

## ***Defining critical data literacy for recommender systems: A media-grounded approach***

Arnaud Claes

*Université catholique de Louvain, Belgium*

Thibault Philippette

*Université catholique de Louvain, Belgium*



### **Peer-reviewed article**

**Citation:** Claes, A., & Philippette, T. (2020). Defining critical data literacy for recommender systems: A media-grounded approach. *Journal of Media Literacy Education*, 12(3), 17-29. <https://doi.org/10.23860/JMLE-2020-12-3-3>

**Corresponding Author:**  
Arnaud Claes  
[arnaud.claes@uclouvain.be](mailto:arnaud.claes@uclouvain.be)

**Copyright:** © 2020 Author(s). This is an open access, peer-reviewed article published by Bepress and distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. JMLE is the official journal of [NAMLE](#).

**Received:** February 1, 2020

**Accepted:** October 8, 2020

**Published:** December 14, 2020

**Data Availability Statement:** All relevant data are within the paper and its Supporting Information files.

**Competing Interests:** The Author(s) declare(s) no conflict of interest. The ALG-OPINION research project is funded by the Brussels Institute for research and innovation (Innoviris), ref. 2018-ANTICIPATE-24.

[Editorial Board](#)

### **ABSTRACT**

The digital processing of massive data is becoming a central component of our technological infrastructures. While being able to use these tools efficiently is an issue that cannot be ignored, it appears crucial to provide citizens with the means to control their technical environment. Recommender systems and personalization technologies are currently being blamed for the destabilization of users' informational ecosystems and a growing polarization of opinions. However, a critical review of the current literature on the subject indicates that these recommender systems may also be beneficial to the user in specific circumstances. Building on current critical data literacies approaches, key concepts from the philosophy of technology and a media literacy perspective, this paper suggests a framework defining the competences needed to help users assess these technologies and critically include them in their digital ecosystem.

**Keywords:** *recommender systems, data literacy, technology of the self, media education, media competences.*



**Journal of Media Literacy Education**

THE OFFICIAL PUBLICATION OF THE

NATIONAL ASSOCIATION FOR MEDIA LITERACY EDUCATION (NAMLE)

Online at [www.jmle.org](http://www.jmle.org)

## INTRODUCTION

Technological innovations in computer science are slowly being integrated into the organization of our society and becoming a central infrastructure of our environment. Data is seen as the main fuel of this technological paradigm and taking advantage of its full potential appears to be a critical issue (Boyd & Crawford, 2012). Nevertheless, understanding how we can allow individuals to use these infrastructures without complying with them may be just as necessary. Among those technologies, recommender systems are under heavy scrutiny. These algorithms are integrated into a growing number of online platforms to filter and customize the content they deliver to their users. The behavioral data collected by the system is exploited to determine which are the products or contents that could maximize the user interest. Since political news are also available through such systems, they are held responsible for the fragmentation of the public sphere and the circulation of fake news. Consequently, it seems obvious that there is a need to develop a data literacy in order to curb these phenomena.

The concept of data literacy reflects two related but distinct concerns: (1) how to manage digital information in order to make proper decisions using information and communication technologies (Khan et al., 2018; Qin & D'Ignazio, 2010) and (2) how to properly use those technologies in order to be a functioning individual in our society (Buckingham, 2006). With a few exceptions (e.g. D'Ignazio & Bhargava, 2015), this second matter of interest usually comes with the underlying idea that progress is inevitable, and that individuals must blindly comply with the introduction of news tools by learning how to use them as intended (Fastrez & Philippette, 2017). The advocated solution is most of the time based on the acquisition of technical skills or the development of computational thinking which can more or less be defined as the ability to express a problem in such a way that it can be solved using a computer (Papert, 1993; Wing, 2011). Nevertheless, Heidegger (1977) and later researches in philosophy of technology highlight the fact that technological developments, far from being neutral, remain deeply related to specific social and historical contexts (Feenberg, 2010; Simondon, 1958; Stiegler, 1994). Our environment is becoming more and more mediated by technologies, but this mediation is less and less visible because of, among other things, the seamless interfaces and undisputed infrastructures (Plantin et al., 2018; Star & Ruhleder, 1996). A lack of sufficient understanding of this *technonature* (Roqueplo, 1983)

might alienate the users who are reliant on other technically literate people who end up “engineering” their environment without their consent. However, critical perspectives emanating from the technological field itself allow us to investigate issues related to the transparency or accountability of these technologies (Lepri et al., 2018), revealing the processes and technical choices responsible for their implementation. Those perspectives give us opportunity to envisage ways of supporting a more critical data education that allows individuals to debate and influence the uses of the data by technological platforms (Bucher, 2017).

The first part of this article will be dedicated to the presentation of the concerns raised by recommender systems and how data literacy could help prevent some of them. We will then move on to some elements of theoretical framing of data literacy in order to underline our perspective. Using concepts from the philosophy of technology and the French literature in media education, we will spend some time discussing the media literacy model we rely on to improve existing data literacy models. In the last part of this article, we will briefly present how this framework can also complement design initiatives proposing alternative interaction modalities.

### Recommender systems: defining the issue

The increasing use of personalization technologies on the Internet such as recommender systems led to growing concerns regarding the ability of users to maintain a diversity in their news consumption. The terms *filter bubble* and *echo chamber* were respectively coined by Pariser (2011) and Sunstein (2009) to describe a state of intellectual isolation resulting from filtering algorithms that remove from our view content that would go against our ideas and preferences. Since the two concepts are not properly defined by their respective authors, we will refer to Bruns' definitions:

An echo chamber comes into being when a group of participants choose to preferentially connect with each other, to the exclusion of outsiders. The more fully formed this network is (that is, the more connections are created within the group, and the more connections with outsiders are severed), the more isolated from the introduction of outside views is the group, while the views of its members are able to circulate widely within it. (Bruns, 2017, p. 3)

A filter bubble emerges when a group of participants, independent of the underlying network structures of their connections with others, choose to preferentially communicate with each other, to the exclusion of outsiders. The more consistently they exercise this choice, the more likely it is that participants' own views and information will circulate amongst

group members, rather than any information introduced from the outside. (Bruns, 2017, p. 3)

While an echo chamber is a connectivity issue and a filter bubble a communication one, both can be caused by algorithmic curation or pre-selected personalization, and personal choices or self-selected personalization (Zuiderveen Borgesius et al., 2016). However, both Sunstein's and Pariser's critiques are focused on the algorithmic factor. Based on this deterministic perspective, the solutions are either to "fix" the algorithms, to avoid them, or to sabotage them (Bozdog & van den Hoven, 2015).

