Truth in an Age of Information

Alan Dix$^{1,2}$

$^1$Professorial Fellow, Cardiff Metropolitan University, Wales UK
$^2$Director of the Computational Foundry, Swansea University, Wales, UK
alan@hcibook.com, https://alandix.com

Keywords: misinformation, argumentation, provenance, qualitative–quantitative reasoning, fake news, echo chambers

Abstract: Many of the issues in the modern world are complex and multifaceted: migration, banking, not to mention climate change and Covid. Furthermore, social-media, which at first seemed to offer more reliable 'on the ground' citizen journalism, has instead become a seedbed of dis-information. Trust in media has plummeted, just when it has become essential. This is a problem, but also an opportunity for research in HCI that can make a real difference in the world. The majority of work in this area, from various disciplines including data-science, AI and HCI, is focused on combatting misinformation – fighting back against bad actors. However, we should also think about doing better – helping good actors to curate, disseminate and comprehend information better. There is exciting work in this area, but much still to do.

1 INTRODUCTION

Falsehood flies, and truth comes limping after it.
Jonathan Swift, The Examiner No. 14,
Thursday, 9th November 1710

Politicians have always been ‘economical with the truth’ and newspapers have toed an editorial line. However, never in recent times does it seem that confidence in our media has been lower. From the Brexit battle bus in the UK to suspected Russian meddling in US elections, fake news to alternative facts – it seems impossible for the general public to make sense of the contradictory arguments and suspect evidence presented both in social media and traditional channels. Even seasoned journalists and editors seem unable to keep up with the pace and complexity of news.

These problems were highlighted during Covid when understanding of complex epidemiological data was essential for effective government policy and individual responses. As well as the difficulty of media (and often government) in understanding and communicating the complexity of the situation, various forms of misinformation caused confusion. There are obvious health impacts of this misinformation due to taking dangerous ‘cures’ (Nelson, 2020) and vaccination hesitancy (Lee, 2022a), as well as its role in encouraging violence against health workers (Mahase, 2022). In addition, a meta-review of many studies of Covid misinformation identified mental health impacts as also significant (Rocha, 2021).

If democracy is to survive and nations coordinate to address global crises, we desperately need tools and methods to help ordinary people make sense of the extraordinary events around them: to sift fact from surmise, lies from mistakes, and reason from rhetoric. Similarly, journalists need the means to help them keep track of the surfeit of data and information so that the stories they tell us are rooted in solid evidence.

Crucially in increasingly politically fragmented societies, we need to help citizens explore their conflicts and disagreements, not so that they will necessarily agree, but so that they can more clearly understand their differences.

These are not easy problems and do not admit trite solutions. However, there is existing work that offers hope: tracking the provenance of press images (ICP, 2016), ways to expose the arguments in political debate (Carneiro, 2019), even using betting odds to track the influence of news on electoral opinion (Wall, 2017).
I hope that this paper will give hope that we can make a difference and offer challenges for future research.

1. THE B-MOVIE CAST OF MISINFORMATION

Deliberate misinformation is perhaps the most obvious problem we face. There is extensive data science studies by academics and data journalists attempting to understand the extent and modes of spread (e.g. Albright, 2016; Vosoughi, 2018). Crucially false information appears to spread more rapidly than true information; possibly because it is more novel (Vosoughi, 2018). Although there is considerable debate as to the sufficiency of their responses, both Facebook and Twitter are constantly adjusting algorithms and policies to attempt to prevent or discourage fake news (Dreyfuss, 2019; NPR, 2022; Twitter, 2022). Within the HCI community there has been considerable work exploring the human aspects around the spread of misinformation online (Flintham, 2018, Geeng, 2020; Varanasi, 2022), ways to visualise it (Lee, 2022b), tools for end-users to help identify it (Heuer, 2022) and CHI workshops (Gamage, 2022; Piccolo, 2021).

1.1 Bad actors

Much of the focus on misinformation is on ‘bad actors’: extremist organisations, ‘foreign’ powers interfering in elections, or simply those aiming to make a fast buck. In the context of mis-information, ‘bad’ can mean two things:
1. They are intrinsically bad people, bad states, or bad media.
2. They use bad methods and/or spread bad information (including misinformation and hateful or violent content).

The first of these can be relative to clear criteria such as human rights or terrorism, but may simply mean those we disagree with; and, of course, the boundary between the two may often be unclear.

When the two forms of ‘bad’ agree the moral imperative is clear, even though implementation may be harder. Forced in part by government and popular pressure, social media platforms have extensive mechanisms both to attempt to suppress bad information and suspend accounts of those who promulgate it (Guardian, 2018).

