*Article*

# Investigating the Effects of Misinformation as Infopathogens: Developing a Model and Thought Experiment

Roger D. Magarey [1,*], Thomas M. Chappell [2] and Kayla Pack Watson [1]

1 Center for Integrated Pest Management, North Carolina State University, Raleigh, NC 27606, USA; knpack@ncsu.edu
2 Department of Plant Pathology and Microbiology, Texas A&M University, College Station, TX 77843, USA; tom_chappell@tamu.edu
* Correspondence: rdmagare@ncsu.edu

**Abstract:** Previously, it has been shown that transmissible and harmful misinformation can be viewed as pathogenic, potentially contributing to collective social epidemics. In this study, a biological analogy is developed to allow investigative methods that are applied to biological epidemics to be considered for adaptation to digital and social ones including those associated with misinformation. The model's components include infopathogens, tropes, cognition, memes, and phenotypes. The model can be used for diagnostic, pathologic, and synoptic/taxonomic study of the spread of misinformation. A thought experiment based on a hypothetical riot is used to understand how disinformation spreads.

**Keywords:** infopathogen; infodemic; infopocalypse; disinformation; taxonomy

## 1. Introduction

Misinformation (misleading or incorrect information that may also be disinformation (deliberately deceptive misinformation)) has become one of the most important and least well understood problems in modern society. The World Health Organization coined misinformation about COVID-19 an infodemic and the term has been adopted in the scientific literature (Evanega et al. 2020; Zarocostas 2020; Van Der Linden 2022), while another term, infopocalypse (Fallis 2020), was coined to describe misinformation's potentially harmful effects on society. Proposed defining characteristics of misinformation include factual falseness or contradiction with experts (Li 2020; Vraga and Bode 2020), and actions suggested as ways to reduce harm arising from misinformation often assume that alleged misinformation is correctly defined as such. This assumption can be limiting because it sometimes orients research toward investigating the consequences of spreading certain material, versus classifying material based on its functional consequences. Although misinformation is not a new phenomenon, the information economy has greatly increased the velocity and volume of misinformation. Evidence suggests that exposure to negative political information continues to shape attitudes even after the information has been effectively discredited (Thorson 2016) indicating that the functional outcome of exposure to such information is not determined exclusively by the content of the information and its metadata, but by its interaction with consumers of the information. Algorithmic filter "bubbles" and "echo chambers" associated with social media are a primary gateway through which individuals are exposed to misinformation including fictionalized accounts of current events (Rhodes 2021). The spread and consumption of misinformation has the negative consequence of social network segregation and disruption, and is expected to be increasingly difficult to detect as a result (Chen and Rácz 2021).

Misinformation has substantial impacts upon society, and can be found where potential incentives exist for entities to exploitatively manipulate others through the use of misinformation to influence behavior (Del-Fresno-Garcia 2019), especially concerning

contentious political issues. Purported linkages between vaccination and autism have been identified as contributors to "vaccine hesitancy" and a reduced likelihood to comply with public health guidance (Roozenbeek et al. 2020; Kata 2010). A related example is misinformation about alternative treatments for COVID-19 (Blevins et al. 2021) and cancer (Gorski 2019) that are likely ineffective or even harmful. The example of propagating falsehoods to manipulate others into purchasing medicines is so enduring that numerous terms for it have related etymological origin, e.g., charlatan, quack, montebank. A second example of harmful misinformation is what has been called "fake news", (Fernández-Torres et al. 2021) being fictional accounts of current events promulgated to influence opinion often about organizations or public figures, and which cause lasting, critical, and compounding damage to public institutions (Persily 2017; Rodríguez Fernández 2019).

Misinformation can cause suffering. Transmissible causes of suffering in biological systems are called pathogens. A useful framework for use in studying suffering caused by misinformation can be built from the concept of an infopathogen, which is transmissible information that can result in harm, through processes such as social epidemics (Magarey and Trexler 2020). A useful example system in which to develop this framework and its applications is the COVID-19 pandemic which began in 2020. The COVID-19 pandemic involved a great deal of contention concerning what individuals, groups, policymakers, and policy enforcers should do to reduce several kinds of harm, including risks to individuals and public health (Roozenbeek et al. 2020), and competing agendas created opportunities for the generation and spread of misinformation for various purposes. COVID-19 misinformation has itself been described as an "infodemic" (Evanega et al. 2020; Zarocostas 2020; Van Der Linden 2022) and attributed with affecting individuals' responses to the health risks associated with COVID-19 exposure and infection.

Misinformation is now an important research topic and at least two research agendas have been proposed. The first research agenda (Van Der Linden 2022) includes three topics: (i) susceptibility to misinformation including social, political and cultural factors; (ii) how misinformation spreads in social networks; and (iii) immunization treatments against misinformation. A second paper (Li 2020) identified three research agendas: (i) better understanding the impacts of misinformation; (ii) how individual choices impact misinformation consumption; and (iii) understanding how social context such as designs, policies, and socio-technological infrastructure influences the prevalence of misinformation. In order to address these research questions, it is also important to understand the dynamics of misinformation spread. These dynamics can be understood through the use of a biological analogy, beginning with the idea of harm caused by misinformation spreading as a social epidemic (Magarey and Trexler 2020).

Information, including misinformation, has been shown to be contagious (Jin et al. 2013; Tambuscio et al. 2015; Kucharski 2016) and to play a role in social epidemics (Christakis and Fowler 2013). Building on the concept of the meme, an infopathogen has been defined as harmful information that can spread or intensify a social epidemic (Magarey and Trexler 2020). The definition of a meme in the context of cognition is that a meme is an idea, being complex and a memorable unit, that can be spread by vehicles that are physical manifestations of the meme (Brodie 2009). The term "internet meme" represents audiovisual material or images and graphics with or without superimposed text (Wasike 2022) which are one vehicle by which memes are spread. The rate at which memes (including harmful ones) spread is dependent upon multiple factors such as meme "fitness" (a description of its propagation rate in context), the individual sources' competence, social network structures, societal, contextual, geospatial, and technical factors, and the practical outcomes of meme diffusion (Spitzberg 2014). According to this model, memes compete at multiple levels to occupy information niches. In this paper, we discuss the analogy between biological agents and information/memetic agents, as a basis for analyzing infopathogens and their epidemiology. Critically, we pursue avenues for this analysis that do not rely on the determination of whether something is misinformation a priori, instead placing focus on the characteristics of putative infopathogens that have functional impact on observable

outcomes. The infopathogen framework and the related biological pathogen example enhance translation between domains and can be used to better understand the dynamics of misinformation.

## 2. Materials and Methods

### 2.1. Infopathogen Theory

The concept of an infopathogen is built upon the idea of a harmful meme (or group or memes) that is spread from person to person but more importantly through information networks such as social media. Magarey and Trexler (2020) provide a detailed justification for the analogy that memes behave like biological pathogens. This includes examples of research on violence, demonstrating its ability to spread like an epidemic (Slutkin et al. 2018a, 2018b). Magarey and Trexler (2020) also show that biological concepts such as life cycle diagrams and the epidemiological triad can be adapted to infopathogens, reinforcing the utility of the analogy. Biological mechanisms such as resistance can even be shown to account for supposed differences between the spread of viruses and memes (Lerman 2016; Magarey and Trexler 2020). Finally, another support for the analogy is a growing number of papers advocating psychological inoculation (Van Der Linden 2022; Roozenbeek et al. 2022), confirming inferentially that there must be an information agent against which the inoculation is applied. In the sense of content (especially digital) "going viral", the concept aligns with that of a "mind virus" (Robertson 2017), and because viral content may also comprise non-harmful memes (Brodie 2009), the term infopathogen is used in this study. Criteria for a psychological definition of an infopathogenic infection have been developed (Robertson 2017). These criteria can be summarized as follows: (i) a deleterious, unplanned, and observable change in an individual; (ii) the disruption of normal cognitive feedbacks governing volition, distinctness, continuance, productivity, intimacy, social interest, and emotion; (iii) appropriation of the individual's resources to activities that spread the infopathogen; and (iv) uncharacteristic changes in emotional attachments.

