Introduction

I am on the phone with a journalist discussing my research into anti-vaccination. As the conversation winds up, they ask a question I have come to expect: "how big do you think this is?" My answer is usually some version of the following: that we have no way of knowing. I and my fellow researchers can only see the information that is public or in the sunlight. How anti-vaccination information spreads through private networks is dark to us. It is private and necessarily so. This means that we cannot track how these conversations spread in the private or parochial spaces of Facebook, nor can we consider how they might extend into other modes of mediated communication.

Modern communication is a complex and multiplatform accomplishment. Consider this: I am texting with my friend, I send her a selfie, in the same moment I hear a notification, she has DMed me a relevant Instagram post via that app. I move to Instagram and share another post in response; we continue our text message conversation there. Later in the day, I message her on Facebook Messenger while participating in a mutual WhatsApp group chat. The next day we Skype, and while we talk, we send links back and forth, which in hindsight are as clear as hieroglyphics before the Rosetta stone. I comment on her Twitter post, and we publicly converse back and forth briefly while other people like our posts. None of these instances are discrete conversational events, even though they occur on different platforms.
They are iterations on the same themes, and the archival properties of social media and private messaging apps mean that neither of you forgets where you left off. The conversation slides not only between platforms and contexts but in and out of visibility. Digitally mediated conversation hums in the background of daily life (boring meetings, long commutes and bad dates) and expands our understanding of the temporal and sequential limits of conversation.

In this article, I will explore digitally-mediated cross-platform conversation as a problem in two parts, and how we can understand it as part of the 'dark social'. Specifically, I want to draw attention to how online spaces are part of our everyday communicative practices and are not necessarily synonymous with the illicit, illegal, or deviant.

I argue that the private conversations we have online are also part of the dark social web, insofar as they are hidden from the public eye. When I think of dark social spaces, I think of what lies beneath the surface of murky waters, what hides behind in backstage areas, and the moments between platforms. In contrast, 'light' (or public) social spaces are often perceived as siloed. The boundaries between these platforms are artificially clean and do not appear to leak into other spaces. This article explores the dark and shadowed spaces of online conversation and considers how we might approach them as researchers. Conversations occur in the backchannels of social media platforms, in private messaging functions that are necessarily invisible to the researcher's gaze. These spaces are distinct from the social media activity analysed by Marwick and boyd. Their research examining teens' privacy strategies on social media highlights how social media posts that multiple audiences may view often hold encoded meanings. Social media posts are a distinct and separate category of activity from meditated conversations that occur one to one, or in smaller group chat settings.

Second is the disjunction between social media platforms. Users spread their activity across any number of social media platforms, according to social and personal logics. However, these movements are difficult to capture; it is difficult to see in the dark. Platforms are not hermeneutically sealed off from each other, or the broader web. I argue that understanding how conversation moves between platforms and in the backstage spaces of platforms are two parts of the same dark social puzzle.

Conversation Online

Digital media have changed how we maintain our social connections across time and space. Social media environments offer new possibilities for communication and engagement as well as new avenues for control. Calls and texts can be ignored, and our phones are often used as shields. Busying ourselves with them can help us avoid unwanted face-to-face conversations. There are a number of critiques regarding the pressure of always-on contact, and a growing body of research that examines how users negotiate these demands. By examining group messaging, Mannell highlights how the boundaries of these chats are porous and flexible and mark a distinct communicative break from previous forms of mobile messaging, which were largely didactic. The advent of group chats has also led to an increasing complication of conversation boundaries. One group chat may have several strands of conversation sporadically re-engaged with over time. Manell’s examination of group chats empirically illustrates the complexity of digitally-mediated conversations as they move across private, parochial, and public spaces in a way that is not necessarily temporally linear.

Further research highlights the networked nature of digitally mediated interpersonal communication and how conversations sprawl across multiple platforms (Burchell). Couldry (16, 17) describes this complex web as the media manifold. This concept encompasses the networked platforms that comprise it and refers to its embeddedness in daily life. As we no longer “log on” to the internet to send and receive email, the manifold is both everywhere and nowhere; so too are our conversations. Gershon has described the ways we navigate the communicative affordances of these platforms as "media ideologies" which are the "beliefs,
attitudes, and strategies about the media they [individuals] use" (391). Media ideologies also contain implicit assumptions about which platforms are best for delivering which kinds of messages.

