A Perfect Storm: Social Media News, Psychological Biases, and AI

PRATIM DATTA and MARK WHITMORE, Kent State University

JOSEPH K. NWANKPA, Miami University

In an age where news information is created by millions and consumed by billions over **social media (SM)** every day, issues of information biases, fake news, and echo-chambers have dominated the corridors of technology firms, news corporations, policy makers, and society. While multiple disciplines have tried to tackle the issue using their disciplinary lenses, there has, hitherto, been no integrative model that surface the intricate, albeit “dark” explainable AI confluence of both technology and psychology. Investigating information bias anchoring as the overarching phenomenon, this research proposes a theoretical framework that brings together traditionally fragmented domains of AI technology, and human psychology.

The proposed Information Bias Anchoring Model reveals how SM news information creates an information deluge leading to uncertainty, and how technological rationality and individual biases intersect to mitigate the uncertainty, often leading to news information biases. The research ends with a discussion of contributions and offering to reduce information bias anchoring.

CCS Concepts: • **Human-centered computing** → *Collaborative and social computing; Collaborative and social computing theory, concepts and paradigms*; **Social networks**;

Additional Key Words and Phrases: AI, XAI, DLP, bias, information bias, anchoring, fake news

ACM Reference format:

Pratim Datta, Mark Whitmore, and Joseph K. Nwankpa. 2021. A Perfect Storm: Social Media News, Psychological Biases, and AI. *Digit. Threat.: Res. Pract.* 2, 2, Article 15 (April 2021), 21 pages.

<https://doi.org/10.1145/3428157>

1 INTRODUCTION, MOTIVATION, AND RESEARCH QUESTION

“For this can generally be said of men: that they are...fickle...and as long as you serve their welfare, they are entirely yours.” (Machiavelli, *Il Principe*)

In October 2018, Apple CEO Tim Cook, speaking at the 40th International Conference of Data Protection and Privacy Commissioners, specifically noted “Our own information—from the everyday to the deeply personal—is being weaponized against us with military efficiency” and “these scraps of data, each one harmless enough on its own,” can “serve up increasingly extreme content, pounding our harmless preferences into harm.” Tim Cook

Authors’ addresses: P. Datta and M. Whitmore, College of Business, Kent State University, 800 E. Summit St. Kent, Ohio 44242; emails: {pdatta, mwhitmo4}@kent.edu; J. K. Nwankpa, Farmer School of Business, Miami University, 800 E High St. Oxford, Ohio 45056; email: nwankpjk@miamioh.edu.

Updated author affiliations: PRATIM DATTA, Kent State University, USA and University of Johannesburg, South Africa; MARK WHITMORE, Kent State University, USA; JOSEPH K. NWANKPA, Miami University, USA.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, or to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

© 2021 Association for Computing Machinery.

2576-5337/2021/04-ART15 \$15.00

<https://doi.org/10.1145/3428157>

Digital Threats: Research and Practice, Vol. 2, No. 2, Article 15. Publication date: April 2021.

further warns that “Platforms and algorithms that promised to improve our lives can actually magnify our worst human tendencies... rogue actors and even governments have taken advantage of user trust to deepen divisions, incite violence, and even undermine our shared sense of what is true and what is false. This crisis is real. It is not imagined, or exaggerated, or crazy” (emphases added).

Tim Cook’s mention of technology magnifying our worst human tendencies or proclivities are of particular interest to this research. While human biases predate **artificial intelligence (AI)**, AI can amplify and entrench (anchors) biases, leading to faster and often deadlier instances of fake news internalization and propagation, especially in light of a SM news deluge. Instead of treating AI machine learning as a black-box, this research explains how AI can reinforce and anchor human biases, thus adding to the **explainable AI (XAI)** repertoire. In what follows, this research surfaces the perfect storm by explaining the process by intersection of human tendencies (biases) and AI in exacerbating **information bias anchoring (IBA)** when faced with information uncertainty.

Anchoring is a cognitive bias that allows people to internalize and entrench (or anchor) their decision-making on particular information that appears to match (or amplify) their subjective arbitrary reference points (e.g., biases, beliefs, experiences, and hypotheses) [e.g., Norman and Deflin 2012].

It is well established that IBA increases susceptibility to fake news consumption and sharing [Lazer et al. 2018; Jost et al. 2020]. In a world replete with links to SM (most webpages carry a “share” link to SM sites such as Facebook, Twitter...), higher susceptibility to fake news along with a growing propensity to share the fake news creates a maelstrom effect. Therefore, it comes as no surprise that fake news and false rumors reach more people, penetrate deeper into the social network, and spread much faster than accurate stories [Silverman 2016]. A news content with sensational yet completely fictional news will generate more clicks and will be shared more on SM [Vosoughi et al. 2018] and surprisingly, people tend to believe fake news especially if the contents are aligned with their prior belief [Silverman and Singer-Vine 2016].

Understanding the factors that drive IBA and the underlying conditions through which IBA enables the propagation of fake news in SM is an important phenomenon of interest for various reasons. First, fake news has become a key part of our online news consumption on SM leading to undesirable consequences. Yet, the majority of the SM platform providers continue to struggle with ways to reduce the impact of fake news on their platforms [Shane 2017; Kim and Dennis 2018; Torres et al. 2018; Zhou et al. 2020].

First, one reason for platforms’ struggle in mitigating negative effects of fake news is that the burden lies to SM users to critically evaluate the information obtained from SM news feeds. Thus, understanding how users’ actions can inherently create confirmation biases is an important step toward mitigating the effects of fake news.

Second, several studies have noted the inherent content bias associated with search engines, SM news feeds, and AI-based recommendations [Bakshy et al. 2015]. These content providers use sophisticated algorithms to produce news feeds and contents tailored to our profiles. In an era where new AI algorithms and big data can create detailed user profiles and target SM users with news contents that are in line with their general belief and opinions on a topic, the role of AI-based technologies and its implication on IBA has not received adequate scrutiny.

Third, while there lies a rich body of extant research on understanding the consequences of and interventions for information biases [Ruokolainen and Widen 2020; George et al. 2000], there is little research that models antecedent biases leading to IBA. The conditions and mechanisms through which digital news media and SM platforms can enable IBA, has not been adequately investigated.

This article aims to contribute to the growing body of literature on fake news by integrating AI-based technological and psychology-based behavioral perspectives to build a theoretical model of information anchoring. This research forwards a conceptual framework that investigates how intersecting antecedent psychological and technological factors and conditions that shape information bias anchoring as a preface to fake news propagation.

This article argues that, in the context of digital news media, IBA is caused by the juxtaposition of the environment, AI-based personalization, and confirmation biases that drives information anchoring. Specifically, we ask, what are the environmental, technological, and psychological antecedents to IBA and its downstream effects on information anchoring in digital media?

In addressing this important yet under-represented question, this article describes how various SM news information aspects, namely, velocity, variety, and anonymity, create information overload, leading to a sense of uncertainty. Facing uncertainty, the article explains how individuals rely on how AI works in tandem with our own biases, proclivities, or preferences to offer us apparently “relevant” information that quells the uncertainty. Thus, **AI-based Deep Preferential Learning (AI DPL)** helps us make sense of a world in flux but, in the process, accentuates our biases by creating echo-chambers.

The research is organized as follows. We begin with a brief theoretical underpinning and literature review that succinctly covers individual cognition under conditions of information overload and uncertainty. Next, the proposed IBA model illustrates how SM news information overload, a function of SM news information velocity, variety, and anonymity, increases uncertainty. The framework further illustrates how, in the presence of uncertainty, individuals tend to capitulate to inherent biases and how AI helps perpetuate such biases, leading to an entrenched information bias anchoring. The research concludes with a discussion of the framework and its subsequent implications, limitations, and future research directions.

1.1 IBA in the Fake News Context

Fake News. While the concept of “fake news” has pervaded our social tapestry, its definition remains elusive and specious. Yet, given the predominant impact and mention of the term, the broad swathes that define fake news deserve attention and clarification. Allcott and Gentzkow [2017] defined fake news as “new articles that are intentionally and verifiably false and could mislead readers.” However, the study acknowledged the limitations of their definition, which excluded unintentionally reporting mistakes, misleading but not outright false statements, conspiracy theories, and rumors. Prior studies agree that fake news overlaps with misinformation and information disorder regardless of the intentions. A review of academic articles between 2003 and 2017 found the operationalization of the term fake news to include manipulation, propaganda, fabrication, satire, and parody [Tandoc et al. 2018]. An attempt at a multidisciplinary definition of fake news defined *Fake news* as “fabricated information that mimics news media content in form but not in organizational process or intent” [Lazer et al. 2018, p. 1094]. However, this definition, which is the result of a consensus among some researchers in the area, is very limiting and would exclude the use of “Fake News” for many kinds of information commonly associated with this term. Standard definitions of misinformation (information that is incorrect or misleading), or disinformation (false information that is deliberately and covertly spread to influence public opinion or obscure the truth (Merriam-Webster Dictionary)), is more inclusive. Furthermore, withholding vital information would also be considered a form of deception along with the distortion and fabrication of information in the definitions mentioned above. Yet, these forms of fake news are not captured in the [Lazer et al. 2018] definition of fake news.