A framework for analyzing and mitigating harmful memes based on the concept of infopathogens provides a starting point (Magarey and Trexler 2020). The framework included the development of an epidemiological tetrad and a lifecycle diagram for infopathogens. The tetrad is based upon the disease triangle, which is a widely used mnemonic that represents the multipartite conditionality of infectious disease manifestation: presence of an agent (the pathogen), a susceptible host (here, the person), and a conducive environment (Gordis 2009). For example, the occurrence of the disease malaria requires a susceptible human host, a disease-causing agent from the Plasmodium group, and environmental conditions conducive to infection. Generalization of the classical disease triangle allows additional conditionalities to be represented, for example the additional requirement of mosquito vectors for the transmission of malaria-causing Plasmodium, and thus also the environmental conditions required for conduciveness to mosquito-based transmission, and host susceptibility to mosquito feeding. The epidemiological triad has been adapted for use in the study of epidemics of non-communicable diseases by including physical agents and vectors in place of biological agents and vectors. The epidemiological triad is expanded to a tetrad by including information agents (i.e., infopathogens) and vectors that can help the roles information technologies and media play in the development of social epidemics.

Another tool for understanding the epidemiology of infopathogens is the diagrammatic representation of zoonotic disease as a life cycle, typically as a directed cyclic graph and used for targeting epidemiological control points for intervention (Lynteris 2017). An example life cycle diagram was developed for the spread of violence caused by neo-Nazism (Magarey and Trexler 2020). A recent review confirmed the veracity of this approach in that misinformation epidemics can be summarized on three dimensions: susceptibility, spread, and immunization (Van Der Linden 2022).

Psychological inoculation (Van Der Linden 2022; Roozenbeek et al. 2022), also known as infovaccination (Magarey and Trexler 2020), cognitive immunity (Roozenbeek and van der Linden 2019), or prebunking (Van Der Linden 2022), is equivalent to the biological

concept of vaccination through which individuals are exposed to weakened or partial host-manipulative pathogens (misinformation), for the purpose of eliminating immunological naïveté without causing the full harm of unmitigated infection. According to Van Der Linden (2022), "psychological inoculation consists of two core components: (1) forewarning people that they may be misled by misinformation (to activate the psychological 'immune system'), and (2) prebunking the misinformation (tactic) by exposing people to a severely weakened dose of it coupled with strong counters and refutations (to generate the cognitive 'antibodies')".

Descriptions of infopathogen dynamics and development of mitigations are useful but exist downstream of the identification of infopathogens. The etiological question of infopathology is similar to that for biological pathology: among microbes, pathogens are few; among information, misinformation and infopathogens may be relatively few. Viruses are one type of pathogen that is well-suited to this biological analogy, because they are strictly dependent on the manipulation of host resources for propagation and spread. The study of biological viruses often involves the sequencing of the viral genome to understand the information it contains, being what is required for the generation of a virus' functional parts—a genome is the complete set of information required to build an organism (Goldman and Landweber 2016). Genomic analysis, or understanding what genes are present and their functions, enables virologists to not only better understand the functions and expected epidemic dynamics of the virus but also to devise mitigations for the purpose of reducing transmission (e.g., exposure management, sanitation) or of increasing immunity (e.g., vaccination). For example, genomic analysis of a viral genome involves several steps, including DNA extraction, shearing, replication, sequencing, and DNA analysis to compare sequenced genes against gene libraries (CDC 2021). Knowing what genes are present in a pathogen helps researchers design specific mitigations. For example, plants can be genetically modified to produce viral coat proteins that will encapsulate viral particles to create disease resistant varieties of plants (Beachy 1997), and vaccines can be developed that deliver incomplete parts (proteins) of viruses determined to be involved in pathogenesis through genomic analysis.

### 2.2. Model Development

To develop the analogy between biological and informational pathogens, a model of five infopathogen-descriptive components is proposed (Table 1). The five components include: (i) infopathogenic (describing infopathogens), (ii) tropic (describing tropes), (iii) cognitive (describing cognitive processes in the form of mind maps), (iv) menomic (describing memes analogous to the biological genome), and (v) phenotypic (describing text and graphics associated with the meme that can be observed and analyzed). Some of these components have been combined in existing tools, including an application to find online trolls (Jachim et al. 2020). The first two components are taxonomic: infopathogens, being identifiable units of transmissible information that have hypothetically pathogenic effects, and tropes, being similar to organismal body plans, or structure-based virus classification. These components can be used for diagnostic, pathologic, and taxonomic study. The next three components are functions: host–pathogen interactions, which in the case of infopathogens are the ways in which information-based behavior is affected. A critical distinction to make between biological and informational pathogens concerns the origins of their information content, which are not directly observable in the host the way that phenotypes are. Biological pathogens occur as a result of natural processes favoring pathogen adaptation to persist, in general. In contrast, informational pathogens' content is generated largely by humans, creating/combining it not necessarily for the persistence of the informational pathogen, but for diverse reasons relating to the intent of the creator. For an example of detecting this intent in distributed creators using what we are here calling phenotypes, see (Kaghazgaran et al. 2019). This distinction emphasizes that though informational pathogens cause dynamics similar to those caused by biological pathogens, informational

pathogens with traits maladaptive for long-term persistence can be repeatedly introduced by humans. Next, each of these components are discussed in detail.

**Table 1.** Proposed five-component model for study of infopathogens.

| | Component | Description | Purpose | Potential Utility |
|---|---|---|---|---|
| 1 | Infopathogenic—analogous to a biological pathogen | Represents infopathogens that can spread harmful information and are a causal agent in social epidemics. | Taxonomic to allow infopathogens to be described, studied, and organized. | Taxonomic classification Case study histories Observed harmful behaviors Specific mitigations |
| 2 | Tropic—analogous to an organismal body plan | Represents tropes that summarize the overall impact of cognitive processes that are influenced by harmful memes. | Synoptic generalization to allow for description of potential harmful behavior and possible mitigations. | Generalizations on observed harmful behaviors Generalized mitigations |
| 3 | Cognitive—analogous to an infected host's morphology | Represents the impact of memes on human cognition and behavior. | Analytic to understand how memes can cause harmful behaviors. | Mind maps linking memes to tropes |
| 4 | Menomic—analogous to a biological genome | Represents memes that can spread harmful information through interpersonal contact and social media. | Informatic to track the flow of harmful information through networks. | Meme frequency Date of first observation Rate of increase/decrease |
| 5 | Phenotypic—analogous to the phenotypic traits | Represents the text and graphics that accompany a meme. | Diagnostic to determine the harmful attributes of a meme. | Sentiment analysis Harmfulness analysis |

Infopathogens in this analogy are composed of memes with meme products, share functional traits that can be understood as tropes, and have causal influence on communicating entities through processes comparable to host–pathogen interactions. Infopathogens are represented by combinations of tropes that can be placed into taxonomic classifications, and studied to understand their memetic complement, origins, and dissemination. A symptomological approach to identifying and characterizing an unknown infopathogen based on the components of this analogy is explored in the following thought experiment.

Tropes in turn are a functional basis for aggregative cognitive processes. The methods useful for such a grouping would be similar to those used in biological systematics, through which the organization of trait similarity is studied for purposes such as inferring evolutionary history. Literary tropes (a form of information) are themes or conventions used to convey concepts efficiently—the convention establishes that the concept's typical context should be assumed. Tropes can be found not just in written and spoken language but in all forms of communication (Baldick 2008). Their efficiency in conveying ideas may explain why the frequency of tropes has increased in movies in the last five years (García-Ortega et al. 2020). Tropes have also been used as a tool to explain common tactics used by the so-called anti-vaccination movement (Kata 2012), or for detecting online trolls (Jachim et al. 2020). Tropes can enhance ascertainment of complex and specific information by signaling to communicating individuals that nuanced context should be assumed; conversely, tropes characterizing infopathogens can be used to subtly enhance urgency or threat through the same mechanism of connecting implied context. For infopathogens, common scams or modes of exploitation are examples of tropes and establish context that is a precursor for exploitation or manipulation.