Multiple meta-researches indicate that this scenario is unlikely to unfold for common users and that there is enough empirical evidence to reject these theories (Bruns, 2019; Guess et al., 2018; Moeller & Helberger, 2018). Despite algorithmic filtering, several external factors counterbalance the risks of intellectual isolation. Social media are not always primarily used to stay updated on political subjects. By establishing social relations based on other topics, cross-ideological connections are therefore difficult to avoid (Litt & Hargittai, 2016; Messing & Westwood, 2014; Bakshy et al., 2015) and users may benefit from accidental exposure to political news (Fletcher & Nielsen, 2018). Additionally, in comparison to an offline environment, the user is still exposed to a greater diversity of content even with filtering processes occurring (Fletcher & Nielsen, 2017). It also appears that people are not limiting themselves to online sources and are still relying on traditional media (Newman et al., 2020). Finally, many studies on these issues are carried out in the United States, where the political landscape is highly polarized. Thus, it is still unclear whether online polarization phenomena appear under the same conditions in a more heterogeneous media and political landscape (Boczkowski & Mitchelstein, 2013; Garrett, 2013).

According to this quick review, remaining in a filter bubble or an echo chamber requires considerable efforts. As Bruns points out, "this would require extreme homophily, coupled with equally extreme heterophobia, of ourselves and our fellow travelers severing any existing contacts to non-adherents, online and offline, through which outside views could reach us" (Bruns, 2019, p. 34). While companies, technologies and user communities combine to produce unique information environments with their inner logics privileging specific information and values (Madsen, 2016), these environments are not perfectly sealed echo chambers or filter bubbles. According to Bruns (2019), the real issue

is not fragmentation but "polarization", which is not primarily caused by technology but by deeper social and political issues.

However, the claim that personalized news environments do not cause a fragmentation of the audience is based on the premise that people maintain a diversified media consumption regime. The steady decline in news consumption on older media during the past years, although it stabilized recently, contradicts this assertion (Newman et al., 2020). Maintaining a balanced (online) "information diet" therefore remains an important concern for media literacy. Moreover, even though Bruns (2019) sees the problem of polarization as a purely social issue, the studies cited do not allow us to completely reject any form of technological influence. While it appears empirically plausible to refute the deterministic and mono-causal approach favored by Pariser and Sunstein, we cannot ignore the fact that social media and personalization technologies are framing the way we interact with online news. For example, according to an investigation of the New York Times (Horwitz & Seetharaman, 2020), internal research studies in Facebook were conducted to evaluate the effect of their recommender system on the polarization of their community. Apparently, the user engagement metric favored by the system allowed hyperactive partisan users to gain a lot of visibility on their platform. Since those studies were shut down and never made public, we cannot verify the veracity of these claims. However, those statements align with the current literature on the subject and highlight the need for independent studies on this issue to provide a more accurate model than the filter bubble or echo chamber theories.

## DATA LITERACY: A LITERATURE REVIEW

Data literacy is a concept that has emerged particularly from two main strands of research: studies on information literacy and statistical literacy. For authors in the field of information literacy, data literacy is a subset of specific individual skills and knowledge needed to transform data into information (that can be seen as data with a meaning), including the ability to access, manage, interpret, critically evaluate or ethically use data (Carlson et al., 2011; Koltay, 2017; Mandinach & Gummer, 2013). It also includes a series of processes aimed at transforming those data into actionable knowledge (Koltay, 2017), through the ability to make hypotheses from a database or in order to build one, to identify a series of problems or even to monitor a

transformation in the social sphere, what some call “data-driven decisions” (Mandinach & Gummer, 2013). As Koltay points out: “Data curation raises questions related to ownership of the data, its retention, maintenance, access to it, its openness and costs” (Koltay, 2017, p. 5)

Schild (2004) stresses that data literacy is also closely related to statistical literacy described as “the ability to understand and critically evaluate statistical results that permeate our daily lives - coupled with the ability to appreciate the contribution that statistical thinking can make in public and private, professional and personal decisions” (Wallman, 1993, as cited by Prado & Marzal, 2013, p. 125). However, the author considers that statistical literacy goes beyond the ability to process data, but also includes critical thinking: “statistics are more than numbers. Statistics are numerical summaries about things in reality. The nature of things being summarized can make the difference” (Schild, 2004, p. 7). Gould goes even further by considering statistical literacy as a citizenship issue “since democracies require informed debate, and almost all policy discussions require some statistical understanding” (Gould, 2017, p. 22).

While those perspectives are focused on the lifecycle of data management, other approaches consider similar issues by focusing primarily on the underlying infrastructure enabling the flow of digital information. The purpose is to help people better understand variations in representation and how data are altered by computers to allow those representations (Gould, 2017). This computer procedural literacy (Mateas, 2005) does not necessarily mean that individuals must completely “crack the code” of computer tools, but to grasp the links between culturally-embedded practices and abstract technically-mediated processes. For example, as for social networks and the Internet, algorithmic skills can lead to the understanding of the role of the algorithms in making content visible online and how particular recommender systems such as Google Pagerank or Facebook Edgerank work (Cardon, 2015; Klawitter & Hargittai, 2018). However, this perspective on data literacy puts aside some critical issues of citizenship in favor of individual – even complex and reflexive – coding/decoding capabilities through computational systems.