Integrating a multiplicity of definitions, this article defines fake news as “news-related information originating from a purported news-outlet that is intentionally fabricated, distorted, disguised, withheld, or misinterpreted with or without an explicit or implicit motive to misinform, disinform, or deceive.”

Zhou et al. [2020] identify three components that must exist for information to be defined as fake news. These are authenticity, which refers to information that is not factual; intention, meaning aiming to mislead; and whether the information is news, i.e., relevant and current and originating from a seemingly credible entity within a network. The proposed definition encompasses all three of these components.

The definition proposed differs somewhat from this typology in that it acknowledges that fake news can be an intentional act to deceive (explicit) or can be disseminated by a sender without the knowledge of deception or bias and therefore unintentional (implicit). The original creation of the fake news is most likely an intentional act, but the sharing of this information may not be inherently intentional. For example, distortion and information bias are more pronounced in SM news than they are in national news-outlets [Budak et al. 2016]. However, SM news information consumers are more likely to share news without knowledge of its bias or lack of veracity [Chakraborty et al. 2016].

SM is replete with fake news, both created and shared, including current instances related to the ongoing COVID-19 crisis. According to University of Oxford's Reuters Institute [Brennen et al. 2020], the total *volume* of different kinds of coronavirus misinformation rose by 900% between January and March. COVID-19 fake news instances included how COVID-19 is not heat-resistant, vodka and bananas as a cure, some in the West chastising Asians as carriers, some Asians chastising Africans as carriers, and some Africans believing that the virus solely infected Caucasians. Brennan et al. [2020] found that, in their sample, 88% of the misinformation originated from SM news.

Health-related fake news is a particular concern. A study conducted on the spread of fake medical news in Poland found that about 40% of the medical news shared on SM was fake. The highest percentage of fake news stories were on cancer followed by neoplasm, vaccinations, heart attack, and HIV/AIDS [Waszack et al. 2018].

Fake news and politically motivated exaggeration of news has also been identified in a number of major social and political issues in the UK including Brexit, benefits fraud, and climate change [Gavin 2018]. A kind of "fake news" falsification of identities or "soft puppet" Twitter accounts were used in the months leading up to the Brexit referendum. Most of the messages from these fake accounts were for leaving the EU. The researchers stated that the volume of information from false accounts provided a false amplification effect [Bastos and Mercea 2019]. This is analogous to a false consensus heuristic originally proposed by Tversky and Kahneman [1973].

In 2019, news of a child trafficking and pornography ring out of the Comet Ping Pong pizza restaurant in Washington, DC went viral. The news scandal, referred to as Pizzagate, was alleged to being run by Hillary Clinton and John Podesta, a former White House Chief of Staff. The accusations originated on 4Chan, an online bulletin board that was passed as news on SM sites such as Facebook and YouTube. The incident resulted in death threats to the owner of the pizza parlor and eventually, a shooting at the pizza parlor by one of the followers of the fake news. An investigation by law enforcement found no evidence of any child trafficking, but despite reports of no wrongdoing, both the stories and threats continued [Kang and Goldman 2016].

Recently, researchers from different fields have started paying more attention to the role of anchoring on fake news with the SM context. In an experimental design, Jost et al. [2019] established a direct effect between anchoring and fake news. The study notes that anchoring was particularly problematic in the context of SM fake news as readers of the fake news could still be influenced by the anchor even when they knew the information was false. Similarly, Bago et al. [2020] notes that anchoring effects on political misinformation and fake news was moderated by the political implications of the news or the degree to which the news aligned with the reader's political belief.

2 THEORETICAL OVERVIEW

2.1 Prospect Theory and IBA

The prospect theory of Kahneman and Tversky [1979] adds theoretical credence to the IBA tapestry. Prospect theory offers two fundamental premises that shed light on IBA. First, individuals do not assess utility or value based on absolutes but on relative differences to individual baselines (e.g., behavioral biases, proclivities, interests, partisanship). Second, individuals are more risk-averse about gains and risk-takers averse related to uncertainty, since individuals tend to dislike losses more than they like gains. Thus, when facing uncertainty, human behavioral heuristics regress toward inherent biases.

According to Tversky and Kahneman [1973], important, complex decisions based on beliefs concerning uncertain events are reduced to simple judgmental heuristics that build biases and can lead to severe and systematic errors. The authors remark how "people apply heuristic rules to their fallible impressions... rarely aware of the basis of their impressions and they have little deliberate control over the processes by which those impressions are formed."

Consider a Windows XP user asked to upgrade to the latest version of Microsoft Windows 10 OS that offers greater reliability and less cybersecurity vulnerabilities. The user might be hesitant or even reluctant about

Fig. 1. Information bias anchoring model.

upgrading because (i) the user might consider the uncertainty related to the learning costs of installing and migrating from a long-term use of Windows XP (the baseline status quo), however unreliable, and (ii) feel more risk-averse about the upgrade and more risk-seeking about maintaining Windows XP as the status quo, even when aware of cybersecurity vulnerabilities inherent to Windows XP.

The same uncertainty-driven bias from loss aversion is palpable in fake news and IBA. Individuals with stronger biases and partisanship are more likely to display a more defined baseline, thus prompting greater inertia in assessing SM news information. SM news information that is closer to the baseline will be more easily accepted and anchored, thus compounding IBA behaviors. Similarly, even when faced with more rational and unbiased SM news information, individuals with stronger biases and partisanship are more likely to discredit SM news information, however factual. This reflects how individuals, when conditioned by baselines, tend to be risk-averse toward positive factual SM news information, while tending to be risk-taking by capitulating to more biased or partisan news, reflecting loss-aversion, regardless of credibility.

3 IBA MODEL

Building on the background literature pertaining to information anchoring bias and fake news, this study proposes an IBA model (Figure 1). The model delineates SM news information velocity, variety, and anonymity as key characteristics that coalesce into news information overload. The overload generates news information uncertainty as users struggle to discern trust from fake news.

Prior research notes that information uncertainty within the SM triggers the need to seek out additional information to resolve the conflicting information [Karduni et al. 2018]. As a result, users are compelled to process information more systematically in a bid to find answers [Tiedens and Linton 2001]. Given the high costs of acquiring accurate information in an uncertain environment [Vaccari and Chadwick 2020], users simply apply cognitive coping mechanisms that introduce individual biases [Tiedens and Linton 2001]. Consequently, faced with uncertainty, individual biases and trust in AI DPL emerge as filtering mechanisms culminating in information bias anchoring.

Figure 1 provides the proposed model underlying our study. The subsequent sections explain the rationale of the model and discuss specific propositions.

3.1.1 SM News Information and Information Overload Intensity

We live in an information sharing economy, replete with burgeoning news creation, consumption, and propagation. According to Pew Research [2018], more than 68% of SM users use SM for news consumption. However,

with tens of thousands of news information (articles, opinions, videos, snippets, tweets, and retweets) created and published by a multitude of sources of varying credibility and origin every day, SM news is characteristic of its velocity (rate of new news information), variety (divergence in opinions), and anonymity (unknown source identities). Increasing SM news information velocity, variety, and anonymity coalesce into SM news information overload.

It has been long known that an individual's decision-making and reasoning performance correlates positively with the amount of information received—but only up to a point. User cognitive processing is limited [Miller 1956] and “brains have difficulty processing all the relevant information” [Mintzberg 1997, p. 37]. If further information is provided beyond this point, the individual's performance rapidly declines [Chewning and Harrell 1990]. Information received beyond this point is no longer integrated into the decision-making process, creating information overload [Chewning and Harrell 1990; O'Reilly 1980].

Velocity. News information velocity is the ratio of news information volume created over time. According to Anderson and De Palma [2012], the information economy produces a surfeit of information compared to what an individual is able to process, characterizing information overload. As Simon [1970] remarks, “A wealth of information creates a poverty of attention.”

With hundreds of terabytes of news information created, shared, and consumed on SM everyday, news information is characteristically “big data.” An essential aspect of Hill et al.'s [2015] remark is the sheer number of sources creating news information. As Leetaru [2019] notes in Forbes, “One of the great ironies of the ‘big data’ revolution is how little we know about the data we use. The world of ‘big data’ might be better called the world of ‘big imagination.’”

Hill et al. [2015] classify big data in terms of velocity (the speed of data generation) and variance (the disparity and dispersion about the same data from different sources) and remark how “preponderance and speed of data from several sources increases velocity and volume but also increases variance and questions the veracity” of the information (*ibid.*, p. 34).

The velocity of a specific SM news topic is operationalized as the rate, i.e., the volume of news generated related to the specific news topic over a certain period of time.¹ In examining Computer Mediated Communications, Hiltz and Turoff (1985, p. 680) remark that “when communications volume has built up but users have not had a chance to develop screening skills” within that period of time, users suffer from information overload that underpins their uncertainty and dampens decision-making. Hiltz and Turoff [1985] essentially refers to information velocity, i.e., the amount of information generated within a certain period of time.

Users are likely to be able to process voluminous information when given a longer time to consume information. However, when faced with voluminous information within a short time span, information velocity can trigger uncertainty from feeling overwhelmed and overloaded with information [Kerr and Hiltz 1982]. Inasmuch as increased velocity can cause information overload, requiring a reader to consume a large amount of SM news information in a short period of time, we argue that SM news velocity increases uncertainty.