Host–pathogen interactions in the analogy involve those between functionally individual entities, primarily through social media, and infopathogens. In general, infopathogens affect the behavior of social entities. Cognitive processes elicited following exposure to

a given meme could be studied using mind mapping. Mind mapping essentially joins memes together in a network or organized structure that can help understand their cognitive relationships for a specific individual. This technique has been described as mapping cognition processes analogous to a vehicle and a (road) network (Hills and Kenett 2022). For example, mind mapping has been used to show pathways to suicide and progress from depression towards remediation (Robertson and McFadden 2018). Mind mapping thus supplies a method to aggregate memes at the level of cognitive processes. In terms of the analogy, mind mapping describes the infected host and can be used to describe changes experienced by the host following the incorporation of information, and thus to help infer cause from, or the source of, hypothetically harmful information.

Memes are analogous to genes, together composing a hypothetical infopathogen menome, which is considered to be the memetic equivalent of the genome (Villanueva-Mansilla 2017). Because they are a loose analog of genetic code, memes represent the informatics component of this model: they are units of information whose dynamics can be observed in detail, such as date of first appearance, frequency of replication, sources, and diffusion patterns (Ratkiewicz et al. 2010; Schlaile et al. 2021). In the analogy, memes have observable characteristics, just as genes do, but none are strictly informative with respect to function without also considering hosts. Memes also undergo four stages in replication: assimilation, retention, expression, and transmission (Heylighen 1998). Memes differ in their fitness or their ability to be selected by the human brain and replicate (Spitzberg 2014). Why some memes go viral and others do not may come down to factors such as simplicity, distinctiveness, emotional appeal, connection to celebrities, and network topology, although none of these factors can be used as a blueprint for a successful meme (Schlaile et al. 2018). To understand the causal influence of a meme, the way in which the meme is processed must be considered. Commonly, evidence of meme processing into products manifests as direct or broadcast communication via social media, whereby text or graphics or other memes are used to represent the response to or development of the subject meme. Thus, meme products are analogous with phenotypes, which are observable characteristics of an organism or virus (Martin and Hine 2008), for example pathogen morphology such as spore size and color. Evidence of meme processing on internet social networks that function as "public square" fora may not be readily apparent as networks become increasingly fractured, segregated, and used to conduct highly personalized manipulative attacks—especially when personalization can be enhanced by artificial intelligence, and misinformation can be used to conceal coordination.

### 2.3. Hypothetical Case Study

A hypothetical case study illustrates the potential for using the infopathogen concept and related symptomology to understand an event affected by the presence and transmission of misinformation. The setting is entirely fictional, a metaphor, and not unprecedented: societal hostility toward an institution whose nominal function is the curation and provision of information—a library. This scenario can be generalized to others, such as involving news organizations, agencies, or schools. A public library is chosen as an example because it represents a social institution intended to do public good, but also subject to contentious contemporary debate concerning socially acceptable content. Information about ongoing debate is not primarily disseminated by individual libraries, instead being developed and distributed by news organizations and through social media, such that the example highlights the need for a source-independent understanding of misinformation. The thought experiment follows analytic steps to infer the functions of misinformation related to the event, and what can be learned to anticipate similar incidents. The incident depicted involves misinformation spread initially by manipulative actors and thereafter by affected entities, and is assumed to be deleterious from the perspective of the affected entities in informed hindsight. By depicting a library, the example also compels consideration of the question, what information is appropriate? Because disagreeability and incorrectness are not the same reason to deem information inappropriate, the example

also highlights the need to understand hypothetical misinformation-induced harm versus information-induced change.

Our study assumes that many individuals in the case study are exposed to polarized information on social media as their primary news source. The proportion of adults in the United States who get their information online, including social media, has been consistently around 70% for the last 10 years (Newman et al. 2023). By consuming information through online media, users are empowered to choose news sources, potentially resulting in a restriction of narratives congruent with pre-established (and thus subject to polarization) viewpoints (Nikolov et al. 2015; Sikder et al. 2020). This can result in confirmation bias (Jonas et al. 2001) at large scales (Del Vicario et al. 2016; Sikder et al. 2020), leading to the potential manipulation of individuals to perform harmful or anti-social actions.

## 3. Results

The thought experiment begins with concern about books in a library collection (hereafter known as the "bad books"). Some citizens consider content in the books to be harmful to children. Although this case study is fictional, public concern about children's access to media and information is an enduringly contentious issue, with sexuality, humor, violence, and danger being common concerns (Bickford and Lawson 2020; Doyle 2017). In the case study, a social media campaign is started by concerned citizens to pressure the library to remove books from circulation. Momentum for this social media campaign increases following several memes that become popularized such as "#Banbadbooks".

At some stage, this legitimate exercise of free speech is exploited through the promulgation of disinformation and conspiracy theories by manipulative actors. This results in the disruption of legitimate and factually informed online discourse about the potential impacts of particular books and content on children, which would be expected to resolve with some compromise. This disinformation is amplified by remote actors that seek to sow social discord, a phenomenon that has been previously described (Weintraub and Valdivia 2020). This social media disinformation (which is distinguished from misinformation in that it is deliberately developed and promoted for the purpose of deception) suggests that exposure to certain content is associated with exaggerated and unsubstantiated negative consequences, for example, the disinformation alleges without basis or evidence that the books cause delayed-onset mental health issues. The disinformation is false but not unrealistic. At this stage in the thought experiment, the exploitative goal of the generators and promoters of disinformation is realized when memes that promote harm begin to circulate. The idea connected to disinformation meme products at this point is the degradation of trust in the library itself, and libraries as institutions, and the manipulation of actors results in diffuse meme products such as "#BurnBadBooks" and "#LibraryRiot". The combination of these memes into an infopathogen contributes influence to a complex situation that results in a riot at the library and the burning of the "bad books" by angry citizens (Figure 1). Like a biological infectious agent, the assembly of several parts (here, memes) into the whole is required for symptom manifestation, all depending on the host's potential to be affected.

Following the riot, analysis of the public communications of the individuals involved in the event leads to the identification of content and memes. Memes that are consistently associated with the social media accounts of confirmed perpetrators are identified. These memes constitute a candidate infopathogen menome. Next, memes associated with the event are characterized on an informatic basis such as meme popularity/prevalence, date of origin, history of propagation, and analytics from the associated meme products for each meme such as word cloud descriptions and sentiment scores. Sentiment scores are a natural language processing method to assess the emotional characteristics of memes from both text and images (Pang and Lee 2008; Alluri and Krishna 2021). Common sentiment scores are positive, neutral, and negative, as well as ones based on emotions such as anger, joy, disgust, fear, sadness, and surprise (Prakash and Aloysius 2021; Ratkiewicz et al. 2010). More recently, sentiment analysis has been advanced using ontological models (Dragos