Similarly, Burchell argues that the relational ordering of available media technologies is "highly idiosyncratic" (418). Burchell contends that this idiosyncratic ordering is interdependent and relational, and that norms about what to do when are both assumed by individuals and learnt in their engagement with others (418). The influence of others allows us to adjust our practices, or as Burchell argues, "to and regulate one's own conduct ... and facilitate engagement despite the diverse media practices of" (418). In this model, individuals are constantly learning and renegotiating norms of conversation on a case by case, platform by platform basis. However, I argue that it is more illuminating to consider how we have collectively developed an implicit and unconscious set of norms and signals that govern our (collective) conduct, as digitally mediated conversation has become embedded in our daily lives. This is not to say that everyone has the same conversational skill level, but rather that we have developed a common toolbox for understanding the ebb and flow of digitally mediated conversations across platforms. However, these norms are implicit, and we only have a partial understanding of how they are socially achieved in digitally-mediated conversation.

**What Lies Beneath**

Most of what we do online is assumed not to be publicly visible. While companies like Facebook trace us across the web and peer into every nook and cranny of our private use patterns, researchers have remained focussed on what lies above in the light, not below, in the dark. This has meant an overwhelming focus on single platform studies that rely on the massification of data as a default measure for analysing sentiment and behaviour online. Sociologically, we know that what occurs in dark social spaces, or backstage, is just as important to social life as what happens in front of an audience (Goffman). Goffman's research uses the metaphor of the theatre to analyse how social life is accomplished as a performance. He highlights that (darkened) backstage spaces are those where we can relax, drop our front, and reveal parts of our (social) self that may be unpalatable to a broader audience.

Simply, the public data accessible to researchers on social media are “trace data”, or “trace conversation”, from the places where conversations briefly leave (public) footprints and can be tracked and traced before vanishing again. Alternatively, we can visualise internet researchers as swabbing door handles for trace evidence, attempting to assemble a narrative out of a left-behind thread or a stray fingerprint.

These public utterances, often scraped through API access, are only small parts of the richness of online conversation. Conversations weave across multiple platforms, yet single platforms are focussed on, bracketing off their leaky edges in favour of certainty. We know the social rules of platforms, but less about the rules between platforms, and in their darker spaces.

Conversations briefly emerge into the light, only to disappear again. Without understanding how conversation is achieved and how it expands and contracts and weaves in and out of the present, we are only ever guessing about the social dynamics of mediated conversation as they shift between light, dark, and shadow spaces. Small things can cast large shadows; something that looms large may be deceptively small. Online they could be sociality distorted by disinformation campaigns or swarms of social bots.

**Capturing the Unseen: An Ethnomethodological Approach**

Not all data are measurable, computable, and controllable. There is uncertainty beyond what computational logics can achieve. Nooks and crannies of sociality exist beyond the purview of computable data.
This suggests that we can apply pre-digital social research methods to capture these “below the surface” conversations and understand their logics. Sociologists have long understood that conversation is a social accomplishment. In the 1960s, sociologist Harvey Sacks developed conversation analysis as an ethnomethodological technique that seeks to understand how social life is accomplished in day-to-day conversation and micro-interactions. Conversation analysis is a detailed and systematic account of how naturally-occurring talk is socially ordered, and has been applied across a number of social contexts including news interviews, judicial settings, suicide prevention hotlines, therapy sessions, as well as phone conversations (Kitzinger and Frith).

Conversation analysis focusses on fine-grained detail, all of the little patterns of speech that make up a conversation; for example, the pauses, interruptions, self-corrections, false starts, and over-speaking. Often these too are hidden features of conversation, understood implicitly, but hovering on the edges of our social knowledge. One of the most interesting uses of conversational analysis is to understand refusal, that is, how we say 'no' as a social action. This body of research turns common-sense social knowledge – that saying no is socially difficult – into a systemic schema of social action. For instance, acceptance is easy to achieve; saying yes typically happens quickly and without hesitation. Acceptances are not qualified; a straightforward 'yes' is sufficient (Kitzinger and Frith). However, refusals are much more socially complex. Refusal is usually accomplished by apologies, compliments, and other palliative strategies that seek to cushion the blow of refusals. They are delayed and indirect conversational routes, indicating their status as a dispreferred social action, necessitating their accompaniment by excuses or explanations (Kitzinger and Frith). Research by Kitzinger and Frith, examining how women refuse sexual advances, illustrates that we all have a stock of common-sense knowledge about how refusals are typically achieved, which persists across various social contexts, including in our intimate relationships.