PROPOSITION 1A. *An increase in the velocity (rate of news information generated per unit of time) of SM news information related to a phenomenon will increase individual perceptions of uncertainty related to the phenomenon.*

Variety. News information variety refers to the heterogeneity, variance, and divergence in interpreting a news phenomenon. Online media is under a hitherto unknown competitive flux, a rapidly growing phenomena with petabytes of news created anonymously at an unprecedented rate, with high levels of divergent information. The number of online news articles has grown exponentially across multiple platforms and channels along with an increased rate of information delivery. This increases information velocity. Bot accounts offer a useful example.

¹For example, assume 30,000 COVID-19 SM news posts from January to March (3 months) and 24,000 COVID-19 SM news posts from April to July (4 months). Operationalized, COVID-19 SM news information velocity was 10,000 COVID-19 SM news per month from January to March and 6,000 COVID-19 SM news per month from April to July.

Bot accounts are used by companies to automatically create polarizing multiple, anonymous Twitter feeds every few minutes with self-retweets across other anonymous accounts, reaching a large swathe of audiences in a very short period of time. As a result, there is little convergence of opinion, leading to an increased variance in news coverage of a single phenomenon, exemplified by growing volatility of sentiments and reactions.

Consider the **GE (General Electric)** fraud allegation on August 14th, 2019. Harry Markopolos, a forensic accountant, published a report alleging that GE, a large American conglomerate, was practicing accounting fraud. The report went viral with more than 7,000 news snippets, with a little more than 60% of news information believing in the fraud and nearly 40% suspecting the fraud allegation as market manipulation. Faced with such news information variance and with GE already facing negative investor sentiment, GE's stock dropped 11% on August 15th, wiping out \$9 billion in GE's market value in one single day. Investors soon learned of the fraudulent claim, but it took 21 days before GE's stock price returned to its pre-market-manipulation-report levels.

PROPOSITION 1B. *An increase in the variety of SM news information related to a phenomenon will increase individual perceptions of uncertainty related to the phenomenon.*

Anonymity. Anonymity is defined as the user's inability to determine the true identity or a source due to the source's assumption of pseudonyms for non-malevolent (e.g., username) or malevolent reasons (e.g., spoofing or phishing to deceive users by assuming a more credible identity).

Today, news sources and consumers are largely anonymous, masked by Twitter handles, assumed usernames, or avatars, simply multiple nameless and faceless entities producing, sharing, and resharing news and opinions as news information on the fly. As Ribeiro et al. (2018, p. 1) remark, "A key characteristic of news on SM is that anyone can register as a news publisher without any upfront cost (e.g., anyone can create a Facebook page claiming to be a newspaper or news media organization)."

Anonymity increases verification costs [Datta and Chatterjee 2008]. Verification is a time-consuming matter, forcing individuals to expend time and effort beyond handling the information overload posed by the speed and variety of news information.

Anonymity amplifies the information overload resulting from volume and variety of news information. As a result, anonymity increases users' cognitive costs-to-scale. Faced with a high velocity and variety of SM news information, source anonymity increases cognitive costs from processing an enormous volume and variety of SM news information for credibility and verification.

The recent spread of SM news related to the COVID-19 pandemic offers a case in point. The case highlights the compounding effects of anonymity on uncertainty. The rapid spread of the coronavirus, totaling over 200 deaths and more than 10,000 infections, has caused doldrums in SM news. Anonymous SM feeds have been crafting fake news remarking on (i) how the Coronavirus was grown in a government laboratory to begin a biological warfare and (ii) how afflicted patients can show "asymptotic shedding," and even (iii) requesting personal information in return for providing information access to infected victims. Such anonymous rumors have led to growing uncertainty and paranoia among the public [McCurry 2020].

Processing anonymity of distributed and assumed/masked/phished source information comes at a high cognitive cost. Cognitive processing of anonymous, assumed, or masked (and phished) news information sources requires expending effort in deciphering the credibility of the anonymous source. Anonymity in the realm of digital news media encompasses a variety of attributes. First, anonymity may be a result of a masking or phishing behavior where a non-credible entity could assume a credible identity by masking itself online or misattributing information from one legitimate source to another.

For example, a Twitter feed could carry the name "The Times of Wellington" or "Fox News Haifa" without being credibly associated with either The Times, Fox News, Wellington, or Haifa. Second, anonymity may be the result of perceived source novelty. In an era where hundreds of thousands of bloggers, podcasters, citizen reporters, and activists can create online digital media and news snippets on the fly, the likelihood of

credibly knowing, identifying, or authenticating each source is simply infeasible. The lack of knowledge, identification, or authentication of an information source adds to perceived anonymity. Third, anonymity occurs where source information is advertently or inadvertently missing owing to SM relays, retweets, and focused editing of content.

In summary, the individual effects of SM news information velocity, variety, and anonymity of uncertainty are non-linear. Hiltz and Turoff [1985] found that information overload perceptions started increasing at an intermediate point of information consumption. As users consume more information, processing costs rapidly increase, from the costs of cognitively parsing (i) a large volume of SM news information within a short time and (ii) a diverse, heterogeneous and often conflicting high information variety. Independently, SM news information velocity and variety imply an increasing non-linear influence on information overload.

The accentuated impact of news information anonymity comes from the additional cognitive processing costs of parsing source credibility in the face of information velocity and variety, leading to a compounding effect. Consider the additional costs of checking the source credibility of every anonymous news information post wholly expending bounded cognitive processing capabilities on reconciling overwhelming SM news information velocity and variety. SM news information anonymity therefore adds costs of checking source credibility frequency for both information velocity and variety.

Thus, we argue that when SM news information input surpasses an individual's information processing capacity, the individual experiences an information overload, manifested as uncertainty (e.g., Tushman and Nadler 1978; Datta and Chatterjee 2008; Hill et al. 2015).

PROPOSITION 1C. *An increase in the anonymity of SM news information sources related to a phenomenon will compound the combined effects of SM news information velocity and variety on individual perceptions of uncertainty related to the phenomenon.*

PROPOSITION 2A. *SM news information uncertainty is a function of the product of SM news velocity and variety and compounded by news-source anonymity such that*

$$\text{SM News information uncertainty} = (\text{News Info}_{\text{velocity}} \times \text{News Info}_{\text{variety}})^{\text{anonymity}}$$

PROPOSITION 2B. *SM News information uncertainty follows a sigmoidal curvilinear function such that initial marginal increases in SM news information velocity, variety, and anonymity exponentially increase uncertainty followed by a plateauing effect on marginal uncertainty once an information overload threshold is achieved.*

3.2 Uncertainty, Psychology, and Technology

The confluence of SM news information velocity, variety, and anonymity create a perfect storm of increased uncertainty.²

According to Datta and Chatterjee (2008, p. 15), “uncertainty as the quandary of the principally unknowable, unpredictable, and uncontrollable future owing to ‘neither ignorance nor complete and perfect information but partial knowledge’ [Knight 2012, p. 199].”

But where do individuals surmise their “partial knowledge”? This research assumes that an individual’s biases (proclivities, partisanship) underpins an individual’s partial knowledge. Datta and Chatterjee [2008] note how, when facing uncertainty, individuals need to rely on heuristics to temper the uncertainty in order to progress in their decision-making.

In the presence of news information uncertainty, two elemental forces of psychology and technology come into play and converge. In conditions of uncertainty, human beings rely on psychology (internal information) and technology (as an external validator) as an uncertainty reduction tactic.

3.2.1 The Role of Psychology and Reliance of Internal Biases. Second, the psychology of news media consumers, based on individual tendencies and proclivities, underpins behavior. Although news media consumers intend to access and read factual information as news, it is well established in literature that individual intentions do not automatically translate into behavioral action indicating that not all individuals follow through on their stated intentions [Ajzen 1985; 1991; Godin and Kok 1996; Venkatesh et al. 2012]. Instead, news media consumers tend to anchor their beliefs to news that accentuates their biases and proclivities, thus leading to IBA.

There has been a multiplicity of research in modeling IBA, notably as (i) an economic account of social actions propounded as value rationality by economic philosophers such as Max Weber, John Rawls, and Amartya Sen; (ii) a behavioral economic choice guided by prospect theory [Kahneman and Tversky 1979]; and (iii) a psychological choice forwarded by Ajzen et al.’s [2004] hypothetical bias.

Economic philosophy commonly implies that IBA is based on value rationality, determined by a conscious belief in a value the user assigns to a behavioral decision based on some ethical, aesthetic, or experiential aspects, independent and regardless of rational intentions or probabilities of success [e.g., Weber 1978].

The prospect theory of Kahneman and Tversky [1979] implies IBA as a function derived from an evaluation of utilitarian “expected value” from competing choices. We argue that confirmation bias is central to users’ assigning “expected value” across competing digital news media choices. Four centuries earlier, in 1620, Francis Bacon referred to confirmation bias as follows: “Once a man’s understanding has settled on something (either because it is an accepted belief or because it pleases him), it draws everything else to support and agree with it. And if it encounters a larger number of more powerful countervailing examples, it either fails to notice them, or disregard them, or makes fine distinctions to dismiss and reject them, and all this with much dangerous prejudice, to preserve the authority of its first Conceptions.”