et al. 2018) based on appraisal theory (Martin and White 2003). This model was applied to detect alleged hate speech in France based on the emotional content and subjective judgments of the individual posts (Battistelli et al. 2020). Ideally, an advanced form of sentiment analysis could be used to characterize each meme in the menome, for the purpose of organizing them based on potential effect, and identifying meme combinations that may have an influential effect only when considered together.
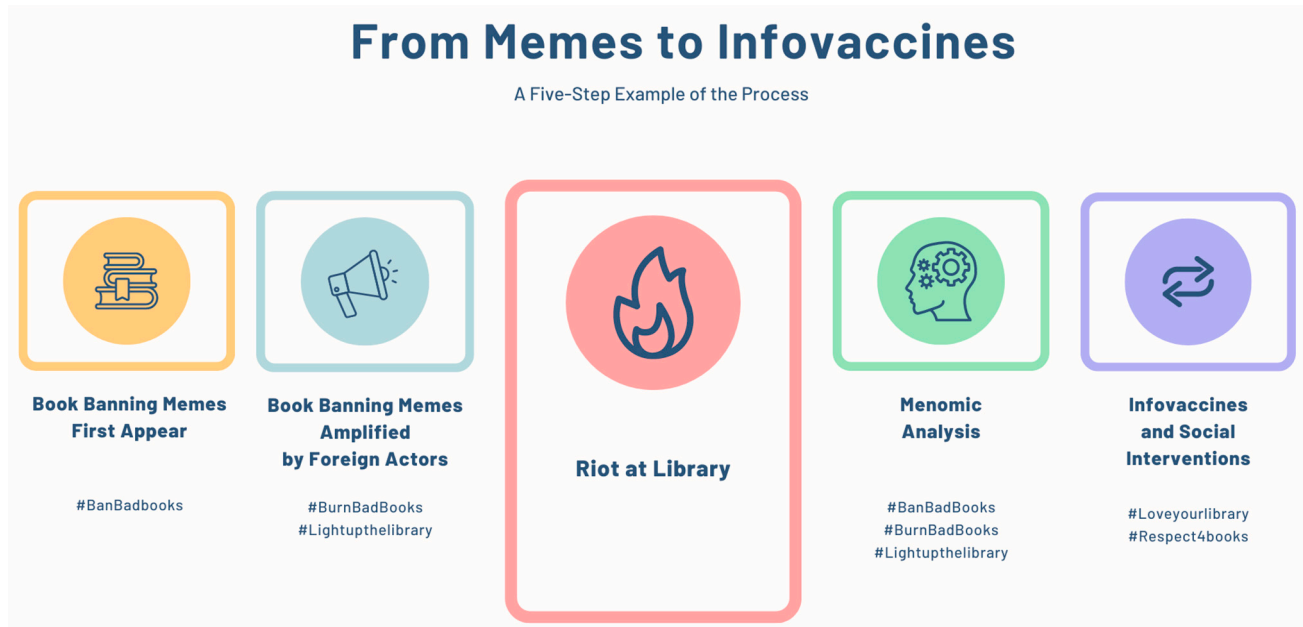


**Figure 1.** Illustration of the case study of a riot at a public library showing bad book memes, amplification by malign actors, the riot, infopathogenic analysis, and potential mitigations.

Because an infopathogen menome could be large, a third step is to analyze the structure and organization of the infopathogen using cognitive mapping (Figure 2). The individual memes associated with the riot are mapped with arrows leading to similar cognitive states which in turn lead to tropes. For example, "#libraryRiot" and "#BurnBadBooks" both translate to the imperative "attack the library", which is in turn mapped to the "Quest" trope. In the case study, the infopathogen is composed of three tropes. The first trope "Right makes might", allows for the aggregation of memes and cognition around the concept that banning bad books is a just cause. The "Right makes might" trope pits adherents of two opposing ideologies against each other, after which they "duke it out", with the outcome indicating whose position is correct (TV Tropes 2022). The second trope "Harmful to minors" aggregates memes and cognition around the concept that bad books harm children. This trope includes things such as violence and sexuality that moral guardians believe kids should not be exposed to until they have sufficient maturity (TV Tropes 2022). The third trope, "the Quest", aggregates cognition and memes around the concept that a physical assault on the library is not only justifiable but also heroic. The Quest trope features the hero and a supporting party traveling on a mission with a firm goal in mind, for example defeating the "big bad" (TV Tropes 2022). Critically, all three of these tropes are required for the menome to be infopathogenic, because responding to any one or two of the memes as though they are valid does not lead to the riot.

The fourth step is to characterize individuals interacting with the memes, in order to understand the sources of misinformation and the patterns of these sources' activities. Individuals interacting with the memes are potential infopathogen-affected individuals, but the degree to which a putative infopathogen will have causal influence on an affected individual is determined by numerous factors especially including the individual's past interaction with memes. For example, weak claims regardless of source can be highly contagious on social media (Pennycook and Rand 2020) but people are more credulous

of misinformation when it comes from a source they consider to be both credible and politically congruent (Traberg and van der Linden 2022). The conduits through which information sources interact with other entities, especially internet platforms such as social media and video sharing sites, are similar to infectious disease vectors such as pathogen-transmitting arthropods (Magarey and Trexler 2020). Internet vectors greatly expand the rate at which infopathogens can spread. Useful data on the source might include the audience size and host demographics for the platform, channel or discussion groups with which the infopathogen is associated. Depending on the size of the corpus and the number of individuals, the host and source ranges could be underestimated due to a limited sample size. This issue is also a problem in biology, for example, the host range of a novel or invasive pest or pathogen may be unknown, although specialized analysis can provide clues (Gilbert et al. 2005). Using data on infectious disease host susceptibility and sources, it is possible to calculate the proportion of infected and at-risk individuals in the population, and this possibility translates to the consideration of infopathogens. The infected proportion is individuals who are probably susceptible and likely to have been exposed to the infopathogen and the at-risk proportion is from individuals who are relatively more likely to act on the infopathogen's associated content as though it is valid, but who have not yet been exposed.
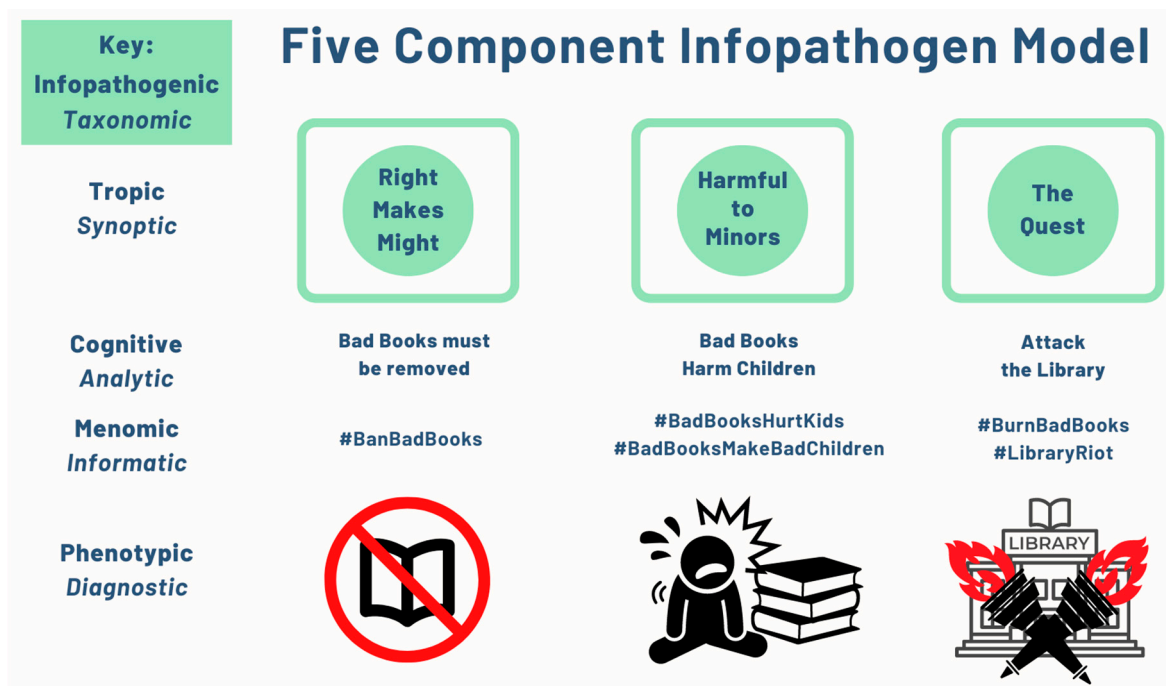


**Figure 2.** Proposed structural composition of the badbook infopathogen consisting of infopathogenic, tropic, cognitive, menomic, and phenotypic components. These components correspond with the taxonomic, synoptic, analytic, informatic, and diagnostic functions of the model system.