Probably the most high-profile example of the later was Twitter’s suspension of @realDonaldTrump. This was both met with widespread relief, but also caution due to its potential impact on free speech (Noor, 2021), especially given Twitter’s arguments for why it was suspended when it was (Twitter, 2021).

Of course, sometimes bad actors may spread true (or even good) information.

In some cases this is simply because few are altogether bad. For example, those who believe and then promulgate Covid conspiracy theories; many will be well meaning, albeit deeply misguided, and some of the information may be accurate.

However, true information can also be cynically used to give credence to otherwise weak or misleading arguments; for example a recent study of cross-platform misinformation (Micallef, 2022) found a substantial proportion of cases where a YouTube video with true information about Covid was referenced by a tweet or post that in some way mis-interpreted the material or used it out of context. In addition, many Astroturfing accounts will distribute accurate information as a means to create trust before disseminating misinformation. It can be hard to distinguish these and it is not uncommon for politicians or other campaign groups to inadvertently re-tweet or quote true or at least defensible information that originated from very unsavoury groups, thus giving them credence.

1.2 When good actors spread bad information

As we saw in the last example, those we regard as ‘good’ actors can also sometimes spread bad information.

Sometimes this is deliberate. An extreme case is during war when misinformation campaigns in an enemy country are regarded as a normal and indeed relatively benign form of warfare (Shaer, 2017). In peace time deliberate misinformation is likely to be less extreme and more often stretching or embroidering the truth, or selectively reporting. It may also be accidental. For example, Figure 1 shows a “Q&A” (form of fact check) on the BBC news web site following a claim made by Boris Johnson in January 2018 regarding UK contributions to the EU budget. The overall thrust of the Q&A is correct, the net amount that was sent to the EU at that time was substantially less than the £350 million figure that Johnson claimed, but the actual figures are wrong, the Q&A suggested that around 2/3 of the gross figure was returned, when the actual figure was close to a half. This is probably because at some point a journalist lost track of which figure the half was
referring to, but the overall effect was to create a substantially incorrect figure.

Figure 1: Q&A on BBC new web site January 2018 (BBC, 2018). Note this Q&A pop-up is no longer in the news item; instead there a link to a ‘Reality Check’ page which is correct, but with no explicit retraction.

In between are the subtle biases are simply assumptions of journalists that play out in the selection of which stories to report and also in the language used. For example, in crime or conflict reporting passive language may be used (“the assailant was shot”, or “shells fell on”) compared with active language (“AAA shot BBB” or “XXX fired shells on”) depending on which side is doing the shooting or bombing.

Personally, while I may despair or be angry at the misinformation from those with whom I disagree, I am most upset when I see poor arguments from those with whom I agree. This is partly pride, wanting to be able to maintain a moral high ground, and partly pragmatic, if the arguments are poor then they can be refuted.

In an age of adversarial media, any mistakes, misrepresentation or hyperbole can be used to discredit otherwise well-meaning sources and promote alternatives that are either ill-informed or malicious. This was evident in the US during the 2016 presidential campaign when many moderate Republican supporters lost faith in the reputable national press in favour of highly partisan local papers; a trend which has intensified since (Gottfried, 2021; Meek, 2021)

2. SEEKING TRUTH

2.1 The full cast

We have already considered the ‘B-movie’ bad/good guy roles, of the producers and influencers, both of whom can mislead whether ill-intentioned or ill-informed. In reality even the ‘bad’ actors may be those with genuinely held, albeit unfounded, beliefs about 5G masts or communist take-over of US government. Of course, those of us who would consider ourselves ‘good’ actors, may still distort or be selective in what we say albeit for the best of reasons.

In addition, those who receive misinformation and are confused or misled by it may differ in levels of culpability. It is easier to believe the things that make life easier, whether it is the student grasping at suggestions that the impact of Covid may be over exaggerated in order to justify a party, or the professional accepting climate change scepticism to justify buying that new fuel-hungry car.

Of course, the purveyors of news and information are under pressure, and may not be wholly free in what they say, or may run risks if they do. Even in the last year we have seen many journalists, bloggers and authors arrested, sanctioned, stabbed and shot.

Perhaps more subtle is the interplay within the ecology of information: journalists and social media modify what and how they present information in order to match the perceived opinions and abilities of their readership.