Confirmation bias has been shown to occur in one of two main ways:

- (1) Selective seeking of information that is consistent with one’s prior beliefs, expectations, and hypotheses [Snyder and Swann 1978; Trope and Bassok 1982; Kunda 1999; Oswald and Grosjean 2004; see Klayman and Ha 1987 and Nickerson 1998 for reviews], a process known as selective searching.

²It could be argued that fake news potentially drives uncertainty. However, at the time “ t_0 ,” the user is unaware that the news is fake. As a result, even when fake news is presented to users, such information will go through the proposed IBA model process as news information characterized in terms of Information velocity, variety, and anonymity. The consumer of the SM news information will have to process the information at a high cognitive processing cost and subsequently (time $t_0 + i$) succumbing to anchoring, inadvertent or otherwise.

- (2) Biased interpretation of ambiguous information in a manner that enhances confidence in one's prior beliefs, expectations, or hypotheses [Klapper 1960; Risen and Gilovich 2007; Zillmann and Bryant 1985], a process known as biased interpretation.

Two theories, cognitive dissonance and affordance theory, reify user confirmation biases in the context of SM news information.

Cognitive dissonance is an internal, attitudinal drive to reconcile one's beliefs and behaviors, especially in the face of inconsistent information. Given that users have limited cognitive capacities [Miller 1956]. When faced with many varied (conflicting) choices from anonymous sources in a short period of time (velocity), users restore cognitive harmony through either selectively seeking supporting information or through ignoring countervailing information in beliefs or between beliefs and behaviors that create dissonance which promotes either a change in attitude or behavior to restore harmony [Festinger 1957]. Cognitive dissonance thus drives both confirmation bias where users readily accept information that supports their views and its complement, disconfirmation bias where users question and refute countervailing information. Thus, cognitive dissonance underscores confirmation bias as uncertainty avoidance, consequently leading to anchoring behaviors.

As formulated by Gibson [2015], Affordance Theory refers to how users perceive and interact with the "affordances" offered by the environment, further explaining how the actions predicated by a proximal environment can shape user perceptions and behavior. SM technology platforms offer a variety of actions as affordances—functions that users perceive as properties that define the way we interact and navigate within the platform. For example, a SM platform may feature affordances such as "Like," "Dislike," "Love," "Share," or "Subscribe" to encourage user interaction. However, such interactions predicate user affinities and proclivities, allowing SM platforms to build their user-preference repository. This user-preference repository or database is then used to filter information to match user preferences, thus creating personalized filter bubbles that further confirmation biases.

Taken together, users have specific cognitive toolkits that vary from individual to another. While one user might have a richer set of cognitive tools that allows for a broader and deeper processing of the proximal environment (digital and physical), another user might carry a sparser set of cognitive tools that relies on heuristics and biases for processing the proximal environment. Notwithstanding individual differences, environmental uncertainty can increase cognitive processing costs, ultimately forcing some degree of reliance on intrinsic biases.

PROPOSITION 3. SM news receivers with higher levels of uncertainty are more likely to effectuate higher levels of trust in receivers' internal biases.

3.2.2 Uncertainty and Reliance on AI DPL. Anonymity, velocity, and variety builds up measurement noise and information overload in the environment, leading to uncertainty [Datta and Chatterjee 2008]. Information overload severely depletes cognitive resources and slow-down decision-making, corresponding to heightened environmental uncertainty.

On the technology front, research has well established that human decision quality suffers from uncertainty once an information overload threshold is achieved [e.g., Datta and Chatterjee 2008]. If further information is provided beyond this point, the decisional performance of an individual rapidly declines [Chewning and Harrell 1990]. AI systems are built to mitigate this apparent uncertainty by deducing human preferences using DPL. Akin to converting uncertain browsers into anchored buyers, AI systems capitalize on the uncertainty to trace patterns of directional search and **Natural Language Processing (NLP)** to define preference profiles that match the user's inherent cognitive biases. This is the buy-in.

In the age of pervasive digitization of content and ubiquitous computing, AI is the emerging heuristic as a mechanism to solve decisional quandaries. In the face of uncertainty, AI uses preferential learning as a principled approach toward mitigating the uncertainty and aid decision-making. After all, "the key distinction between problems in which a probabilistic approach is important and problems which can be solved using

non-probabilistic machine learning approaches is whether uncertainty plays a central role” [Ghahramani 2015, p. 11]. The authors further note: “Uncertainty plays a fundamental role in all of this. Observed data can be consistent with many models, and therefore which model is appropriate given the data is uncertain. Similarly, predictions about future data and the future consequences of actions, are uncertain.” (*ibid.*, p. 12).

The Role of Technology. The economic intersection of AI demand and supply in news media is particularly intriguing. With globalization and technology rapidly flattening the world, digital news media sources over online channels have burgeoned. More than 60% of U.S. news readers get their news online and that number is expected to increase by 6% annually while print newspaper subscriptions have fallen by 8% annually. Digital mobile-device news advertising revenue’s rapid growth also continued in 2018, increasing from \$57 billion in 2017 to \$71 billion in 2018 [Pew Research 2018].

In short, individuals, today, often turn to SM for news information querying and gathering to guide their cognitive schema.

The aforementioned statistics reveal that a majority of individuals today rely on SM news information as a de facto source of facts, often unaware that SM news information is deeply reliant on AI-based technological platforms and algorithms. The study by Pega [2018] of worldwide consumers captures the pervasiveness of AI in society. The study reported that, even though 33% of consumers thought that they were using AI in their everyday technology and Internet services, 77% of consumers were actually using AI. The statement underscores how AI is pervasive in its influence, covertly or overtly. Furthermore, according to a PwC [2017] survey, 63% of businesses believed that AI can offer a superior one-on-one personalized experience.

Interest from companies to personalize and users to rely on SM news information signify an increase in both AI demand and supply. 77% of consumers using AI, wittingly or unwittingly, underpins AI demand while 63% of companies believing that AI can offer a richer customer experience underpins AI supply.

Intuitively, digital news lucrativeness has led to the creation of thousands of “digitally native” (purely digital) news sources served across digital platforms such as Facebook, Google, Twitter, Apple News, Yahoo, and Pandora. With a multiplicity of digital media and digital channels vying for consumer attention and engagement as a preface to advertising revenues, news media competition has heightened to a fever pitch. In this context, the speed of gathering news, processing news, and distributing news have become existential issues for the media industry.

Taken together, the devaluation and demise of print media, the growth of digitally native news sources and channels, and the shifting focus of news advertising in the past decade have particularly deconstructed the media landscape where broader coverage and speed-to-market are prioritized over veracity of the news.

AI DPL and Personalization. The lack of reasoning supports our model because we argue that AI starts reasoning for people when people fail to reason and simply follow their biases. Davis [2019] remarks, “Technology encourages us to believe we can all have first-hand access to the ‘real’ facts—and now we can’t stop fighting about it.”

A vital role played by AI in the news media is in the area of personalization [Allcott et al. 2019]. Information on the Internet is estimated to double every 72 hours, creating a need on the part of the consumer to curate its content [Bhargava 2009]. The overload of news can create a sense of fatigue, or reaction to avoid news. This promotes the desire for individuals to rely on news platforms to curate information for them [Song et al. 2017]. This has led to a proliferation of platforms for social news curation [Schneider et al. 2017]. Content curation is a manual or automatic process of selectively gathering, organizing, and presenting the news in a manner that represents an interest, point of view, or topic [Schneider et al. 2017]. The curator can be a number of different actors including journalists, SM or online news platform, or an individual. It is argued that in order to protect the “democracy of the web,” platforms hosting news should be decentralized and self-regulated to protect the privacy of the curator [Jaing et al. 2014].

Unsurprisingly, prior research has cautioned against the “filter bubble” [Pariser 2011], a mechanism to overcome news information overload by increasingly relying on information curated by like-minded others in the social network [Pentina and Tarafdar 2014]. Arguably, his curation approach contributes to a proliferation of fake news on the Internet.

While traditional media relied on a common, standard content for its readers, AI is able to assign relevant personalization for a sample-of-one. While such personalization appears to be apparently welcome in an age of competition, there is a dark side to such granular personalization.

The structure and operational dimensions of AI DPL are based on minimizing $\hat{y}$ for an example DPL heuristic shown below. The neural network AI heuristic, hereunder, relies on weights W (W_1 and W_2) assigned to individual biases b (b_1 and b_2) to predict the outcome $\hat{y}$. In the scope of this research, each individual bias b would carry a corresponding weight W . The AI would assign each individual a specific W and b based on the individual’s previous news information search and clickstream behaviors (clickstream is the sequence and type of news information clicked by a user to visit the page, share, or download). By doing so, the AI DPL uses neural networks to predict every individual’s personal proclivities, preferences, or biases.

$$\hat{y} = \sigma(W_2\sigma(W_1x + b_1) + b_2)$$

The algorithmic function is a generalized formulation of weighted decision-making based on preference matrices. The AI uses that the output $\hat{y}$ is fed as a “feedforward” to predict individual news information preferences. If the individual only visits a few of the news information sites on the predicted “feedforward” outcome $\hat{y}$ or a different set of news information sites, the AI considers it a “loss” or “deviation” between the AI’s predicted outcome $\hat{y}$ and the individual’s actual behavior y . The AI then checks for common preferences from the news information the individual visits and reads in order to revise and reassign weights W and biases b to minimize the loss function between predicted news information output $\hat{y}$ and the individual’s actual news information preferences y (i.e., minimize $(\hat{y}, y)$ using techniques such as reducing sum-of-squares errors between y and $\hat{y}$).