A social perspective emerges when technical systems are questioned regarding their asymmetrical

effects (e.g. Lepri et al., 2018; Brunton & Nissenbaum, 2015, as cited in Pangrazio & Selwyn, 2018), and when the competences of individuals are considered both for personal development and more collectively in the context of political or ethical debates. For example, Pangrazio & Selwyn (2018) mobilize the concept of personal data literacies with a reference to a critical perspective aimed at making individuals aware of the external processing of their personal data by technologies, and the social, political or economic implications of those processes. The educational purpose is to encourage people to develop resistance tactics.

In another article, Pangrazio uses the concept of critical digital literacy in reference to the ability of individuals to use digital networks as resources to shape their own social identities (Pangrazio, 2016). By using the term “creative data literacy,” D’Ignazio argues that “non-technical learners may need pathways towards data which do not come from technical fields” (D’Ignazio, 2017, p. 6). She proposes different tactics to educate people without special technical skills to have a better control over their data (D’Ignazio, 2017). In some ways, the “critical” dimension of this proposition is similar to the critical thinking supported by Schild (2004) regarding statistical literacy.

A critical data literacy can also be understood as the ability to make critical judgements about the ideology underlying data-driven innovation. Aside traditional data literacy which “includes the ability to read, work with, analyze and argue with data”, D’Ignazio (2017) also advocates for a big data literacy that includes skills<sup>1</sup> to identify and address critical problems such as system transparency, personal data extraction, technological complexity or social control (D’Ignazio & Bhargava, 2015). This perspective is overlapping with older media literacy concerns about social issues (e.g. Bazalgette, 1989; Buckingham, 1998; Hobbs, 1998; Livingstone, 2003) while being also close to critical media literacy, which refers to questions of dominant representations and ideological systems (e.g. Alvermann & Hagood, 2000; Kellner & Share, 2007; Ellis & Eberly, 2015).

Despite the undeniable contribution of these different critical perspectives, even when they include a more social or economic dimension (D’Ignazio and Bhargava, 2015; Pangrazio and Selwyn, 2018), the appropriation of data is widely regarded as a personal step towards regaining control (e.g. through critical key

---

<sup>1</sup> In general, we use the term “skills” in reference to functional and operational know-how. However, we prefer the term “competences” when there is a more analytical or critical

dimension. Then, we would use the term “competences” in reference to what D’Ignazio and Bhargava describe, but we keep the authors’ terms here.

questions), omitting the prospective “common good” issues behind the management of digital data. This perspective is, however, largely questioned within the French scientific literature, and we wanted to underline similarities and discrepancies regarding this topic. The French language does not have a direct translation for the concept of literacy and will mostly use the term “culture” as a substitute: for instance, information culture or data culture instead of information and data literacy (Serres, 2009).

In this context, “culture” can refer in the same way as “literacy” to knowledge or skills possessed by an individual regarding a specific domain (Le Deuff, 2009; Serres, 2009). However, this term can also be used with another meaning which is not prevalent in the Anglo-Saxon literature. In this case, “culture” refers to a heritage of norms and values shared by a community (Forquin, 1996). This second meaning invites us to consider the educational process not only as the transmission of competences but also as a collective and reflective practice involving the joint appropriation of a cultural background. For example, in what she called the *hypersphere*, Merzeau (2012) emphasizes the importance of rethinking memory sharing in the digital traces that we must learn to duplicate, recycle and recirculate. It is important for her to overcome the opposition between tracing and protection, since traceability is an integral part of the online sociability of users (Merzeau, 2017). She advocates for a “*mémoire-milieu*” (memory-environment), ascending (i.e., outside the profiling of major economic players) and accompanied (i.e., oriented towards a folksonomy of a long collective memory). In that sense, a data “culture” is the ability to (re)write one’s own traces, playing with their density, tone or reliability, to manage methodically online data, profiles and collective memory (Merzeau, 2017). However, it is obvious that the techno-economic paradigm currently promoted by the major web actors limits the opportunities of carrying out these actions. Merzeau’s theoretical reflection is all the more important in order to reflect on the risk of balkanization triggered by the computational, probabilistic and amnesic processes of the recommender systems designed by big data companies.

### **Limits of current data literacies**

Contemporary studies done in philosophy of technology have stressed the social nature of technological developments. According to Feenberg’s critical theory of technology, social meanings are

attributed to technological artifacts by social groups influencing its design. The notion of efficiency is always related to objectives, representations and values of those specific groups. Thus, choices of design are framing specific ways of life, and technology is imbued with those same representations and values (Feenberg, 2010). As pointed out by Feenberg:

What the object is for the groups that ultimately decide its fate determines what it becomes as it is redesigned and improved over time. If this is true, then we can understand technological development only by studying its meaning for the various groups that influence it (Feenberg, 2010, p.15).

In the case of recommender systems, this social dimension can be easily identified by evaluating the metrics used to define a good recommendation. For example, in 2006, Netflix organized the Netflix Prize, an open competition rewarding the best collaborative-filtering algorithm to predict movie ratings on a 5-star scale. In order to easily evaluate each competitor, the recommendation problem was simplified: the goal was to have a root mean squared error for the predicted rating 10% lower than the algorithm used by the company (Amatriain & Basilico, 2015). In this context, efficiency was defined as the ability to predict the most precisely possible the satisfaction of the user. While this is not an unusual way to frame the recommendation problem, it can have unintended consequences when used in another context. As we discussed previously regarding information on social media, relying on user engagement metrics to define an appropriate recommendation contributes to the growing visibility of highly engaging but also highly polarized contents.

Following a similar perspective, the critical data literacy models described previously are focused on the social and political dimensions of data-driven technologies, with concern about the imbalance between the platform and some users. This asymmetry is attributed to the opacity of the algorithmic processes and users’ poor visibility on data collection procedures, i.e. politics of data (e.g., Boyd & Crawford, 2012), algorithm politics (Bucher, 2018) or politics of platforms (Gillespie, 2010). Based on concepts such as Paulo Freire’s pedagogy of the oppressed (D’Ignazio, 2017) or De Certeau’s strategies and tactics (Pangrazio & Selwyn, 2018), the main purpose of critical data literacy is to identify the social meaning behind specific artifacts in order to highlight the harm done to particular (weaker) users. This is usually achieved by evaluating the metrics involved in the technical process as we did earlier (Beer, 2016). Placing those users in a dynamic of resistance against a dominant actor, suggested

countermeasures are often directed towards sabotage tactics such as obfuscation of personal data. Parisier (2011) suggests for instance to “like” random contents or alternate search engines to confuse the system. While we do agree in a way with the perspective, the suggested behaviors appear disproportionate regarding the issue we are dealing with.