Importantly, this analysis does not at any point assume that exposure to a putative infopathogen is deterministic with respect to behavior or negative outcomes. Indeed, confirmation of the existence of an infopathogen, even at a conceptual level, is not required. What is central to the approach is the study of information characteristics, dynamics, and consumer-/context-dependent effects, for the purpose of describing scenarios that are characteristic of misinformation-affected negative outcomes. It is here that symptomology enters the approach, and communication networks that are newly able to be observed and characterized based on trafficking of memes identified as relevant to a given event. The introduction of misinformation by manipulative actors may cause accelerated communication network segregation as a necessary precursor for action that is unpopular at scales other than isolated incidents. In the example developed here, it is assumed that riots

purposed to damaging libraries are broadly condemnable by society, and that the event discussed is the result of manipulative actors' success in causing local, concentrated discord. Manipulative actors engage in something similar to the process through which a virus harmfully repurposes host molecular machinery: information is introduced and acted upon, to the detriment of the host. In this sense it is not the "bad books" that are the ultimate target of the manipulators, but rather the institution(s) with which the manipulators may be competing for resources (e.g., standing or support).

It is possible to develop mitigations for the spread of misinformation that do not require content to be labeled as such. The same cognitive mapping that reveals information-dependent pathways leading to the riot can be used to identify factors that cause individuals to successfully detect the signature of manipulative intent in content or content-sharing patterns. Methods developed to combat suicide likely provide a starting point for the application of mapping to arguably irrational behavior (Robertson and McFadden 2018). At-risk subpopulations have likely not been exposed to suspect content, being immunologically naïve as per the biological analogy, and may benefit from exposure attended by informative context. This is equivalent to psychological immunization against misinformation and is the cognitive analog to biological immunization and is achieved through education and controlled exposure to misinformation (Roozenbeek and van der Linden 2019; Roozenbeek et al. 2022; Van Der Linden 2022) as discussed earlier. In the analogy, we view what is called "psychological immunization" as primarily education—similar to learning about the potential harms of germs, that they are not always strictly avoidable, and that awareness of symptoms can be useful. This education is more like influence and the development of the ability to discern threats, than it is a pre-compensatory manipulation that has functional similarities with exploitative disinformation.

Potential mitigations could be based on the combined characteristics of the infopathogen menome and targeted hosts. Concrete examples of this exist in communication that explains tropes related to scams, such as the common advance-fee scam. In the case of the library riot, individuals' being aware of the relevant misinformation-related tropes confers immunity. The importance of the tropes—functional aspects of misinformation-driven manipulation which are shared between diverse local situations and targeted groups—lies in their being recognizable even in association with a novel event, and independently from their source. Content does not need to be classified as misinformation and suppressed for groups to be effectively immune to related misinformation, if the patterns of the misinformation's influence and the signatures of manipulative intent indicated by the combination of meme types are understood by the groups. In fact, it can be argued that there is a "healthy" amount of misinformation that should circulate in the media so that a baseline level of immunity exists in the population, for two reasons: (1) individual capacity to identify and respond to misinformation is required for first exposure to new content, and (2) an entity able to comprehensively classify and control information to a degree required for the enduring elimination of falsehood would be subject to the same potential errors as the individuals it attempts to protect, with risks of becoming a source of misinformation itself.

## 4. Discussion

The novel contribution of this paper is to outline a strategy to study harmful misinformation using analysis and strategies derived from biology and genetics. It formalizes aspects of the analogy implied by deeming content "viral" and completes the analogy's application to involve aspects that may not be intuitively obvious. The model's components, including infopathogens, tropes, cognition, memes, and phenotypes, can be used for diagnostic, pathologic, and synoptic/taxonomic study of the spread of misinformation. We make the case for a holistic response to misinformation based on epidemiology rather than simply strategies that try to combat the spread of misinformation. Having developed these concepts for a hypothetical case study, the next step is to evaluate the utility of the analogy for a real-world case study with empirical data. Specifically, these studies could focus on case studies where known disinformation or misinformation has spread widely throughout

social media networks. This could provide the needed data to test if the model's components could be adequately described and if the same infopathogen could be identified in similar but distinct case studies.

*4.1. Limitations of the Model*

The framework developed here emphasizes commonalities of informational and biological misinformation, to facilitate translation of useful knowledge concerning biological agents to the study and management of emerging informational agents. Though the use of the analogy elaborated here provides a translation opportunity, it also comes with limitations and caveats. First, the notion that individuals can be infected by information is not completely comparable to biological infection, and thus the key commonality (that both viruses and misinformative memes are information that may exploitatively repurpose host activity) underlying the analogy warrants further empirical exploration and scrutiny. By elaborating the analogy, we also translate difficulties relevant to addressing biological viruses to the domain of misinformation: the conflicting goals of medical vs. public health interventions have principally to do with individual vs. collective agency and priority, and differences between individuals' assessments of risk to themselves and to their respective communities is exquisitely illustrated by the COVID-19 virus, its associated pandemic, and the numerous instances of dis/misinformation surrounding it. While a long evolution of the medical practice in treating disease as metaphysical or self-wrought preceded the modern germ-theory basis for interventions, parallels in the management of misinformation are relatively undeveloped: the contemporary focus is on an analog of sanitation and the elimination of information, versus enhancement of an analog of dynamic immune response. The former is arguably useless in the face of strictly novel misinformation, again underscored by the impacts of COVID-19 ("novel coronavirus"), whereas the latter is useful when the source or exploitative intent of the source of information is unknown.

There are also limitations to applying the biological analogy to the study of disinformation versus misinformation in terms of meme spread. It is well known that memes evolve and propagate as they spread through networks ([Beskow et al. 2020](); [Schlaile et al. 2018]()), and that propagation of a meme depends on its relative fitness compared to other memes ([Spitzberg 2014]()). A critical difference between biological and informational agents is that information agents—especially disinformative ones—are very commonly accompanied by intent, whereas biological agents arise due to human intention exceedingly rarely. This fact strains the analogy between organic agents that evolve depending on survival, and infopathogens that can be created by malefactors despite potential maladaptedness and thus ephemerality of the creations. For example, in the case of elections or climate change, the classification of what is considered misinformation can be a polarized topic, and the classification is itself subject to the influence of intent. A likely challenge for applying the model to health care may be that some of the needed input data may be restricted by privacy laws or the analysis may be inhibited by ethical considerations. Another limitation of the model is that it currently does not have the capabilities to address differences between individuals and their preferences for consumption of information, especially on social media.

In addition, there are also several limitations that impact model performance that require additional research. The first is a need to improve the ability to classify the harmful nature of memes by improved analysis of meme products. Current sentiment analyses are admittedly coarse, as were early genomic analyses that used slow and labor-intensive techniques such as gel electrophoresis to read the four-letter genetic code. Since then, genomic analysis has become far more efficient to the point that whole genomes can be sequenced or decoded quickly and inexpensively ([Heather and Chain 2016]()). Current sentiment analysis techniques allow for a simple description of emotions engendered by memes, but more specific diagnostics are needed. One challenge is to use natural language processing and other techniques to identify infopathogenic memes according to the four criteria defined by [Robertson](https://) ([2017]()). The challenge of this task is that it involves a complex

analysis of emotional states and disruptions to cognition. A recent study of language signatures in social media has demonstrated that the analysis of changes in emotion and cognition is possible, for example proceeding and following a romantic breakup, even in unrelated posts (Seraj et al. 2021). As for the third criteria (appropriation of resources), this could be assessed by analyzing the frequency of infopathogenic memes in users' social media posts. In addition, natural language processing techniques that can identify content associated with anti-social and/or self-destructive behavior are needed, so that the context in which this content appears to cause harm may be described and, ideally, improved such that the content's influence is obviated.