Conversation analysis shows us how conversation is achieved and how we understand each other. To date, conversation analysis techniques have been applied to spoken conversation but not yet extended into text-based mediated conversation. I argue that we could apply insights from conversation analysis to understand the rules that govern digitally mediated conversation, how conversation moves in the spaces between platforms, and the rules that govern its emergence into public visibility. What rules shape the success of mediated communication? How can we understand it as a social achievement? When conversation analysis walks into the dark room it can be like turning on the light.

How can we apply conversation analysis, usually concerned with the hidden aspects of plainly visible talk, to conversation in dark social spaces, across platforms and in private back channels? There is evidence that the norms of refusal, as highlighted by conversation analysis, are persistent across platforms, including in people's private digitally-mediated conversations. One of the ways in which we can identify these norms in action is by examining technology resistance. Relational communication via mobile device is pervasive (Hall and Baym). The concentration of digitally-mediated communication into smartphones means that conversational norms are constantly renegotiated, alongside expectations of relationship maintenance in voluntary social relationships like friendship (Hall and Baym). Mannell also explains that technology resistance can include lying by text message when explaining non-availability. These small, habitual, and often automatic lies are categorised as “butler lies” and are a polite way of achieving refusal in digitally mediated conversations that are analogous to how refusal is accomplished in face-to-face conversation. Refusals, rejections, and, by extension, unavailability appear to be accompanied by the palliative actions that help us achieve refusal in face-to-face conversation. Mannell identifies strategies such as “feeling ill” to explain non-availability without hurting others' feelings. Insights from conversation analysis suggest that on balance, it is likely that all parties involved in both the furnishing and acceptance of a butler lie understand that these are polite fabrications, much like the refusals in verbal conversation.

Because of their invisibility, it is easy to assume that conversations in the dark social are chaotic and disorganised. However, there are tantalising hints that the reverse is true. Instead of arguing that
individuals construct conversational norms on a case by case, platform by platform basis, I suggest that we now have a stock of common-sense social knowledge that we also apply to cross-platform mediated communication. In the spaces where gaps in this knowledge exist, Szabla and Blommaert argue that actors use existing norms of interactions and can navigate a range of interaction events even in online environments where we would expect to see a degree of context collapse and interactional disorganisation.

Techniques of Detection

How do we see in the dark? Some nascent research suggests a way forward that will help us understand the rhythms of cross-platform mediated conversation. Apps have been used to track participants' messaging and calling activities (Birnholtz, Davison, and Li). This research found a number of patterns that signal a user's attention or inattention, including response times and linguistic clues. Similarly, not-for-profit newsroom The Markup built a Facebook inspector called the citizen browser, a "standalone desktop application that was distributed to a panel of more than 1000 paid participants" (Mattu et al.). The application works by being connected to a participant's Facebook account and periodically capturing data from their Facebook feeds. The data is automatically deidentified but is still linked to the demographic information that participants provide about themselves, such as gender, race, location, and age. Applications like these point to how researchers might reliably collect interaction data from Facebook to glimpse into the hidden networks and interactions that drive conversation. User-focussed data collection methods also help us, as researchers, to sever our reliance on API access. API-reliant research is dependent on the largesse of social media companies for continued access and encourages research on the macro at the micro's expense. After all, social media and other digital platforms are partly constituted by the social acts of their users. Without speech acts that constitute mediated conversation, liking, sharing GIFs, and links, as well as the gaps and silences, digital platforms cease to exist. Digital platforms are not just archives of “big data”, but rather they are collections of speech and records of how our common-sense knowledge about how to communicate has stretched and expanded beyond face-to-face contexts.

A Problem of Bots

Ethnomethodological approaches have been critiqued as focussing too much on the small details of conversation, on nit-picking small details, and thus, as unable to comment on macro social issues of oppression and inequality (Kitzinger and Frith 311). However, understanding digitally-mediated conversation through the lens of talk-as-human-interaction may help us untangle our most pressing social problems across digital platforms.