Our previous overview of the filter bubble theory highlighted the excessive weight given to algorithmic factors and their impact on online disinformation. It even appears that such systems can bring informational benefits to users in specific contexts (Bozdog & van den Hoven, 2015). Therefore, while resistance is highly relevant for issues such as data privacy, adopting tactics of sabotage against recommender systems could be counterproductive.

An “instrumental” viewpoint that focuses solely on the development of functional skills leads to a representation of the user as someone who must keep up-to-date with technological innovation and does not allow the discussion of non-deterministic issues (Buckingham, 2006).

As we expose with our review of filter bubble and echo chamber theories on the previous pages, the main issue regarding recommender systems is not the biases and defects of this technology in itself, but how those devices are defining the activity they organize, in our case news consumption. As Feenberg points out, in critical theory of technology, “technologies are not seen as mere tools but as frameworks for ways of life” (Feenberg, 2006, p. 14). These frameworks are constrained by what he calls a “cultural horizon”, which is a set of “culturally general assumptions that form the unquestioned background to every aspect of life” (Feenberg, 2010, p. 16). According to his idea, rationalization, defined as the use of scientific knowledge to acquire a greater control of our environment including our social life, is our modern horizon (Feenberg, 2010).

Our digital ecosystem is constrained by this horizon. In *The Invisible Computer*, Norman (1998) argues for the development of human-centered technologies that are transparent to the user in order to refocus his attention on the work to be done. The technical mediation must be as invisible as possible to ease the experience of the user. The technological background behind a great part of our social interactions is slowly fading out of view unless we consciously focus on it to allow for a greater efficiency (Plantin et al., 2018; Sandvig, 2013; Van Den Eede, 2011).

This situation has resulted in a vision of news consumption as a passive process where the user input is mostly limited to social contacts and the data gathered by the platform (Spohr, 2017). As a side effect, many users are not aware of the presence of filtering tools on social networks (Eslami et al., 2015; Hamilton et al., 2014).

When Feenberg develops his theory, it is with the intent to politicize and democratize technological design. What is interesting in this perspective is the role of dilemmas when the cultural horizon enters a phase of negotiation and transition. Previously unrepresented interests get the chance to defend their case. Once the case is settled, technological updates are integrated into the technical code reflecting the new cultural horizon involved in the design process. Personalization technologies such as recommender systems are currently experiencing this state of transition. While being inaccurate, the theories of filter bubble and echo chamber have contributed to the popularization of this issue and to bringing it into the public debate. In addition, a growing number of news’ organizations are entering the field with their own resources, goals and values. (Bodó, 2018). All of this is contributing to the diversification of the technical landscape and the development of new perspectives.

Although the design of recommender system and their interfaces is not the core of our contribution, we believe that media and data literacy have a role to play in this mechanism by encouraging users to challenge current horizons and to seek alternative tools. Users have a degree of maneuverability that does not always correspond to the initial intent of the designers. In the long run, these unexpected uses end up being formalized in the technical code and integrated into the social meaning of the device (Feenberg, 2010).

As Spohr (2017) emphasizes, users need to be aware that high-quality information is not something that reach the users passively. On the contrary, users should actively engage with their media environment to shape it in a way that suits their needs instead of relying on unconscious mechanisms. Regarding recommender systems, we believe that they can be used in a way that contradicts the passivity they promote by default. Ultimately, encouraging this mental shift could contribute to the evolution of design choices. Using concepts borrowed from media literacy, the next section will be dedicated to the definition of competences that could promote an active engagement with personalized information ecosystems.

**Towards a media-grounded data literacy (for online news)**

Using Feenberg’s critical theory of technology, we previously emphasized how technological artifacts reflect the intents of their designers. Based on these premises, we argue that recommender systems can also be studied as media therefore allowing us to borrow concepts and practices from media literacy. According to Anderson and Meyer, a medium is “a recognizable human activity that organizes reality into readable texts for engagement” (Anderson & Meyer, 1988, p. 316). From this perspective, recommender systems, their companies and communities can be analyzed as unique newsrooms. By organizing information according to specific values, interest and constraints, both designers of recommender systems and journalists of a newsroom are composing a frame through which we see the world around us. None of these frames is better than the other but they can complete each other under the condition that their structure is made visible.

Unlike many traditional media literacy models which are dimensional, associating either dimensions

(media format, audience, genre, etc.) or operations (use, understand, analyze, evaluate, etc.) with a set of pre-formulated key questions, the framework we mobilize is matrix-based. Conceptually, this model allows an evolving problematization in relation to emerging (mediation) issues —such as the “problem” of filter bubble on informative platforms, by articulating informational dimension with more social and societal dimensions, while analyzing the mediating role played by digital technologies. This model is the one that has been adopted as a reference framework by the Higher Council for Media Education in French-speaking Belgium.

In this model, media literacy is defined as a set of competences “required to perform different tasks (reading, writing, navigating, and organizing) on a variety of media considered as informational, technical and social objects” (Tilleul et al., 2015, p. 76). Through this matrix of generic competences (Fastrez, 2010, p. 48), media activities can be analyzed as intersections of four types of media tasks and three perspectives about the media.