AI models and algorithms are a function of their users’ inputs. AI algorithms learn based on what is input to decipher user preferences. In the wake of the Hong Kong protests against Chinese mainland control where news media is vastly divided and partisan, AI creates a different preferential learning profile for a user that predominantly clicks on and reads mainland China news versus a preferential learning profile for a user that predominantly clicks on and reads Hong Kong and Western news sites and news. AI uses each candidate’s profile to deliver custom content that matches each individual’s proclivity as a way to mitigate their news information uncertainty. The AI sends the user interested in mainland China news information that plaudits mainland Chinese tolerance toward Hong Kong and China’s attempts at maintaining unity under the “One Country-Two Systems” policy. On the other hand, the AI sends the user interested in Western news information that plaudits Hong Kong protests against mainland Chinese oppression and the unnecessary violence being meted out by mainland China under the pretense of control.

In a 2018 Vox interview³ about AI-based preferential learning on search and SM sites, the interview notes that AI on search and SM sites is not meant to be paternalistic, ideal neutral platforms but as a platform that treats its users as customers. “There is safety in saying the consumer is going to make the decision. Yes, there is an algorithm here, but the algorithm is to the consumer. The algorithm is just saying, “What do you want? Okay, we’re going to give you more of that.”

Given that AI-based search is impersonal, search engines prioritize preference-matching over relevance or truth, creating a misinformed heuristic that the reader believes is true. In trying to alleviate user uncertainty, AI focuses on rational choice with a variance reduction objective. Parkes and Wellman (2015, p. 267) succinctly surmise: “AI strives to construct—out of silicon (or whatever) and information—a synthetic *homo economicus*, perhaps more accurately termed *machina economicus*.”

³<https://www.vox.com/technology/2018/2/19/17020310/tristan-harris-facebook-twitter-humane-tech-time>.

This impersonal AI preference-optimizing algorithm thus condones and accentuates individual biases, leading individuals to anchored AI heuristics, subsequently helping propagate biases, creating a Petri dish for fake news.

PROPOSITION 4. *SM news receivers with higher levels of uncertainty are more likely to effectuate higher levels of trust in AI DPL.*

PROPOSITION 5. *SM news receivers with higher levels of trust in AI DPL are more likely to anchor their information biases.*

While internal biases increase partisanship, such partisanship can be conditioned and tempered by unbiased news from credible media. A traditional non-AI world was equally replete with biases. Yet, there were a few established and mostly credible news sources to offer requisite news information veracity. The shift to consuming news information on SM along with the growing use of AI for DPL has changed information bias anchoring behaviors. With DPL AI using sophisticated neural networks to analyze individual SM news consumption behavior, intrinsic biases coupled with a trust in AI can accentuate rather than temper biases. As remarked earlier, DPL AI can filter, forward, and even create news information that matches our biases and inclinations, creating *echo-chambers*. Accentuated biases increase the loss aversion baseline, leading to more entrenched partisan biases and anchoring behaviors.

PROPOSITION 6. *SM news receivers with higher levels of internal biases combined with higher levels of trust in AI DPL are more likely to anchor their information biases more than SM news receivers with only higher levels of internal biases or SM news receivers with only higher levels of trust in AI DPL.*

3.3 The Moderating Effects of User Information Specificity

Although AI DPL serves as a mitigating cue toward anchoring, anchoring remains incomplete. AI-led anchoring attempts are complemented by the level of user's information specificity, a psychological dimension that moderates the direct impact of AI DPL on the level of anchoring. Simply put, a user with low susceptibility toward fake news will dampen AI DPL effects on the level of anchoring, whereas a user with high susceptibility toward fake news will accentuate or exacerbate AI DPL effects on the level of anchoring. This is the lock-in.

There is a distinct difference between users' interest in an issue versus users' information specificity. User interest and/or involvement on COVID-19 vaccines may be a part of a broader swathe of user interest in a topical area. However, users' information specificity offers a more directed approach toward a specific type of information such as "when are COVID-19 vaccines likely to be accessible?" versus user interest and involvement on COVID-19 vaccines. Thus, we argue that users' information specificity offers a vector, which corresponds with an inherent vector present in fake news.

Suppose two readers are reading about a new cancer drug and its side effects across many SM articles, with a wide variety of conflicting information. If one reader is related to someone suffering from cancer, the reader has high information specificity. Conflicting news about the cancer drug's side effects will increase the reader's uncertainty. The reader with low information specificity might take an arm's length approach to the conflicting news that will not impact the reader's uncertainty.

It can be argued that users that are more knowledgeable may be better positioned to deal with inconsistent information. However, users with high information specificity are neither more knowledgeable nor biased. Rather, high information specificity users place a higher level of importance on specific news topics.⁴ As a result, when

⁴The interplay between information specificity and motivated reasoning is of note. Motivated reasoning, i.e., skewed reasoning based on a certain internal or external motivation, can underpin information specificity. A trial lawyer, for example, is motivated to win a case and will search for "specific" precedents and arguments only in support of the client. As a result, a trial lawyer's information specificity is high but does not essentially, reflect confirmation bias. Were the lawyer to represent an opposing side as a client in another situation, motivated reasoning related to high information would skew toward supporting an opposing viewpoint.

presented with SM news information with high variety (variance), anonymity, and velocity, users will suffer from a heightened state of uncertainty that will inevitably trigger a shift toward trust in AI DPL.

Consider an investor heavily invested in Tesla's stock. A Tesla stockholder has high information specificity related to Tesla news. Frequent and many conflicting newsfeeds, some negative such as Elon Musk, Tesla's chairman, smoking marijuana and some positive such as Tesla production reducing Model 3 backlogs, will create more uncertainty for the reader who is a Tesla stockholder. Owning Tesla stock (assuming that the buyer is pro-Tesla because stock ownership reflects a partisanship for the company) increases the reader's information specificity and not knowing how investors will react to the negative versus positive Tesla news will increase uncertainty.

Partisan attachment is inherent to users' information specificity. Susceptibility to partisan attachment plays a major role in what information is attended to, how it is received, processed, stored, and recalled. For example, motivated reasoning, which is a process whereby one's personal motives and goals direct the attentional focus and memory. James and Van Ryzin [2017] differentiate two primary goals: accuracy and directional. Accuracy goals are non-partisan and involve seeking out relevant information in order to achieve a correct picture of the situation. If motivated by directional, partisan goals, people seek and pay attention to only that information that confirms a prior belief, thus confirming a bias that they already possess. The authors conducted an experiment involving people's reactions to the Affordable Care Act and found that respondent's reactions to the act largely fell along political party identification.

Even knowledgeable users are cognitively lazy [Pennycook and Rand 2019], adding to users' reliance on AI DPL's algorithmic conveniences. In line with the classical reasoning of Pascal's wager presented earlier, "humans are cognitive misers, in that resource-demanding cognitive processes are typically avoided" [Pennycook and Rand 2019, p. 39]. When faced with a multiplicity of choices (owing to news information variety), information overload, and uncertainty, individuals tend to succumb to inherent biases that require less cognitive effort [Pentina and Tarafdar 2014]. Information specificity implicated by hidden layers of cognitive laziness from partisan attachment thus interdicts reasoning efforts and influences a greater reliance on AI DPL to find confirmatory news information that can accentuate inherent biases and reduce reasoning costs.

In short, individuals faced with uncertainty and with high information specificity will often rely more on AI DPL to reason for themselves in order to reduce reasoning costs and simply follow their biases (or information vectors).

PROPOSITION 7. Under conditions of news information uncertainty, users' information specificity influences their trust on AI DPL such that users with high information specificity are more likely to shift their trust to AI DPL compared to users with low information specificity.

4 RATIONALIZING THE IRRATIONAL: DISCUSSION, CONTRIBUTIONS, AND IMPLICATIONS

While there is a preponderance of examples and impacts of information biases and fake news, little is known about what leads to information bias anchoring, and thus susceptibility to fake news. The proposed conceptual research framework bridges the lacunae by offering a theoretical framework on the downstream effects of SM news information overload on uncertainty and the downstream effects of uncertainty on internal biases, external AI-driven information, and subsequently, anchoring behaviors.

This research integrates hitherto fragmented domains of psychology, technology, economics, and SM to surface the rich tapestry of interconnections related to the creation and information anchoring of digital media news.

Our research surfaces the implicit interweaving of the SM news information overload (the environment), cognitive biases (intrinsic psychology), and AI-based DLP (technological intermediation). Taken together, the framework underscores ways to mitigate information bias anchoring and fake news propagation that can be

useful to XAI research and practice. The proposed framework further contributes to our understanding in several ways.

IBAM as an XAI Core. This proposed IBAM adds to the rich body of XAI by highlighting the process and intersections of human bias and technical reliance and interventions. The proposed IBA model helps decipher the confluence of AI and psychological black boxes of information consumption. Understanding this process is beneficial to the XAI research corpus.