A second performance limitation is the need for research to catalog tropes associated with information-based deleterious contagion, along with associated definitions, expressed symptomology (if any) in affected communication networks, and potential treatments in the form of description of these tropes and related manipulative tactics in their context. Classification of common scams based on their attributes and not necessarily their sources is in line with this research area, which has not only clarified the nature of the scams themselves from the perspective of pathology but has also enabled identification of groups targeted by those attempting the scams, leading to similarly targeted mitigative efforts that have been successful. A related need is to create a taxonomy of infopathogens using data generated from the preceding three research areas. The framework could also include case histories of selected outbreaks and the success and failures of mitigations that were used. Such a taxonomy provides a framework for the development of educational materials that can help individuals recognize infopathogenic patterns and reduce their susceptibility to harm. This taxonomy could be built upon existing frameworks that include typologies and dimensions including information type, motive, facticity, and verifiability (Kapantai et al. 2021).

*4.2. Ethical Considerations*

The proposed infopathogenic analysis suggested in this paper raises several ethical issues. In the 2002 American movie "Minority Report" based on the Philip K. Dick novella "The Minority Report", ethical issues are examined in a fictional setting at the intersection of criminology and predictive forecasting. In the movie, three clairvoyant individuals, or "precogs", predict who will commit homicides briefly before they occur. This prediction allows the police to arrest the implicated individuals before they can commit the forecasted crime. Analysis intended to identify potentially influential misinformation may allow for identification of individuals who have consumed that misinformation; however, none of these individuals may be affected at all. Indeed, the potential for misclassification of content as misinformation, and the potential for intentional misclassification (which is perhaps meta-disinformation) underscores the need for frameworks that do not rely on accurate a priori classification of content before studying or mitigating its influence. Here a basic model is suggested to this effect, focused on the characterization of misinformation as part of a pathogen analog that is highly host-dependent in terms of its effects. Ultimately in the movie, the forecasting accuracy of the precogs is questioned, leading to the abandonment of the program. Concerns about U.S. government control of information on the ostensible basis of misinformation prevention raised concerns that resulted in the announcement of disbanding a short-lived "Disinformation Governance Board" at the federal level, again underscoring the need for means to combat misinformation without relying on a centralized classifier of content without civil oversight. Hierarchically directed efforts to limit the spread of infopathogens could lead to societal damage exceeding those caused by misinformation. Consequently, proposed mitigations for infopathogens should focus on human welfare and be deployed through distributed networks in the same way that mitigations for computer viruses are not centrally or hierarchically controlled.

As was discussed by Magarey and Trexler (2020), the behavior of infopathogens can be described by the Susceptible–Exposed–Infected–Recovered model (SEIR) (Hethcote 2000). The transition from an exposed to an infected individual in this epidemiological model

has an analog with the stages in meme replication: assimilation, retention, expression, and transmission (Heylighen 1998). However, we believe a focus on SEIR analyses for infopathogens could be problematic. Unlike viral genes which enter the host in a capsule, hosts may be exposed to memes from multiple sources, often over a period of time. A second concern is that attempting to analytically define an infection condition of an individual runs immediately into privacy concerns, irrespective of whether the infection status of an individual could even be determined. Instead, we believe focusing on the development of mitigations and socially-based interventions as described above would be more appropriate research topics.

We believe the framing of the analytics of negative outcomes as having a pathogenic causal component can also provide a positive social framework for improving mitigations. Individuals who participate in the library riot are rightfully prosecuted through the criminal justice system, and in the fictitious example developed here, it is assumed that these individuals generally regret participating in the riot. This sets the stage for effective intervention against misinformation in the form of simply studying and explaining the processes through which misinformation causes communication network changes and results in negative outcomes. Pre-conditioning to avoid negative outcomes could focus on specific hypothetical agents or content but this may provide only a narrow range of immunity. Instead, pre-conditioning through shared experience and transfer of knowledge may be a more robust approach because it encourages individuals to innovate in avoiding and managing exposure.

## 5. Conclusions

The potential for the spread of harmful misinformation will continue to increase due to new technologies and information overload (Tandoc and Kim 2022). For example, "deepfakes" are realistic videos created using algorithms that depict people saying and doing things that they did not actually say or do (Fallis 2020). There has been increasing concern that deep fakes will result in an "infopocalypse" where truth cannot be discerned from fiction (Fallis 2020). The responsibilities of social media companies, the roles they play in potentially amplifying misinformation, and their societal impacts have been noted by other authors (Trim 2022; Reisach 2021). Given the profit motives of online media companies to use algorithms that feed users with polarized content irrespective of veracity, it is individual users who must take steps to discern the function and perhaps intent of presented content. Ultimately, research must enhance the ability to detect misinformation (Guo et al. 2019) if it is to accurately classify content as misinformation and inform consumers of its presence. For example, through software tools such as browser extensions (Sharma et al. 2021) that counter exploitative media-selection algorithms without relying exclusively on source classification. The heuristics provided by software antivirus programs, in which potential computer viruses are detected not based on their source or their content but rather on the changes to the host (computer) that they induce, provide a good example of how the analogy between biological infection and misinformation can be elaborated to develop useful interventions that are robust and do not rely on or require control of information sources or consumers. Encouraging developments such as block chain technologies can help limit the spread of misinformation including deepfakes (Fraga-Lamas and Fernández-Caramés 2020). It would however be a mistake for society to limit its response to misinformation purely to combatting its spread in social media and other internet networks. Instead, a holistic approach that includes infopathogenic approaches inspired by epidemiology (Baker et al. 2005) will likely be needed to reduce the harmful impacts of misinformation on society.

**Author Contributions:** Conceptualization, R.D.M., T.C. and K.P.W. Methodology, R.D.M., T.C. and K.P.W.; writing—original draft preparation, R.D.M., T.C. and K.P.W.; writing—review and editing, T.C., R.D.M. and K.P.W.; visualization, K.P.W. and R.D.M. All authors have read and agreed to the published version of the manuscript.

# References

Alluri, Nayan Varma, and Neeli Dheeraj Krishna. 2021. Multimodal Analysis of memes for sentiment extraction. *arXiv* arXiv:2112.11850.

Baker, Edward L., Jr., Margaret A. Potter, Deborah L. Jones, Shawna L. Mercer, Joan P. Cioffi, Lawrence W. Green, Paul K. Halverson, Maureen Y. Lichtveld, and David W. Fleming. 2005. The public health infrastructure and our nation's health. *Annual Review of Public Health* 26: 303–18. [CrossRef]

Baldick, Chris. 2008. *The Oxford Dictionary of Literary Terms*, 3rd ed. Oxford: Oxford University Press. Available online: https://www.oxfordreference.com/view/10.1093/acref/9780199208272.001.0001/acref-9780199208272-e-1172 (accessed on 8 August 2022).

Battistelli, Delphine, Cyril Bruneau, and Valentina Dragos. 2020. Building a formal model for hate detection in French corpora. *Procedia Computer Science* 176: 2358–65. [CrossRef]

Beachy, Roger N. 1997. Mechanisms and applications of pathogen-derived resistance in transgenic plants. *Current Opinion in Biotechnology* 8: 215–20. [CrossRef] [PubMed]

Beskow, David M., Sumeet Kumar, and Kathleen M. Carley. 2020. The evolution of political memes: Detecting and characterizing internet memes with multi-modal deep learning. *Information Processing & Management* 57: 102170. [CrossRef]

Bickford, John Holden, and Devanne R. Lawson. 2020. Examining Patterns within Challenged or Banned Primary Elementary Books. *Journal of Curriculum Studies Research* 2: 16–38. [CrossRef]

Blevins, Jeffrey Layne, Ezra Edgerton, Don P. Jason, and James Jaehoon Lee. 2021. Shouting Into the Wind: Medical Science versus "B.S. in the Twitter Maelstrom of Politics and Misinformation About Hydroxychloroquine. *Social Media + Society* 7: 20563051211024977. [CrossRef]

Brodie, Richard. 2009. *Virus of the Mind: The New Science of the Meme*. Carlsbad: Hay House, Inc.