<b>Media as... Tasks:</b>	Informational objects	Technical objects	Social objects
<b>Reading</b>	The referential, semiotic and formal aspects of media contents	The media technical processes and apparatus	The social effects, the ethical issues, the production and reception contexts, the original intentions, etc.
<b>Writing</b>			
<b>Navigating</b>			
<b>Organizing</b>			

Figure 1. *Matrix definition of media literacy: the dimensions (see Tilleul et al., 2015, p. 41)*

A medium can be analyzed in terms of its informational, technical and social dimensions (see Figure 1). As media messages and devices, they are “informational objects, designed to represent things, real or fictitious, different from themselves, through the use of different sign systems” (Tilleul et al., 2015, p.77). For example, a visualization or a graphic has often a referent, and visual choices (colors, shapes, etc.) are a way of saying something about or interpreting that referent.

The media are also the result of technical processes and apparatus and/or are designed to create other media objects (Tilleul et al., 2015). Last but not least, media are social objects, and as such, they establish or translate

power relationships, they are concerned with ethical issues, they depend on institutional contexts of production, they have various effects regarding cultural contexts of reception, etc. These dimensions are interrelated.

If we refer to media tasks (see Figure 2), they can be about a single media object (e.g., reading a book or writing a text message) or about a set of media (e.g., navigating a website and organizing a social network). Each task can be crossed with each dimension resulting in twelve competences defining media literacy. This matrix model is further explained in an article published in French by Fastrez (2010).

<b>Media as... Tasks:</b>	Informational objects	Technical objects	Social objects
<b>Reading</b>	Decode, understand and evaluate media		
<b>Writing</b>	Create and disseminate media productions		
<b>Navigating</b>	Search (activity with a pre-established objective) and/or explore (open activity)		
<b>Organizing</b>	Categorize using <i>ad hoc</i> typologies and/or implementing tools for organizing documents/technologies/social relations		

Figure 2. Matrix definition of media literacy: the tasks (see Tilleul et al., 2015, p. 41)

By looking at this matrix model, we can identify what is already covered and what is still missing from our data literacy literature review. Data literacy and some other information-based literacies discussed earlier are applied to reading and writing competences at a technical level. Those competences are focused on the ability of the user to technically retrieve or produce information; and critical perspectives add a social dimension into the mix. Thus, a social reading of data-driven technologies involves decoding the politics of data we discussed previously; and a social and technical writing relates to the tactics of resistance and sabotage where the user consciously alters her behavior to disturb the system. Returning to our main study object, we can note that these critical approaches have essentially considered algorithms in recommender systems as technical objects (i.e., filtering and automating a complex set of processes) or as social objects (i.e., profiling their users for the benefit of media companies), but that the informational dimension (i.e., the type of data, the design of the search engine, the formatting of the results) is considered relatively rarely. Yet this informational dimension is essential to the social meaning of the search requests and results activities. The coding or decoding capabilities of individuals cannot be dissociated from their will and intentions. It is precisely by giving them the means to understand the gap between their intention (social) and their behavior based on affordances (informational) and constraints (technical) that their agency can be improved. In other words, this model serves to diagnose problems of use and help to avoid the risk of trying to educate from the perspective of so-called good practices based on social issues that do not correspond or correspond imperfectly to the cultural horizon of particular people.

By being centered on reading and writing tasks, these literacies are focused on linear and causal relations with a single medium. How much of my newsfeed is

altered by algorithm A? How is it influencing me? At which point is company C directly responsible? How can I configure A to optimize the results? On the subject of recommender systems, we tried to emphasize how most of these questions can only be answered in a local context depending on the environment in which the user is situated (i.e., platforms used, socio-economic context, information exposure habits, media landscape). Moreover, several major approaches in philosophy of technology including Feenberg's critical theory defend an ecological understanding of technological artifacts (Van Den Eede, 2019). They invite us to disentangle ourselves from a narrow perspective on technology to consider the relations influencing the specific situations in which technological determination occurs.

A filter bubble or echo chamber effect can happen on a single platform under specific conditions. However, this phenomenon can also be turned into an opportunity if it is counterbalanced by adequate habits and tools weaving a richer network of sources and filtering processes. Whether those systems have positive or negative effects depends on their context of use and how they are associated with other media (O'Hara & Stevens, 2015). Developing organization and navigation competences as presented in this matrix can therefore be insightful and help users develop a more nuanced and refined control of their digital ecosystem.

Organization competences include the ability to categorize multiple media according to their technical components, personal benefits to the user, communities, values or goals. In a complementary fashion, navigation competences involve being able to search and explore media according to well-defined criteria, as well as making them searchable within a community (which induces an articulation between organization and navigation). These competences can adopt multiple shapes depending on the matrix dimension. It can be about recognizing and classifying how different

platforms leverage multiple kind of information to tailor their recommendations and how each of these modalities can complement each other. It is also balancing news sources provided by personalized contents and more traditional mass media in order to get rid of monolithic technical or editorial choices. In addition to comparing sources, which is a rather classical aspect of media and information literacy, it is a question of being able to enrich, modify and recirculate data to improve the relevance of information for oneself and for the community (or communities) to which one belongs. Finally, it also means developing strategies or tools aimed at reorganizing these data.

### **Developing agency with regard for recommender systems**

Organization and navigation competences are not only applied to manage the informational space delimited by a media ecosystem. They can also contribute to assessing the influence of automated recommendations over time and help the user fine-tune the system according to his goals and expectations. By studying last.fm and their recommender system, Karakayali et al. (2018) observed that users are not necessarily oriented towards definite music tastes but are constantly asked to engage with their library and make a conscious effort to redefine their tastes. Using Foucault's work on the care of the self, the authors defend the idea that recommender systems must not only be apprehended as technologies of control but also as means of self-care and self-cultivation (Karakayali et al., 2018). By filtering content online, they offer to the user a space of possibilities framed by specific rules with which to experiment. Nonetheless, users must be willing and competent enough to actively explore this space through richer interfaces. Thus, if a traditional critical media or data literacy aims at suppressing the dangers of oppression and control, our complementary (critical) approach aims at analyzing and developing the potential of those technologies of the self (Karakayali et al., 2018; Reigeluth, 2017).