Germane to our discussion and framework are the design and role of XAI in the domain of SM news information. Among the portfolio of outcomes, XAI seeks to bridge AI design and development and human behavior and decision-making. However, companies, groups, communities of practice, and individuals design the initially embedded and often overarching and inherited AI logic. This introduces an inherent AI design paradox where the designers' innate biases can skew the AI. In discussing the role of software and human agents in knowledge-creation and transformation processes, Datta [2007] discussed how an agent's inherent worldview influences the software embedded logic, creating a design paradox that is fundamentally biased. Miller et al. [2017] explains this paradox by highlighting how AI design coordination relies on social attribution reliant on in-group belief structures or sample-dependent causal connections where "discount some events and not consider their counterfactuals, while being consistent with what an explainee would expect" (*ibid.*, p. 3). Thus, sample choice, socio-lingual variations can severely skew AI training and XAI is relegated to discounting certain inherent biases.

While Zhou et al. [2020] offer an excellent discussion of fake news, the primary focus of the research is the news content. Zhou et al. [2020] parameterize SM news information content in terms of (i) false knowledge, (ii) writing style (e.g., sensationalist, provoking), (iii) propagation patterns, and (4) source and sharer credibility.

This research offers a theoretical framework that underscores the transparency central to XAI design and research. Our research offers a complementary cue by surfacing the fundamental attributes rather than the content of SM news information, asserting that these combined attributes create a vicious environment that increases reader uncertainty from information overload, thus triggering greater reliance on one's internal biases that AI algorithms accentuate. Given that AI algorithms can be easily trained to detect human biases (e.g., from browsing, search streams), the ability to delineate bias prefaces AI's ability to deliver personalized news that reinforces corresponding biases. This compounded effect amplifies biases, leading to anchoring behaviors and promoting echo-chambers. Thus, even biased news can often serve as a gateway to fake news.

Inculcate an Interdisciplinary Approach. The proposed theoretical framework integrates hitherto fragmented areas of knowledge across economics to surface individual behavior as a confluence of multiple vectors in play. One of the issues in studying Fake News is the fragmentary nature of the research spread across many different academic disciplines such as economics, communication and journalism, political science, public policy, consumer behavior, technology, and psychology to name a few. The diversity of views and perspectives has been disruptive to progress in the study of this phenomenon. Different disciplines have their own jargon and language, which has made it difficult to find commonalities and replication of research findings. In an attempt to provide a more coherent approach to the study, a number of researchers have called on the need for multidisciplinary approaches. As an example, one of the biggest challenges has been to come up with a common definition of "Fake News." An illustration of this is the publication in *Science Monthly*, "The science of fake news: Addressing fake news requires a multidisciplinary approach" [Lazer et al. 2018] where 16 researchers across multiple disciplines came together to develop a common definition. Another example of addressing this issue is universities, professional societies, and institutions hosting events on the topic to gather researchers from different universities and academic areas to discuss and share research approaches and findings, such as "Fake News": Causes, Effects, Solutions, sponsored by the **Center for Ethics and Human Values (CEHV)** at The Ohio State University in October of 2018. This involved researchers from eight different academic institutions and nine different disciplines. While these efforts are laudable, we feel that in order to better understand this phenomenon, models

such as the one we propose that present Fake News from a multidisciplinary perspective are crucial for XAI advancements.

Increase Transparency and Reduce Uncertainty. Serious steps need to be taken to reduce news information anonymity and increase transparency: While anonymity online has afforded a voice to anyone around the world, it has masked source credibility, leading to spoofing and phishing behaviors that can sway thinking and disrupt operations. In particular, news information anonymity becomes Damocles' sword. While anonymity allows for the swift creation and delivery of news information, anonymity increases skepticism and places credibility at risk. The framework shows how anonymity exponentially increases cognitive processing and information overload, leading to a sense of uncertainty. Understanding that reducing uncertainty is a crucial aspect in relying on internal biases and external AI DPL, this research highlights the urgent need to increase source transparency without sacrificing anonymity.

It is imperative that we reduce news information velocity. A lot of what is published as news information on SM is a variation of opinions, judgments, blogs, without much justification, reasoning, or logic. Requiring credibility is a preface to reducing volume over time, and thus velocity of SM news.

Technology policies and platforms can aid reducing information velocity. One approach is mirroring best practices from existing industry protocols such as the ones from **Payment Card Industry (PCI)** SSL encryption, and **certification agencies (CAs)**.

News sites need to reduce anonymity and check for news-source credibility using things such as CAs commonly used in online transactions to verify business identity. Another option is to create a hypertext system similar to https (s stands for secure connection). A verified media article site could have a hypertext such as "*httpm*" (m standing for verified media). All other published information can then be marked as "Opinion Pieces; Check for Verification." This can dramatically reduce uncertainty.

Another option is to create a news information clearing house that works like an ACH as a "news information firewall" with predesignated filtering mechanism algorithms that can ensure source and content credibility before distribution. Companies such as Facebook have discussed requiring more legislative guidelines for standardized news information distribution. Creating a news ACH gateway along with a "*httpm*" designation that verifies the source credibility of any news content can temper issues with algorithmic biases, thus reducing both news information velocity and its consequent impact on uncertainty.

Retrain AI Logic to Reduce Bias. The growing importance of XAI lends credence to retraining AI and humans to reduce compounding cognitive skews with AI. However, recent research on SM news strategy proposes how "media systems will need to tune their published content according to not only the popularity/recency of news-stories, but also the user's browsing habit and preferred information diet" [Chakraborty et al. 2015, p. 4].

To combat compounded biases, we argue that news generators and agglomerators should incorporate XAI in design. Instead of using AI for DPL (DPL) to highlight news we like, technology companies should use DPL to understand the user's information specificity. If the AI knows what information is important, the AI can offer more "unbiased" relevant and true search results related to the topic.

Finally, "with so many news sources available, it is essential that journalists thoroughly review their sources of information before publishing news stories. But, in an industry with such tight deadlines and reduced resources, this is now more of a challenge than ever before" [MyNewsDesk 2017] Of the 3000 journalists surveyed, 79% and 65% of journalists, with less than 7 years of experience, used Facebook and Twitter for news reporting, respectively. With journalists' reliance on Facebook and Twitter as news sources, AI algorithms stand to pronounce the journalists' inherent biases, thereby perpetuating information biases. In addition, if journalists meant to increase news information credibility end up perpetuating information biases, a vicious cycle ensues.

Underpinning XAI is the need to understand various paradoxes of choice. There is a paradox of choice between AI deciding on SM news information choices versus individuals deliberately choosing news information.

Consider the privacy paradox, where users can search incognito to reduce AI algorithms predicting user biases. However, such a choice results in a loss of convenience as an opportunity cost. Thus, even privacy-minded users opt for convenience, therefore triggering the proposed IBA model.

Similarly, future research can shed light on an optimal balance between what news information we should consume versus what news information we want to consume. On the one hand, the boundary of AI-driven decisions may mitigate human biases but impose the AI's own preferential-learning logic, depriving us of certain information and creating "filter bubbles." On the other hand, human deliberation on choices may simply accentuate biases, creating "echo-chambers." Thus, it is significant how SM delineates and curates new. Twitter's recent policy offers a compromise where all tweets are presented to the user while remarking a warning on misinformation tweets.

Another area of future research is emerging technology that compounds common illusory beliefs that further underscores biases and anchoring behaviors. Deep fakes offer a case in point. Deep fakes are an advanced and complex superimposition of false images, videos, or voices over real source material using AI deep learning meant to deceive the user. Deep fakes use generative AI to supplant faces, voices, and images, intensifying common illusory beliefs that form the core of biases and can augment users' information specificity owing to more partisan feelings. With newer immersive technologies such as **augmented reality (AR)** meant to heighten experience, a combination of AR and deep fakes can compound individuals', specifically children's common illusory beliefs from an early age. Our increasing appetite for consuming online content makes it easy to create and deliver deep fakes on existing and AR devices that can amplify biases, partisanship, and anchoring.

For policymakers and developers to successfully spot and stymie fake news information in SM news, research and practice needs to respond to these questions prior to building algorithms and training models to solve this grave issue in a world where individuals appear to succumb more and more into relying on SM platforms to make sense of the world.

Reduce Intrinsic Biases. Society and policymakers should expend significant efforts to reduce intrinsic biases. Finally, reducing uncertainty is the most important aspect in reducing information bias and fake news propagation. Psychologically, there are several early childhood and adult interventions. Psychological antecedents to anchoring start in childhood with parental beliefs about what is real and what is make-believe and whether beliefs must be predicated on facts or whether fiction, fantasy, and faith is sufficient. Parental encouragement of fantasy and make-believe is due to several reasons. Parents reinforce a child's skill in pretending as a way to cope with the fictions inherent within the culture and society [Fisher and Fisher 1993]. Parents are motivated to encourage children's make-believe as a way to maintain control over them [Sherrod and Singer 1979]. The parent's own beliefs influence their children's beliefs, which persist into adulthood.