2.2 Two paths

The greatest effort currently appears to be focused on fighting back against bad actors. This includes algorithms to detect and counter mis-information, such as Facebook’s intentions to weed out anti-vaccination. These are predominantly aimed at the bad actors.

However, in addition we need to think about doing better, ways for the good actors to disseminate and understand information so that they are in a better position to evaluate sources of information and ensure that they do not inadvertently create bad information.

We’ll look briefly at four areas where appropriate design could help us to do better:

• echo chambers and filter bubbles
• better argumentation
• data and provenance
• numeric data and qualitative-quantitative reasoning

These are not the only approaches, but I hope they will stimulate the reader to think of more.

2.3 Echo chambers and breaking filter bubbles

Social media was initially seen as a way to democratise news and information sharing and to allow those in the ‘long-tail’ of small interest groups to find like-minded people in the global internet. However, we now all realise that an outcome of this has been the creation of echo chambers, where we
increasingly only hear views that agree with our own. In some ways this has always been the case, both in choices of friendship groups for informal communication and the audiences of different newspapers. However, social media and the personalisation of digital media has both intensified the effect and made it less obvious – you know that a newspaper has a particular editorial line, but do not necessarily recognize that web search results have been tuned to your existing prejudice.

This is now a well-studied area with extensive work analysing social media to detect filter bubbles and understand the patterns of communication and networks that give rise to them (Terren, 2021; Garimella, 2018; Cinelli, 2021). Notably, one of these studies (Garimella, 2018) highlighted the role of ‘gatekeeper’, people who consume a broad range of content, but then select from this to create partisan streams. Perhaps more sadly, the same study notes that those who try to break down partisan barriers pay a “price of bipartisanship” in that balanced approaches or multiple viewpoints are not generally appreciated by their audiences.

In addition, there has been work on designing systems that in different ways attempt to help people see beyond their own filter bubbles (e.g. Foth, 2016; Jeon, 2021), but on the whole this has been less successful, especially in actual deployment. Indeed, attempts to present opposite arguments can end up deepening divides if they are too different and too soon.

### 2.4 Argumentation

It is easy to see the flaws in arguments with which we disagree, we know it is wrong and can thus hunt for the faults – the places where our intuitions and the argument disagree are precisely the places where we are expecting holes in the reasoning. Of course, we all create bad arguments. It is very hard to notice the gaps in one’s own reasoning, but also the fallacious arguments of others when one agrees with their final conclusions.

Of course, those who disagree with us will notice the gaps in our arguments, thus increasing their own confidence and leading them to discount our opinions!

It is crucial therefore to have tools that both help the public to interrogate the arguments of politicians and influencers, and also to help those who are aiming to create solid evidence-based work (including academics) to ensure valid arguments.

There is of course long-standing work on argumentation systems, such as IBIS (Noble, 1988) and work in the NLP community to automatically analyse arguments. Much of this is targeted towards more professional audiences, but there are also steps to help the general public engage with media, such as the Deb8 system (Carneiro, 2019) developed at St Andrews, an accessible argumentation system that allows viewers of a speech or debate to collaboratively link assertions in the video to evidence from the web.

This is an area which seems to have many opportunities for research and practical systems aimed at different audiences including the general public, journalists, politicians, academics, and fact checkers. This could include broad advice, for example, ensuring that fact checkers clearly state their interpretation of a statement before checking it to avoid inadvertently debunking a strawman misinterpretation. Similarly, we could imagine templates for arguments, for example, given an implication of the form “if A then B”, it is important to keep track of the assumptions. In particular, while more formal logics and some forms of argumentation schemes focus on low-level argumentation, it seems that the tools needed perhaps need to focus on the higher-level argumentation, the information and assumptions that underly a statement, more than the precise logic of the inference.

In addition, in the AI community there are now a variety of tools to help automatically detect possible bias in data or machine learning algorithms. Maybe some of these could be borrowed to help human reasoning, for example shuffling aspects of situations (e.g. gender, political party or ethnicity), to help us assess to what extent our view is shaped by these factors.

### 2.5 Data and provenance

One of the forms of misinformation is the deliberate or accidental use of true information or accurate data divorced from its context. For the spoken word or text, this might be a quotation, for photographs or video the choice of a still, segment or even parts edited together that give a misleading impression. Indeed the potential for digital media to be compromised in different ways lead some to look for technology such as blockchains to prevent tampering, or the use of analogue or physical representations (Haliburton, 2021).