CDC. 2021. Whole Genome Sequencing (WGS). Centers for Disease Control and Prevention. PulseNet Methods & Protocols. Available online: https://www.cdc.gov/pulsenet/pathogens/wgs.html (accessed on 29 October 2021).

Chen, Mayee F., and Miklós Z. Rácz. 2021. An adversarial model of network disruption: Maximizing disagreement and polarization in social networks. *IEEE Transactions on Network Science and Engineering* 9: 728–39. [CrossRef]

Christakis, Nicholas A., and James H. Fowler. 2013. Social contagion theory: Examining dynamic social networks and human behavior. *Statistics in Medicine* 32: 556–77. [CrossRef]

Del Vicario, Michela, Alessandro Bessi, Fabiana Zollo, Fabio Petroni, Antonio Scala, Guido Caldarelli, H. Eugene Stanley, and Walter Quattrociocchi. 2016. The spreading of misinformation online. *Proceedings of the National Academy of Sciences* 113: 554–59. [CrossRef]

Del-Fresno-Garcia, Miguel. 2019. Information disorders: Overexposed and under informed in the post-truth era. *Profesional de la Informacion* 28: e280302. [CrossRef]

Doyle, Robert P. 2017. *Banned Books: Defending Our Freedom to Read*. Chicago: American Library Association.

Dragos, Valentina, Delphine Battistelli, and Emmanuelle Kelodjoue. 2018. Beyond sentiments and opinions: Exploring social media with appraisal categories. Paper presented at 2018 21st International Conference on Information Fusion (FUSION), Cambridge, UK, July 10–13; pp. 1851–58. [CrossRef]

Evanega, Sarah, Mark Lynas, Jordan Adams, Karinne Smolenyak, and Cision Global Insights. 2020. Coronavirus misinformation: Quantifying sources and themes in the COVID-19 'infodemic'. *JMIR Preprints* 19: 2020.

Fallis, Don. 2020. The Epistemic Threat of Deepfakes. *Philosophy & Technology* 34: 623–43. [CrossRef] [PubMed]

Fernández-Torres, María Jesús, Ana Almansa-Martínez, and Rocío Chamizo-Sánchez. 2021. Infodemic and Fake News in Spain during the COVID-19 Pandemic. *International Journal of Environmental Research and Public Health* 18: 1781. [CrossRef] [PubMed]

Fraga-Lamas, Paula, and Tiago M. Fernández-Caramés. 2020. Fake News, Disinformation, and Deepfakes: Leveraging Distributed Ledger Technologies and Blockchain to Combat Digital Deception and Counterfeit Reality. *IT Professional* 22: 53–59. [CrossRef]

García-Ortega, Rubén Héctor, Pablo García Sánchez, and Juan J. Merelo-Guervós. 2020. Tropes in films: An initial analysis. *arXiv* arXiv:2006.05380.

Gilbert, Marius, Sylvain Guichard, Jona Freise, J. C. Gregoire, Werner Heitland, Nigel Straw, Christine Tilbury, and Sylvie Augustin. 2005. Forecasting Cameraria ohridella invasion dynamics in recently invaded countries: From validation to prediction. *Journal of Applied Ecology* 42: 805–13. [CrossRef]

Goldman, Aaron David, and Laura F. Landweber. 2016. What Is a Genome? *PLoS Genetics* 12: e1006181. [CrossRef]

Gordis, Leon. 2009. *Epidemiology*. Philadelphia: Saunders.

Gorski, David H. 2019. Cancer Quackery and Fake News: Targeting the Most Vulnerable. In *Cancer and Society: A Multidisciplinary Assessment and Strategies for Action*. Edited by Eric H. Bernicker. Cham: Springer International Publishing, pp. 95–112.

Guo, Bin, Yasan Ding, Lina Yao, Yunji Liang, and Zhiwen Yu. 2019. The future of misinformation detection: New perspectives and trends. *arXiv* arXiv:1909.03654.

Heather, James M., and Benjamin Chain. 2016. The sequence of sequencers: The history of sequencing DNA. *Genomics* 107: 1–8. [CrossRef]

Hethcote, Herbert W. 2000. The mathematics of infectious diseases. *SIAM Review* 42: 599–653. [CrossRef]

Heylighen, Francis. 1998. What makes a meme successful? Selection criteria for cultural evolution. Selection Criteria for Cultural Evolution. Paper presented at 16th International Congress on Cybernetics Namur: Association Internattional de Cybernétique, Namur, Belgium, June 26–29; Available online: http://cogprints.org/1132/1/MemeticsNamur.html (accessed on 11 November 2017).

Hills, Thomas T., and Yoed N. Kenett. 2022. Is the Mind a Network? Maps, Vehicles, and Skyhooks in Cognitive Network Science. *Topics in Cognitive Science* 14: 189–208. [CrossRef]

Jachim, Peter, Filipo Sharevski, and Paige Treebridge. 2020. Trollhunter [evader]: Automated detection [evasion] of twitter trolls during the COVID-19 pandemic. *New Security Paradigms Workshop* 2020: 59–75. [CrossRef]

Jin, Fang, Edward Dougherty, Parang Saraf, Yang Cao, and Naren Ramakrishnan. 2013. Epidemiological modeling of news and rumors on twitter. Paper presented at 7th Workshop on Social Network Mining and Analysis, Chicago, IL, USA, August 11–14; pp. 1–9. [CrossRef]

Jonas, Eva, Stefan Schulz-Hardt, Dieter Frey, and Norman Thelen. 2001. Confirmation bias in sequential information search after preliminary decisions: An expansion of dissonance theoretical research on selective exposure to information. *Journal of Personality and Social Psychology* 80: 557–71. [CrossRef]

Kaghazgaran, Parisa, Majid Alfifi, and James Caverlee. 2019. Wide-ranging review manipulation attacks: Model, empirical study, and countermeasures. Paper presented at 28th ACM International Conference on Information and Knowledge Management, Beijing, China, November 3–7; pp. 981–90. [CrossRef]

Kapantai, Eleni, Androniki Christopoulou, Christos Berberidis, and Vassilios Peristeras. 2021. A systematic literature review on disinformation: Toward a unified taxonomical framework. *New Media & Society* 23: 1301–26. [CrossRef]

Kata, Anna. 2010. A postmodern Pandora's box: Anti-vaccination misinformation on the Internet. *Vaccine* 28: 1709–16. [CrossRef]

Kata, Anna. 2012. Anti-vaccine activists, Web 2.0, and the postmodern paradigm–An overview of tactics and tropes used online by the anti-vaccination movement. *Vaccine* 30: 3778–89. [CrossRef]

Kucharski, Adam. 2016. Study epidemiology of fake news. *Nature* 540: 525–25. [CrossRef]

Lerman, Kristina. 2016. Information is not a virus, and other consequences of human cognitive limits. *Future Internet* 8: 21. [CrossRef]

Li, Jianing. 2020. Toward a Research Agenda on Political Misinformation and Corrective Information. *Political Communication* 37: 125–35. [CrossRef]

Lynteris, Christos. 2017. Zoonotic diagrams: Mastering and unsettling human-animal relations. *Journal of the Royal Anthropological Institute* 23: 463–85. [CrossRef]

Magarey, Roger D., and Christina M. Trexler. 2020. Information: A missing component in understanding and mitigating social epidemics. *Humanities and Social Sciences Communications* 7: 128. [CrossRef]

Martin, Elizabeth, and Robert Hine. 2008. *A Dictionary of Biology*. Oxford: Oxford University Press, vol. 6.

Martin, James R., and Peter R. White. 2003. *The Language of Evaluation*. Basingstoke: Palgrave Macmillan, vol. 2.