Linguistics further shape intrinsic biases [Costa-jussà 2019]. The author discusses how AI-driven NLP copy and amplify social biases from the conversational and written lingual biases. Given that socially evolved language is the basis for training AI, this raises important questions for future research. Should NLP (including speech recognition), informal conversations, tweets, and SMs be used to train AI and thereby, incorporate the inherent idiosyncrasies of the communication medium? Alternatively, should only formal communication, often edited and even censored, be the sole training set to reduce AI bias? The former can increase bias, given the preponderance of alternative communication media using emojis and messaging shorthands, thus unbiasedly reflect inherent socio-linguistic biases. On the other hand, training AI using more formal language can reduce socio-linguistic biases but remove the authenticity and credibility of socio-linguistic communications.

To summarize, this research focuses on XAI theory-building to surface a cognitively evaluative model that traces the socio-technical intersections behind AI and bias-driven consumption of SM news information. While this research has attempted to ground itself in multidisciplinary theoretical logic, a future empirical investigation could add to model validity. Thus, this research lays the foundation for an empirically validation of the proposed

hypotheses. Future empirical investigations can offer granularity on the type of SM news combined along with each sample's linguistic and cultural mores.

5 CONCLUSION

In an age where most people rely on SM news generated at unprecedented anonymity and frequency, society has to be vigilant about fake news anchoring. This article shows the process by which, in uncertain times, AI leverages individuals' internal biases to drive fake news and entrench biases even further.

Our research offers an initial, albeit important step toward building a theoretical XAI model that integrates the intersections of SM attributes, human behavior, and AI.

In this article, we integrate economic and behavioral perspectives to build a theoretical model of intention-behavior discrepancies. This research, grounded in the aforementioned theories, integrates and extends our understanding of IBA using the lens of psychological bias. Sen [2009], the Nobel Laureate, remarks on the faulty nature of social behavior that can create discrepancies between rational intentions and behaviors as driven by prejudices [biases]: "... prejudices typically ride on the back of some kind of reasoning—weak and arbitrary though it might be. Indeed, even very dogmatic persons tend to have some kinds of reasons, possibly very crude ones, in support of their dogmas" (p. 18). Thus, "of course, decisions are made in a minefield of potential bias, some that increase and others that decrease the valuations [of intentions] (emphasis added) [Kogut and Kulatilaka 2004, p. 104].

In a world replete with technology, the focus must be on reasoning that is less suggestive of and pronounced by biases, lest we spin the perfect storm out of control. For a more equitable and just news information dissemination, the need, therefore, lies in XAI designs driven not by prediction but by justification.

REFERENCES

- I. Ajzen. 1985. From intentions to actions: A theory of planned behavior. In *Action Control*. Springer, Berlin, pp. 11–39.
- I. Ajzen. 1991. The theory of planned behavior. *Organizational Behavior and Human Decision Processes* 50, 2 (1991), 179–211.
- I. Ajzen, T. C. Brown, and F. Carvajal. 2004. Explaining the discrepancy between intentions and actions: The case of hypothetical bias in contingent valuation. *Personality and Social Psychology Bulletin* 30, 9 (2004), 1108–1121.
- H. Allcott, M. Gentzkow, and C. Yu. 2019. Trends in the diffusion of misinformation on social media. *Research & Politics*. Retrieved on July 19, 2020 <https://doi.org/10.1177/2053168019848554>.
- H. Allcott and M. Gentzkow. 2017. Social media and fake news in the 2016 election. *Journal of Economic Perspectives* 31, 2 (2017), 211–236.
- S. Anderson and A. De Palma. 2012. Competition for attention in the Information (overload) Age. *The RAND Journal of Economics* 43, 1 (2012), 1–25.
- B. Bago, D. G. Rand, and G. Pennycook. 2020. Fake news, fast and slow: Deliberation reduces belief in false (but not true) news headlines. *Journal of Experimental Psychology: General* 149, 8 (2020), 1608–1613.
- E. Bakshy, S. Messing, and L. A. Adamic. 2015. Exposure to ideologically diverse news and opinion on Facebook. *Science* 348, 6239 (2015), 1130–1132.
- M. T. Bastos and D. Mercea. 2019. The Brexit botnet and user-generated hyperpartisan news. *Social Science Computer Review* 37, 1 (2019), 38–54.
- R. Bhargava. 2009. The manifesto for the content curator: The next big social media job of the future. *SocialMediaToday*. 2009. Retrieved from on July 19, 2020 <https://www.rohitbhargava.com/2009/09/manifesto-for-the-content-curator-the-next-big>.
- J. C. Brennen, F. M. Simon, P. N. Howard, and R. K. Nielsen. 2020. Types, sources, and claims of COVID-19 misinformation. *Factsheet*, Reuters Institute for the Study of Journalism. University of Oxford. Retrieved August 8, 2020 from <https://reutersinstitute.politics.ox.ac.uk/types-sources-and-claims-covid-19-misinformation#methodological>.
- S. Brennen, F. Simon, P. Howard, and R. Nielsen. 2020. Types, sources, and claims of COVID-19 misinformation. Reuters Institute for the Study of Journalism & University of Oxford. Retrieved on July 19, 2020 <https://reutersinstitute.politics.ox.ac.uk/types-sources-and-claims-covid-19-misinformation>.
- Ceren Budak, Sharad Goel, and Justin M. Rao. 2016. Fair and balanced? Quantifying media bias through crowdsourced content analysis. *Public Opinion Quarterly* 80, S1 (2016), 250–271.
- A. Chakraborty, S. Ghosh, N. Ganguly, and K. P. Gummadi. 2015. Can trending news stories create coverage bias? On the impact of high content churn in online news media. In *Computation and Journalism Symposium*.

- A. Chakraborty, S. Ghosh, N. Ganguly, and K. P. Gummadi. 2016. Dissemination bias of social media channels: On the topical coverage of socially shared news. In *Proceedings of the 10th International AAAI Conference on Web and Social Media (ICWSM'16)*.
- E. C. Chewing, Jr. and A. M. Harrell. 1990. The effect of information load on decision makers' cue utilization levels and decision quality in a financial distress decision task. *Accounting, Organizations and Society* 15 (1990), 527–542.
- M. R. Costa-jussà. 2019. An analysis of gender bias studies in natural language processing. *Nature Machine Intelligence* 1 (2019), 495–496.
- P. Datta and S. Chatterjee. 2008. The economics and psychology of consumer trust in intermediaries in electronic markets: The EM-trust framework. *European Journal of Information Systems* 17, 1 (2008), 12–28.
- P. Datta. 2007. An agent-mediated knowledge-in-motion model. *Journal of the Association for Information Systems* 8, 5 (2007), 20.
- W. Davis. 2019. Why can't we agree on what's true any more? *The Guardian*, September 19th.
- L. Festinger. 1957. *A theory of Cognitive Dissonance*. (9th ed.). Stanford CA: Stanford University Press.
- W. A. Fisher and J. D. Fisher. 1993. A general social psychological model for changing AIDS risk behavior. In *The Social Psychology of HIV Infection*, J. Pryor and G. Reeder (Eds.). Erlbaum, 27–53.
- N. T. Gavin. 2018. Media definitely do matter: Brexit, immigration, climate change and beyond. *The British Journal of Politics and International Relations* 20, 4 (2018), 827–845.
- J. F. George, K. Duffy, and M. Ahuja. 2000. Countering the anchoring and adjustment bias with decision support systems. *Decision Support Systems* 29, 2 (2000), 195–206.
- J. J. Gibson. 2015. *The Ecological Approach to Visual Perception*. Psychology Press Classic Editions, New York, NY.
- G. Hill, P. Datta, and W. Acar. 2015. Shifting perspectives: A process model for sense making under uncertainty. *IJSDS* 6 (2015), 33–52.
- Z. Ghahramani. 2015. Probabilistic machine learning and artificial intelligence. *Nature* 521, 7553 (2015), 452–459.
- G. Godin and G. Kok. 1996. The theory of planned behavior: A review of its applications to health-related behaviors. *American Journal of Health Promotion* 11, 2 (1996), 87–98.
- S. R. Hiltz and M. Turoff. 1985. Structuring computer-mediated communication systems to avoid information overload (PDF). *Communications of the ACM* 28, 7 (1985), 680–689.
- A. H. Jaing, Z. S. Bischof, and F. E. Bustamante. 2014. A Cliq of content curators. SIGCOMM'14, October 17, Chicago, Illinois. Retrieved on July 19, 2020 <http://dx.doi.org/10.1145/2619239.2631470>.
- O. James and G. G. Van Ryzin. 2017. Motivated reasoning about public performance: An experimental study on how citizens judge the affordable care act. *Journal of Public Administration Research and Theory* 27, 1 (2017), 197–209.
- P. J. Jost, J. Pünder, and I. Schulze-Lohoff. 2020. Fake news—Does perception matter more than the truth? *Journal of Behavioral and Experimental Economics* 85 (2020), 101513.
- D. Kahneman and A. Tversky. 1979. Prospect theory: An analysis of decision under risk. *Econometrica* 263–292.
- C. Kang and A. Goldman. 2016. In Washington D.C. fake news brought real guns. *The New York Times*, 2016. Retrieved on July 19, 2020 <https://www.nytimes.com/2016/12/05/business/media/comet-ping-pong-pizza-shooting-fake>.
- A. Karduni, R. Wesslen, S. Santhanam, I. Cho, S. Volkova, D. Arendt, and W. Dou. 2018. Can you verify this? studying uncertainty and decision-making about misinformation using visual analytics. In *Twelfth International AAAI Conference on Web and Social Media*.
- E. B. Kerr and S. R. Hiltz. 1982. *Computer-Mediated Communication Systems: Status and Evaluation*. Academic Press, New York.
- A. Kim and A. Dennis. 2018. Says who?: How news presentation format influences perceived believability and the engagement level of social media users. In *Proceedings of the 51st Hawaii International Conference on System Sciences*.
- Z. Kunda. 1999. *Social cognition: Making sense of people*. MIT press, Cambridge, MA.
- J. T. Klapper. 1960. *The effects of mass communication*. New York, Free Press.
- J. Klayman and Y. W. Ha. 1987. Confirmation, disconfirmation, and information in hypothesis testing. *Psychological Review* 94, 2 (1987), 211.
- F. H. Knight. 2012. *Risk, uncertainty and profit*. Courier Corporation.
- B. Kogut and N. Kulatilaka. 2004. Response: Real options pricing and organizations: The contingent risks of extended theoretical domains. *The Academy of Management Review* 29, 1 (2004), 102–110.
- D. M. Lazer, M. A. Baum, Y. Benkler, A. J. Berinsky, K. M. Greenhill, F. Menczer, and M. Schudson. 2018. The science of fake news. *Science* 359, 6380 (2018), 1094–1096.
- K. Leetaru. 2019. Terabytes of news versus one trillion tweets: Is social media really as big as we think? *Forbes*, September 12th. Retrieved on July 19, 2020 <https://www.forbes.com/sites/kalevleetaru/2019/02/12/terabytes-of-news-versus-one-trillion-tweets-is-social-media-really-as-big-as-we-think/#6ab1164029cc>.
- J. McCurry. 2020. South Korea cracks down on fake news about spread of coronavirus. *The Guardian*, January 2020. Retrieved January 31, 2020 from <https://www.theguardian.com/world/2020/jan/30/south-korea-cracks-down-on-fake-news-about-spread-of-coronavirus>.
- H. Mintzberg. 1997. The manager's job: Folklore and fact. *Leadership: Understanding the Dynamics of Power and Influence in Organizations*. 35–53.
- T. Miller, P. Howe, and L. Sonenberg. 2017. Explainable AI: Beware of inmates running the asylum or: How I learnt to stop worrying and love the social and behavioural sciences. *arXiv preprint arXiv:1712.00547*.

