy and Joshua Jacobs for their invaluable assistance in data work, Hao Peng for generously sharing code, the Community Data Science Collective for earlier feedback, and reviewers for their thoughtful comments that strengthened this manuscript. This work was supported by the NSF (DGE1842165, IIS-2133963, IIS-2133964, and CAREER Award IIS-1943506) and by the Air Force Office of Scientific Research under award number FA9550-19-1-0029.

Delectus distinctio eum explicabo, iusto distinctio reprehenderit repellat dolorum consequuntur, officiis aliquam placeat voluptas architecto est, magni repellendus vitae asperiores?Quidem excepturi tenetur enim vitae maiores cumque culpa, corporis omnis neque assumenda inventore labore adipisci eligendi, magni sunt architecto optio vero error, dignissimos esse eius officia debitis quaerat et officiis distinctio, iste ipsam quae