Newman, Nic, Richard Fletcher, Kirsten Eddy, Craig T. Robertson, and Rasmus Kleis Nielsen. 2023. Reuters Institute Digital News Report 2023. Digital News Report. Available online: https://reutersinstitute.politics.ox.ac.uk/digital-news-report/2023 (accessed on 30 January 2024).

Nikolov, Dimitar, Diego F. M. Oliveira, Alessandro Flammini, and Filippo Menczer. 2015. Measuring online social bubbles. *PeerJ Computer Science* 1: e38. [CrossRef]

Pang, Bo, and Lillian Lee. 2008. Opinion Mining and Sentiment Analysis. *Foundations and Trends® in Information Retrieval* 2: 1–135. [CrossRef]

Pennycook, Gordon, and David G. Rand. 2020. Who falls for fake news? The roles of bullshit receptivity, overclaiming, familiarity, and analytic thinking. *Journal of Personality* 88: 185–200. [CrossRef] [PubMed]

Persily, Nathaniel. 2017. Can Democracy Survive the Internet? *Journal of Democracy* 28: 63–76. [CrossRef]

Prakash, T. Nikil, and Amalanathan Aloysius. 2021. Hybrid Approaches Based Emotion Detection in Memes Sentiment Analysis. *International Journal of Engineering Research and Technology* 14: 151–55.

Ratkiewicz, Jacob, Michael Conover, Mark Meiss, Bruno Gonçalves, Snehal Patil, Alessandro Flammini, and Filippo Menczer. 2010. Detecting and tracking the spread of astroturf memes in microblog streams. *arXiv* arXiv:1011.3768.

Reisach, Ulrike. 2021. The responsibility of social media in times of societal and political manipulation. *European Journal of Operational Research* 291: 906–17. [CrossRef]

Rhodes, Samuel C. 2021. Filter Bubbles, Echo Chambers, and Fake News: How Social Media Conditions Individuals to Be Less Critical of Political Misinformation. *Political Communication* 39: 1–22. [CrossRef]

Robertson, Lloyd Hawkeye. 2017. The infected self: Revisiting the metaphor of the mind virus. *Theory & Psychology* 27: 354–68. [CrossRef]

Robertson, Lloyd Hawkeye, and Robert Cassin McFadden. 2018. Graphing the Self: An application of graph theory to memetic self-mapping in psychotherapy. *RIMCIS: Revista Internacional y Multidisciplinar en Ciencias Sociales* 7: 34–58. [CrossRef]

Rodríguez Fernández, Leticia. 2019. Desinformación y comunicación organizacional: Estudio sobre el impacto de las fake news. *Revista Latina de Comunicación Social* 74: 1714–28. [CrossRef]

Roozenbeek, Jon, and Sander van der Linden. 2019. Fake news game confers psychological resistance against online misinformation. *Palgrave Communications* 5: 12. [CrossRef]

Roozenbeek, Jon, Claudia R. Schneider, Sarah Dryhurst, John Kerr, Alexandra L. J. Freeman, Gabriel Recchia, Anne Marthe Van Der Bles, and Sander Van Der Linden. 2020. Susceptibility to misinformation about COVID-19 around the world. *Royal Society Open Science* 7: 201199. [CrossRef] [PubMed]

Roozenbeek, Jon, Sander Van Der Linden, Beth Goldberg, Steve Rathje, and Stephan Lewandowsky. 2022. Psychological inoculation improves resilience against misinformation on social media. *Science Advances* 8: eabo6254. [CrossRef] [PubMed]

Schlaile, Michael P., Theresa Knausberg, Matthias Mueller, and Johannes Zeman. 2018. Viral ice buckets: A memetic perspective on the ALS Ice Bucket Challenge's diffusion. *Cognitive Systems Research* 52: 947–69. [CrossRef]

Schlaile, Michael P., Theresa Knausberg, Matthias Mueller, and Johannes Zeman. 2021. Viral Ice Buckets: A Memetic Perspective on the ALS Ice Bucket Challenge's Diffusion. In *Memetics and Evolutionary Economics: To Boldly Go Where No Meme Has Gone Before*. Edited by Michael P. Schlaile. Cham: Springer International Publishing, pp. 141–80.

Seraj, Sarah, Kate G. Blackburn, and James W. Pennebaker. 2021. Language left behind on social media exposes the emotional and cognitive costs of a romantic breakup. *Proceedings of the National Academy of Sciences* 118: e2017154118. [CrossRef]

Sharma, Dilip Kumar, Sonal Garg, and Priya Shrivastava. 2021. Evaluation of Tools and Extension for Fake News Detection. Paper presented at 2021 International Conference on Innovative Practices in Technology and Management (ICIPTM), Noida, India, February 17–19.

Sikder, Orowa, Robert E. Smith, Pierpaolo Vivo, and Giacomo Livan. 2020. A minimalistic model of bias, polarization and misinformation in social networks. *Scientific Reports* 10: 5493. [CrossRef]

Slutkin, Gary, Charles Ransford, and Daria Zvetina. 2018a. How the health sector can reduce violence by treating it as a contagion. *AMA Journal of Ethics* 20: 47–55. [CrossRef]

Slutkin, Gary, Charles Ransford, and Daria Zvetina. 2018b. Response to "Metaphorically or Not, Violence Is Not a Contagious Disease". *AMA Journal of Ethics* 20: 516–19. [CrossRef] [PubMed]

Spitzberg, Brian H. 2014. Toward a Model of Meme Diffusion (M3D). *Communication Theory* 24: 311–39. [CrossRef]

Tambuscio, Marcella, Giancarlo Ruffo, Alessandro Flammini, and Filippo Menczer. 2015. Fact-checking effect on viral hoaxes: A model of misinformation spread in social networks. Paper presented at 24th International Conference on World Wide Web, Florence, Italy, May 18–22; pp. 977–82. [CrossRef]

Tandoc, Edson C., and Hye Kyung Kim. 2022. Avoiding real news, believing in fake news? Investigating pathways from information overload to misbelief. *Journalism* 24: 1174–92. [CrossRef] [PubMed]

Thorson, Emily. 2016. Belief echoes: The persistent effects of corrected misinformation. *Political Communication* 33: 460–80. [CrossRef]

Traberg, Cecilie Steenbuch, and Sander van der Linden. 2022. Birds of a feather are persuaded together: Perceived source credibility mediates the effect of political bias on misinformation susceptibility. *Personality and Individual Differences* 185: 111269. [CrossRef]

Trim, Michelle. 2022. Cultivating an ethos of social responsibility in an age of misinformation. *SIGCAS Computers and Society* 50: 13–15. [CrossRef]

TV Tropes. 2022. TV Tropes. Available online: https://tvtropes.org/ (accessed on 19 August 2022).

Van Der Linden, Sander. 2022. Misinformation: Susceptibility, spread, and interventions to immunize the public. *Nature Medicine* 28: 460–67. [CrossRef] [PubMed]

Villanueva-Mansilla, Eduardo. 2017. Memes, menomas e LOLs: Expressão e reiteração a partir de dispositivos retóricos digitais. *MATRIZes* 11: 111–33. [CrossRef]

Vraga, Emily K., and Leticia Bode. 2020. Defining misinformation and understanding its bounded nature: Using expertise and evidence for describing misinformation. *Political Communication* 37: 136–44. [CrossRef]

Wasike, Ben. 2022. Memes, memes, everywhere, nor any meme to trust: Examining the credibility and persuasiveness of COVID-19-related memes. *Journal of Computer-Mediated Communication* 27: zmab024. [CrossRef]

Weintraub, Ellen L., and Carlos A. Valdivia. 2020. Strike and Share: Combatting Foreign Influence Campaigns on Social Media. *OSTLJ Ohio State Technology Law Journal* 16: 701.

Zarocostas, John. 2020. How to fight an infodemic. *The Lancet* 395: 676. [CrossRef] [PubMed]