- MyNewsDesk. 2017. Fake news, financial uncertainty, and lack of trust—What journalists say about the future of media, September 14th. Retrieved on October 15, 2019 from <http://www.mynewsdesk.com/us/mynewsdesk/pressreleases/fake-news-financial-uncertainty-and-lack-of-trust-what-journalists-say-about-the-future-of-media-2145386>.
- R. S. Nickerson. 1998. Confirmation bias: A ubiquitous phenomenon in many guises. *Review of General Psychology* 2, 2 (1998), 175–220.
- E. R. Norman and R. Delfin. 2012. Wizards under uncertainty: Cognitive biases, threat assessment, and misjudgments in policy making. *Politics & Policy* 40, 3 (2012), 371–378.
- M. E. Oswald and S. Grosjean. 2004. Confirmation bias. *Cognitive Illusions: A Handbook on Fallacies and Biases in Thinking, Judgement and Memory*. 79.
- C. A. O'Reilly. 1980. Individuals and information overload in organizations: Is more necessarily better? *Academy of Management Journal* 23 (1980), 684–696.
- E. Pariser. 2011. The filter bubble: How the new personalized web is changing what we read and how we think. New York, Penguin Press.
- D. C. Parkes and M. P. Wellman. 2015. Economic reasoning and artificial intelligence. *Science* 349, 6245 (2015), 267–272.
- PEGA. 2018. What consumers really think about AI: A global study. Retrieved July 19, 2020 from <https://www.pega.com/ai-survey>.
- G. Pennycook and D. G. Rand. 2019. Lazy, not biased: Susceptibility to partisan fake news is better explained by lack of reasoning than by motivated reasoning. *Cognition* 188 (2019), 39–50.
- I. Pentina and M. Tarafdar. 2014. From “information” to “knowing”: Exploring the role of social media in contemporary news consumption. *Computers in Human Behavior* 35 (2014), 211–223.
- Pew Research. 2018. News use across social media platforms, September. Retrieve on June 25, 2020 from <http://www.journalism.org/2018/09/10/news-use-across-social-media-platforms-2018/>.
- Filipe N. Ribeiro, Lucas Henrique, Fabricio Benevenuto, Abhijnan Chakraborty, Juhi Kulshrestha, Mahmoudreza Babaei, and Krishna P. Gummadi. 2018. Media bias monitor: Quantifying biases of social media news outlets at large-scale. In *Proceedings of the 12th International AAAI Conference on Web and Social Media*.
- PwC. 2017. Consumer Intelligence Series: Bot.me. Retrieved on July 19, 2020 from <https://pwc.to/2suROx0>.
- J. L. Risen and T. Gilovich. 2007. Target and observer differences in the acceptance of questionable apologies. *Journal of Personality and Social Psychology* 92, 3 (2007), 418.
- H. Ruokolainen and G. Widén. 2020. Conceptualising misinformation in the context of asylum seekers. *Information Processing and Management*, 57, 3 (2020), 102127.
- D. Schneider, L. F. Oliviera, and J. Souza. 2017. Designing, building, and evaluating a social news curation environment using the action design research methodology. *Cluster Computing* 20 (2017), 1731–1748.
- A. K. Sen. 2009. The idea of justice. Harvard University Press.
- L. Sherrod and J. Singer. 1979. The development of make-believe play. In *Sports, Games, and Play*, J. Goldstein (Ed.). Lawrence Erlbaum Associates, 1e28.
- S. Shane. 2017. From headline to photograph, a fake news masterpiece. *The New York Times*, 2017.
- C. Silverman. 2016. This analysis shows how viral fake election news stories outperformed real news on Facebook. *BuzzFeed News*, November 16, 2016.
- C. Silverman and J. Singer-Vine. 2016. Most Americans who see fake news believe it, new survey says. *BuzzFeed News*, 2016.
- H. A. Simon. 1970. The axiomatization of physical theories. *Philosophy of Science* 37, 1 (1970), 16–26.
- H. Song, J. Jung, and Y. Kim. 2017. Perceived news overload and its cognitive and attitudinal consequences for news usage in South Korea. *Journalism and Mass Communications Quarterly* 94, 4 (2017), 1172–1190.
- M. Snyder and W. B. Swann. 1978. Hypothesis-testing processes in social interaction. *Journal of Personality and Social Psychology* 36, 11 (1978), 1202.
- E. C. Tandoc, Jr., Z. W. Lim, and R. Ling. 2018. Defining “fake news” a typology of scholarly definitions. *Digital Journalism* 6, 2 (2018), 137–153.
- L. Z. Tiedens and S. Linton. 2001. Judgment under emotional certainty and uncertainty: The effects of specific emotions on information processing. *Journal of Personality and Social Psychology* 81, 6 (2001), 973–988.
- R. Torres, N. Gerhart, and A. Negahban. 2018. Epistemology in the era of fake news: An exploration of information verification behaviors among social networking site users. *ACM SIGMIS Database: The DATABASE for Advances in Information Systems* 49, 3 (2018), 78–97.
- Y. Trope and M. Bassok. 1982. Confirmatory and diagnosing strategies in social information gathering. *Journal of Personality and Social Psychology* 43, 1 (1982), 22.
- M. L. Tushman and D. A. Nadler. 1978. Information processing as an integrating concept in organizational design. *Academy of Management Review* 3, 3 (1978), 613–624.
- A. Tversky and D. Kahneman. 1973. Availability: A heuristic for judging frequency and probability. *Cognitive Psychology* 5, 2 (1973), 207–232.
- C. Vaccari and A. Chadwick. 2020. Deepfakes and disinformation: Exploring the impact of synthetic political video on deception, uncertainty, and trust in news. *Social Media+ Society* 6, 1 (2020), 2056305120903408.
- V. Venkatesh, J. Y. Thong, and X. Xu. 2012. Consumer acceptance and use of information technology: extending the unified theory of acceptance and use of technology. *MIS Quarterly*, 157–178.
- S. Vosoughi, D. Roy, and S. Aral. 2018. The spread of true and false news online. *Science* 359, 6380 (2018), 1146–1151.

- P. M. Waszak, W. Kasprzycka-Waszak, and A. Kubanek. 2018. The spread of medical fake news in social media—The pilot quantitative study. *Health Policy and Technology* 7 (2018), 115–118.
- M. Weber. 1978. *Economy and society: An outline of interpretive sociology* (Vol. 1). Univ of California Press.
- X. Zhou, A. Jain, V. V. Phoha, and R. Zafarani. 2020. Fake news early detection: A theory-driven model. *Digital Threats: Research and Practice* 1, 2 (2020), 1–25.
- D. Zillmann and J. Bryant. 1985. *Selective-exposure phenomena*. *Selective Exposure to Communication*. 157–190.

Received November 2019; revised August 2020; accepted October 2020