This perspective requires merging media literacy concerns with questions of design. On this topic, the project Gobo from the MIT Media Lab offers an interesting way to reflect on information filtering and recommender systems (Bhargava et al., 2019). Gobo allows the user to connect multiple newsfeeds from his or her social media accounts, and to display them

simultaneously on the application. Immediately displayed in an ante-chronological manner, the feed can be reorganized afterwards according to different filters that the user can set up. By allowing the end-user to contribute to the filtering process, Gobo serves two missions. Firstly, it makes the user more aware of the information that is pushed towards her/him by the traditional platforms. Secondly, it serves as an open experimental place to evaluate the capacity of users to intervene more actively on the proposed contents according to personal needs rather than externally calculated ones. By offering a technology to think with (see Gobo website<sup>2</sup>), similar interfaces could contribute to the development of a critical media-grounded data literacy by establishing a more transparent relation with the user. This transparency does not mean a thorough understanding of the technical components but a capacity to see oneself in the algorithm as well as his influence on others (Reigeluth, 2017). If properly integrated into media education workshops, similar projects could help stimulate learners' willingness to engage critically with richer interfaces to consciously manage their media ecosystem. Our future studies will therefore be dedicated to the exploration of users' agency by developing a similar platform relying on the principles presented in this paper.

### **CONCLUSION**

By personalizing digital ecosystems and filtering the information, recommender systems transform the way users access information and how users can be informed citizens. These technological changes invite us to rethink the way a critical media education can contribute to tackling this issue and how to define a literacy in a way best suited for these challenges.

Since social media and online news platforms are mainly relying on personal data to operate, we decided to focus our attention on data literacy models that have already tried to deal with similar concerns. Considering that most of these studies have been performed in the field of information and statistical literacy, these models appeared incomplete for us to deal with polarization and fragmentation as a media issue. While a critical data perspective was conceptually adequate, the recommended practices involving sabotage tactics could be counterproductive with regard to the underlying causes of the problem.

---

<sup>2</sup> <https://gobo.social/>

Using concepts coming from philosophy of technology, we rephrased the issue surrounding recommender systems. It is not about a defect in the machine to correct either by education or by design. It is more about understanding the limitations of these technologies to navigate more easily in a wider ecosystem in which we can find balance. Even if it is obvious that technical knowledge as well as critical perspectives are needed to “read” our digital environment, we looked for a model to work on and evaluate the development of a literacy integrating navigation and organization dimensions reflecting a critical and active user behavior towards the configuration of its information ecosystem. Learning how to navigate and organize content across multiple platforms using different algorithms is the first step of a shift in horizon where data is more actively managed and recommendation tools are better integrated in search activities.

This shift in perspective should be supported by media education workshops intertwined with new design principles. This paper is mostly theoretical, but we consider it as a basis for future empirical approaches. Further research is still needed to assess the potential of specific interaction modalities and visualization techniques to encourage active exploration and self-cultivation. Open-source projects such as Gobo offer interesting opportunities for such design initiatives supporting studies aimed at improving what we call a *media-grounded data literacy*. We hope that the ideas developed in this paper will also contribute to applied studies and educational activities regarding access to information online.

## REFERENCES

- Alvermann, D. E., & Hagood, M. C. (2000). Critical media literacy: Research, theory, and practice in “New Times.” *The Journal of Educational Research*, 93(3), 193-205.
- Amatriain, X., & Basilico, J. (2015). Recommender systems in industry: A Netflix case study. In F. Ricci, L. Rokach, & B. Shapira (Eds.), *Recommender Systems Handbook* (2nd ed.) (pp. 385-420). Springer US. <https://doi.org/10.1007/978-1-4899-7637-6>
- Anderson, J. A., & Meyer, T. P. (1988). *Mediated communication: A social action perspective*. Sage Publications.
- Bakshy, E., Messing, S., & Adamic, L. A. (2015). Exposure to ideologically diverse news and opinion on Facebook. *Science*, 348(6239), 1130-1132. <https://doi.org/10.1126/science.aaa1160>
- Bazalgette, C. (1989). *Primary media education: A curriculum statement*. BFI Education.
- Beer, D. (2016). *Metric power*. Palgrave Macmillan.
- Bhargava, R., Chung, A., Gaikwad, N. S., Hope, A., Jen, D., Rubinovitz, J., Saldías-Fuentes, B., & Zuckerman, E. (2019). Gobo: A system for exploring user control of invisible algorithms in social media. *Conference Companion Publication of the 2019 on Computer Supported Cooperative Work and Social Computing*, 151-155. <https://doi.org/10.1145/3311957.3359452>
- Boczkowski, P. J., & Mitchelstein, E. (2013). *The news gap: When the information preferences of the media and the public diverge*. MIT Press.
- Bodó, B. (2018). *Means, not an end (of the world): The customization of news personalization by European news media* (SSRN Scholarly Paper ID 3141810). Social Science Research Network. <https://doi.org/10.2139/ssrn.3141810>
- Boyd, D., & Crawford, K. (2012). Critical questions for big data. *Information, Communication & Society*, 15(5), 662-679.
- Bozdog, E., & van den Hoven, J. (2015). Breaking the filter bubble: Democracy and design. *Ethics and Information Technology*, 17(4), 249-265. <https://doi.org/10.1007/s10676-015-9380-y>
- Bruns, A. (2017, September 15). *Echo chamber? What echo chamber? Reviewing the evidence* [Conference]. 6th Biennial Future of Journalism Conference (FOJ17).
- Bruns, A. (2019). *Are filter bubbles real?* Polity Press.
- Bucher, T. (2017). The algorithmic imaginary: Exploring the ordinary affects of Facebook algorithms. *Information, Communication & Society*, 20(1), 30-44. <https://doi.org/10.1080/1369118X.2016.1154086>
- Bucher, T. (2018). *If...Then: Algorithmic Power and Politics*. Oxford University Press.
- Buckingham, D. (1998). Media education in the UK: Moving beyond protectionism. *Journal of Communication*, 48(1), 33-43. <https://doi.org/10.1111/j.1460-2466.1998.tb02735.x>
- Buckingham, D. (2006). Defining digital literacy—What do young people need to know about digital media? *Nordic Journal of Digital Literacy*, 1(4), 263-276.
- Cardon, D. (2015). *A quoi rêvent nos algorithmes: nos vies à l'heure des big data*. Le Seuil.