Extensive research examines platforms such as Twitter for “inauthentic” behaviour, primarily identifying which accounts are bots. Bots accounts are programmed Twitter accounts (for example) that automatically tweet information on political or contentious issues, while mimicking genuine engagement. Bots can reply to direct messages too; they converse with us as they are programmed to act as “humanly” as possible.

Despite this, there are patterns of behaviour and engagement that distinguish programmed bot accounts, and a number of platforms are dedicated to their detection. However, bots are becoming increasingly sophisticated and better able to mimic “real” human engagement online. But there is as yet no systematic framework regarding what “real” digitally mediated conversation looks like. An ethnomethodological approach to understanding this would better equip platforms to understand inauthentic activity. As Yang and colleagues succinctly state, "a supervised machine learning tool is only as good as the data used for its training ... even the most advanced [bot detection] algorithms will fail with outdated training datasets" (8).

On the flipside, organisations are using chat bots to deliver cognitive behavioural therapy and assist people in moments of psychological distress. But the bots do not feel human; they reply instantly to any message sent. Some require responses in the form of emojis. The basis of therapy is talk. Understanding more
accurately how naturally-occurring talk functions in online spaces could create more sensitive and genuinely therapeutic tools.

**Conclusion**

It is easy to forget that social media have largely mainstreamed over the last decade; in this decade, crucial social norms about how we converse online have developed. These norms allow us to navigate conversations, with intimate friends and strangers alike across platforms, both in and out of public view, in ways that are often temporally non-sequential. Dark social spaces are a matter of intense marketing interest. Advertising firm *Disruptive Advertising* identified the very spaces that are the focus of this article as "dark social": messaging apps, direct messaging, and native mobile apps facilitate user activity that is "not as easily controlled nor tracked". Dark social traffic continues to grow, yet our understanding of why, how, and for whom trails behind.

To make sense of our social world, which is increasingly indistinguishable from online activity, we need to examine the spaces between and behind platforms, and how they co-mingle. Where are the spaces where the affordances of multiple platforms and technologies scrape against each other in uncomfortable ways? How do users achieve intelligible conversation not just because of affordances, but despite them? Focussing on micro-sociological encounters and conversations may also help us understand what could build a healthy online ecosystem. How are consensus and agreement achieved online? What are the persistent speech acts (or text acts) that signal when consensus is achieved?

To begin where I started, to understand the scope and power of anti-vaccination sentiment, we need to understand how it is shared and discussed in dark social spaces, in messaging applications, and other backchannel spaces. Taking an ethnomethodological approach to these conversational interactions could also help us determine how misinformation is refused, accepted, and negotiated in mediated conversation. Focussing on "dark conversation" will help us more richly understand our social world and add much needed insight into some of our pressing social problems.

**References**


**Author Biography**

**Naomi Smith, Federation University Australia**

Dr Naomi Smith is a digital sociologist at Federation University Australia (Gippsland). She has a broad range of scholarly interests, including emerging technology, place and bodies. Primarily, her digital work has focused on the intersection of the internet and bodies (including anti-vaccination), how online communities influence the way we make sense of our bodies, and how we manage them.

**License**

Copyright (c) 2021 Naomi Smith

This work is licensed under a [Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License](https://creativecommons.org/licenses/by-nc-nd/4.0/).

Authors who publish with this journal agree to the following terms:

1. Authors retain copyright and grant the journal right of first publication with the work simultaneously licenced under a [Creative Commons Attribution - Noncommercial - No Derivatives 4.0 Licence](https://creativecommons.org/licenses/by-nc-nd/4.0/) that allows others to share the work with an acknowledgement of the work's authorship and initial publication in this journal.

2. Authors are able to enter into separate, additional contractual arrangements for the non-exclusive distribution of the journal's published version of the work (e.g., post it to an institutional repository or publish it in a book), with an acknowledgement of its initial publication in this journal.

3. Authors are permitted and encouraged to post their work online (e.g., in institutional repositories or on their website) prior to and during the submission process, as it can lead to productive exchanges, as well as earlier and greater citation of published work (see [The Effect of Open Access](https://www.elsevier.com/about/our-research-the-importance-of-open-access)).