One example of work addressing this issue was the FourCorners project (ICP, 2016), a collaboration between OpenLab Newcastle, the International Centre for Photography and the World Press Photo Foundation, which embeds provenance into photographs allowing interrogation such as “what are the frames before and after this photograph?”, “are
there other photos at the same time and place?”. One can imagine similar things for textual quotes, in the manner of Ted Nelson’s vision of transclusion (Nelson, 1981), where segments quoted from one document retain their connection back to the original.

This is an area I’ve worked on personally in the past with the Snip!t system, originally developed in 2003 following a study of user bookmarking practice (Dix, 2003). Snip!t allowed users to ‘bookmark’ portions of a web page and automatically kept track not just of the quoted text, but where it came from (Dix, 2010). Later work in this area by others has included both commercial systems such as Evernote, and academic research, such as Information Scraps (Bernstein, 2008). Currently there is an explosion of personal knowledge management (PKM) apps, some of which, such as Readwise (readwise.io) and Instapaper (instapaper.com), help with the process of annotating documents. However, these systems are mostly focused on retaining the context of captured notes and quotes; we desperately need better ways to retain this once the quote is embedded in another document or web page.

This connection to sources is also important for data. In the example from the BBC in Figure 1, the journalist had clearly lost track of the original data on UK/EU funding and so misremembered aspects.

Can we imagine tools for journalists that would help them keep track of the sources for data and images. Indeed, it would be transformative if everyday office tools such as word processors and presentation software made it easy to keep references to imported images. In work with humanities and heritage, we have noted how file systems have barely altered since the 1970s (Dix, 2022) – the folder structures allow us to store and roughly classify, but there is virtually no support for talking about documents and about their relationships to one another. Semantic desktop research (Sauermann, 2005), which seemed promising at the time, has never found its way into actual operating systems.

Happily there are projects, such as Data Stories (2022) that are helping communities to use data to tell their own stories, so that the online world can allow open discourse and interpretation, whilst connecting to the underlying data on which it is based. Furthermore, one of the popular PKM apps Obsidian (obsidian.md) supports semi-structured meta-data for every note.

### 2.6 Numeric data and qualitative–quantitative reasoning

Going back to the example in figure 1, part of the problem here may well simply be that journalists are often more adept with words than numbers. We are in a world where data and numerical arguments are critical. This was true of Covid where the understanding of exponential growth and probabilistic behaviour was crucial, but equally so for issues such as climate change.

One of the arguments put forward by climate change sceptics, is that it is hard to believe in long-term climate models given forecasters sometimes struggle to predict whether it is going to rain next week. This, at first sight, is not an unreasonable argument; although anyone who has dealt with stochastic phenomena knows that it is often easier to predict long-term trends than short-term behaviour. Indeed, it is also relatively easy to communicate this – we can all say with a degree of reliability that a British winter will be wetter and colder than the summer, even though we’ll struggle to know the weather from day to day.

This form of argument is not about exact numerical calculation, nor about abstract mathematics, but something else – informal reasoning about numerical phenomena. Elsewhere I’ve called this qualitative–quantitative reasoning (Dix, 2021a, 2021b) and seems to be a critical, but largely missing, aspect for universal education. Again this is an area that is open for radical contributions, for example, iVolver (Nacenta, 2017) allows users to extract numerical and other data from visualisations, such as pie charts, in published media. My own work has included producing table recognisers in commercial intelligent internet system On Cue in the dot-com years (Dix, 2000) and more recently investigating ways to leverage some of the accessibility of spreadsheet-like interfaces and simple ways to allow users to combine their own data (Dix, 2016).

### 3. CALL TO ACTION

We are at a crucial time in a world where information is everywhere and yet we can struggle to see the truth amongst the poorly sourced, weakly argued, deliberately manipulated or simply irrelevant. However, there are clear signs of hope in work that is being done and also opportunities for research that can make a real difference.

Of course, as academics we are also in the midst of a flood of scholarly publication, some more scholarly than others! There are calls for us to ‘clean up our own act’ too including rigour of academic argumentation (Bassboll, 2018) and transparency of data and materials (Wacharamanotham, 2020). As well as being a problem we need to deal with within academia, it is also an opportunity to use our own
academic community as a testbed for tools and techniques that could be used more widely.

REFERENCES


Mahase E. (2022). Nearly 1500 health workers were attacked or arrested in 2021, report finds. BMJ 2022; 377:o1315 doi:10.1136/bmj.o1315


NPR (2022) Twitter aims to crack down on misinformation, including misleading posts about Ukraine. NPR Technology. May 19, 2022. https://www.npr.org/2022/05/19/1100100329/twitter-misinformation-policy-ukraine