- Carlson, J., Fosmire, M., Miller, C. C., & Sapp Nelson, M. (2011). Determining data information literacy needs: A study of students and research faculty. *Libraries and the Academy*, 11(2), 629-657. <https://doi.org/10.1353/pla.2011.0022>
- D'Ignazio, C. (2017). Creative data literacy: Bridging the gap between the data-haves and data-have nots. *Information Design Journal (IDJ)*, 23(1), 6-18. <https://doi.org/10.1075/idj.23.1.03dig>
- D'Ignazio, C., & Bhargava, R. (2015, September 28). *Approaches to Building Big Data Literacy*. Bloomberg data for good exchange conference, New York.
- Ellis, A., & Eberly, T. (2015). Critical Literacy: Going Beyond the Demands of Common Core. *Illinois Reading Council Journal*, 43(2).
- Eslami, M., Rickman, A., Vaccaro, K., Aleyasen, A., Vuong, A., Karahalios, K., Hamilton, K., & Sandvig, C. (2015). "I always assumed that I wasn't really that close to [her]": Reasoning about invisible algorithms in news feeds. *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, 153-162. <https://doi.org/10.1145/2702123.2702556>
- Fastrez, P. (2010). Quelles compétences le concept de littératie médiatique englobe-t-il ? Une proposition de définition matricielle. *Recherche en Communication*, (33), 35-52.
- Fastrez, P., & Philippette, T. (2017). Un modèle pour repenser l'éducation critique aux médias à l'ère du numérique. *tic&société*, 11(1), 85-110. <https://doi.org/10.4000/ticetsociete.2266>
- Fastrez, P., Tilleul, C., & De Smedt, T. (2016). Media literacy evaluation: A critical discussion of reliability and usefulness issues. *Media Education Research Journal*, 7(1), 36-55.
- Feenberg, A. (2006). What is philosophy of technology? In J. Dakers, *Defining technological literacy: Towards an epistemological framework*. (pp. 5-16) Palgrave Macmillan US. <https://doi.org/10.1057/9781403983053>
- Feenberg, A. (2010). *Between reason and experience: Essays in technology and modernity*. MIT Press.
- Fletcher, R., & Nielsen, R. K. (2017). Are news audiences increasingly fragmented? A cross-national comparative analysis of cross-platform news audience fragmentation and duplication. *Journal of Communication*, 67(4), 476-498. <https://doi.org/10.1111/jcom.12315>
- Fletcher, R., & Nielsen, R. K. (2018). Are people incidentally exposed to news on social media? A comparative analysis. *New Media & Society*, 20(7), 2450-2468. <https://doi.org/10.1177/1461444817724170>
- Forquin, J.-C. (1996). *École et culture: Le point des sociologues britanniques* (2e éd). De Boeck Université.
- Garrett, R. K. (2013). Selective Exposure: New Methods and New Directions. *Communication Methods and Measures*, 7(3-4), 247-256. <https://doi.org/10.1080/19312458.2013.835796>
- Gillespie, T. (2010). The politics of 'platforms.' *New Media & Society*, 12(3), 347-364. <https://doi.org/10.1177/1461444809342738>
- Gould, R. (2017). Data literacy is statistical literacy. *Statistics Education Research Journal*, 16(1), 22-25.
- Guess, A., Nyhan, B., Lyons, B., & Reifler, J. (2018). *Why selective exposure to like-minded political news is less prevalent than you think* (p. 25). Knight Foundation.
- Hamilton, K., Karahalios, K., Sandvig, C., & Eslami, M. (2014). A path to understanding the effects of algorithm awareness. *CHI '14 Extended Abstracts on Human Factors in Computing Systems*, 631-642. <https://doi.org/10.1145/2559206.2578883>
- Heidegger, M. (1977). *The Question Concerning Technology, and Other Essays*. Harper.
- Hobbs, R. (1998). The seven great debates in the media literacy movement. *Journal of Communication*, 48(1), 16-32. <https://doi.org/10.1111/j.1460-2466.1998.tb02734.x>
- Horwitz, J., & Seetharaman, D. (2020, May 26). Facebook executives shut down efforts to make the site less divisive. *The Wall Street Journal*. <https://www.wsj.com/articles/facebook-knows-it-encourages-division-top-executives-nixed-solutions-11590507499>
- Karakayali, N., Kostem, B., & Galip, I. (2018). Recommendation systems as technologies of the self: Algorithmic control and the formation of music taste. *Theory, Culture & Society*, 35(2), 3-24. <https://doi.org/10.1177/0263276417722391>
- Kellner, D., & Share, J. (2005). Toward critical media literacy: Core concepts, debates, organizations, and policy. *Discourse: Studies in the Cultural Politics of Education*, 26(3), 369-386. <https://doi.org/10.1080/01596300500200169>
- Khan, H. R., Kim, J., & Chang, H.-C. (2018). *Toward an understanding of data literacy*. iConference 2018. <https://www.ideals.illinois.edu/handle/2142/100243>

- Klawitter, E., & Hargittai, E. (2018). "It's like learning a whole other language": The role of algorithmic skills in the curation of creative goods. *International Journal of Communication*, 12, 3490-3510.
- Koltay, T. (2017). Data literacy for researchers and data librarians. *Journal of Librarianship and Information Science*, 49(1), 3-14.  
<https://doi.org/10.1177/0961000615616450>
- Le Deuff, O. (2009). Penser la conception citoyenne de la culture de l'information. *Les Cahiers du numérique*, Vol. 5(3), 39-49.
- Lepri, B., Oliver, N., Letouzé, E., Pentland, A., & Vinck, P. (2018). Fair, transparent, and accountable algorithmic decision-making processes. *Philosophy & Technology*, 31(4), 611-627.  
<https://doi.org/10.1007/s13347-017-0279-x>
- Livingstone, S. (2003). *The changing nature and uses of media literacy* [Monograph].  
<http://www.lse.ac.uk/collections/media@lse>
- Madsen, A. K. (2016). Beyond the bubble: Three empirical reasons for re-conceptualizing online visibility. *MedieKultur: Journal of Media and Communication Research*, 31(59), 22.  
<https://doi.org/10.7146/mediekultur.v31i59.19235>
- Mandinach, E. B., & Gummer, E. S. (2013). A systemic view of implementing data literacy in educator preparation. *Educational Researcher*, 42(1), 30-37.  
<https://doi.org/10.3102/0013189X12459803>
- Mateas, M. (2005). Procedural literacy: Educating the new media practitioner. *On the Horizon*, 13(2), 101-111. <https://doi.org/10.1108/10748120510608133>
- Merzeau, L. (2012). Faire mémoire de nos traces numériques. *E-dossier de l'audiovisuel (INA)*.  
<https://halshs.archives-ouvertes.fr/halshs-00727308>
- Merzeau, L. (2017). Mémoire partagée. In M. Cornu, F. Orsi & J. Rochfeld (Ed.), *Dictionnaire des biens communs*. Presses Universitaires de France.
- Messing, S., & Westwood, S. J. (2014). Selective exposure in the age of social media: Endorsements trump partisan source affiliation when selecting news online. *Communication Research*, 41(8), 1042-1063. <https://doi.org/10.1177/0093650212466406>
- Moeller, J., & Helberger, N. (2018). *Beyond the filter bubble: Concepts, myths, evidence and issues for future debates*.  
<https://dare.uva.nl/search?identifier=478edb9e-8296-4a84-9631-c7360d593610>
- Newman, N., Fletcher, R., Schulz, A., Simge, A., & Nielsen, R. K. (2020). *Reuters Institute Digital News Report 2020* (p. 112).
- Norman, D. A. (1998). *The invisible computer: Why good products can fail, the personal computer is so complex, and information appliances are the solution*. The MIT Press.
- O'Hara, K., & Stevens, D. (2015). Echo chambers and online radicalism: Assessing the internet's complicity in violent extremism. *Policy & Internet*, 7(4), 401-422. <https://doi.org/10.1002/poi3.88>
- Pangrazio, L. (2016). Reconceptualising critical digital literacy. *Discourse: Studies in the Cultural Politics of Education*, 37(2), 163-174.  
<https://doi.org/10.1080/01596306.2014.942836>
- Pangrazio, L., & Selwyn, N. (2018). 'Personal data literacies': A critical literacies approach to enhancing understandings of personal digital data. *New Media & Society*.  
<https://doi.org/10.1177/1461444818799523>
- Papert, S. (1993). *The children's machine: Rethinking school in the age of the computer*. BasicBooks.
- Pariser, E. (2012). *The filter bubble: What the Internet is hiding from you*. Penguin Books.
- Plantin, J.-C., Lagoze, C., Edwards, P. N., & Sandvig, C. (2018). Infrastructure studies meet platform studies in the age of Google and Facebook. *New Media & Society*, 20(1), 293-310.  
<https://doi.org/10.1177/1461444816661553>
- Prado, J. C., & Marzal, M. Á. (2013). Incorporating data literacy into information literacy programs: Core competencies and contents. *Libri*, 63(2), 123-134.  
<https://doi.org/10.1515/libri-2013-0010>
- Qin, J., & D'Ignazio, J. (2010). *Lessons learned from a two-year experience in science data literacy education*. International Association of Scientific and Technological University Libraries, 31st Annual Conference.
- Reigeluth, T. (2017). Recommender systems as techniques of the self? *Le Foucauldien*, 3(1), 7.  
<https://doi.org/10.16995/lefou.29>
- Roqueplo, P. (1983). *Penser la technique: Pour une démocratie concrète*. Seuil.
- Sandvig, C. (2013). The internet as infrastructure. *The oxford handbook of internet studies*.  
<https://doi.org/10.1093/oxfordhb/9780199589074.013.0005>
- Schild, M. (2004). Information literacy, statistical literacy, data literacy. *IASSIST Quarterly Summer*, 282(3). <https://doi.org/10.29173/iq790>
- Serres, A. (2009). Penser la culture informationnelle: Des difficultés de l'exercice. *Les Cahiers du numérique*, 5(3), 9-23.

- Simondon, G. (2012). *Du mode d'existence des objets techniques*. Aubier.
- Spohr, D. (2017). Fake news and ideological polarization: Filter bubbles and selective exposure on social media. *Business Information Review*. <https://doi.org/10.1177/0266382117722446>
- Star, S. L., & Ruhleder, K. (1996). Steps toward an ecology of infrastructure: Design and access for large information spaces. *Information Systems Research*, 7(1), 25.
- Stiegler, B. (1994). *La technique et le temps*. Galilée/Cité des sciences et de l'industrie.
- Sunstein, C. R. (2009). *Republic.com 2.0*. Princeton University Press.
- Tilleul, C., Fastrez, P., & De Smedt, T. (2015). Evaluating media literacy and media education competences of future media educators. In S. Kotilainen & R. Kupiainen (Eds.), *Reflections on Media Education Futures. Contributions to the conference Media Education Futures in Tampere, Finland 2014* (pp. 75–89). Nordicom.
- Van Den Eede, Y. (2011). In between us: On the transparency and opacity of technological mediation. *Foundations of Science*, 16(2), 139-159. <https://doi.org/10.1007/s10699-010-9190-y>
- Van Den Eede, Y. (2019). *The beauty of detours: A Batesonian philosophy of technology*. State University of New York.
- Wing, J. M. (2011). Computational thinking. *2011 IEEE Symposium on Visual Languages and Human-Centric Computing (VL/HCC)*, 3-13. <https://doi.org/10.1109/VLHCC.2011.6070404>
- Zuiderveen Borgesius, F., Trilling, D., Möller, J., Bodó, B., Vreese, C. H. de, & Helberger, N. (2016). Should we worry about filter bubbles? *Internet Policy Review*, 5(1). <https://doi.org/10.14763/2016.1.401>