

## **Appreciating news algorithms: Examining audiences' perceptions to different news selection mechanisms**

**ABSTRACT:** As news selection is increasingly controlled by algorithms, a growing number of scholars are exploring how news recommenders can serve public services. Despite aspirations towards public service algorithms, little is known about which type of news recommender people prefer, let alone about a news recommender that aims to promote societal values. This study aims to give insights into audiences' perceptions to news recommenders and their underlying news selection mechanisms. To do so, we distinguish between three news selection mechanisms, namely between content-based similarity, collaborative similarity and content-based diversity. The first two strive for similarity, respectively between news content and news users, while the third one aims for diversity in the news content consumed. Results of a large-scale survey (n=943) show that people prefer content-based similarity over collaborative similarity and content-based diversity. Audience characteristics, such as news information overload and concerns towards missing challenging viewpoints, explain how audiences evaluate the different news selection mechanisms. We discuss how these results align with concerns about selectivity and how news algorithms can be used to tackle these concerns. We therefore introduce the concept 'personalized diversity' and promote the idea of news recommenders as an individual filter for the growing abundance of online information.

**KEYWORDS:** Algorithmic appreciation; algorithmic perceptions; collaborative similarity; content-based diversity; content-based similarity; news recommenders; news selection; public service algorithms

**WORD COUNT:** 9277 (excl. appendices)

### **Introduction**

Prompted by concerns on selective exposure and disinformation, scholars are increasingly examining how news recommenders work - how they automatically select, sort and rank news items - and what their effects are on diversity dimensions such as topic and sentiment diversity (e.g., Haim et al., 2018; Möller et al., 2018). The rationale underlying these studies is the idea that existing news recommenders are predominantly interest-based, and hence lead people towards 'more of the same' news. Subsequently, such recommenders might fail to provide diversity in the news content people are exposed to (Zuiderveen Borgesius et al., 2016).

Although empirical research currently holds a more nuanced view on this discussion (e.g., Bodó, 2019; Haim et al., 2018; Möller et al., 2018), there is a growing idea among scholars to embed news diversity as a design principle into news recommenders (Helberger, 2019; Helberger et al., 2018). Here, news articles are not recommended on how much clicks or likes they have, but rather on what might make an individual read in a more diverse way. This idea also corresponds with the concept of public service algorithms, in which news organizations make societally relevant use of algorithmic power and preserve societal values (Jones & Jones, 2019).

Despite the potential of this idea, little research has been done on how audiences perceive different news recommenders and their underlying news selection mechanisms. Existing research on algorithmic perceptions mainly focuses on the perceived differences between human and algorithmic recommendations (e.g., Araujo et al., 2020; Thurman et al., 2019), even though different types of news selection mechanisms exist, and largely vary on how they are designed and what they are designed to do. Consequently, it remains unclear if - and why - audiences are willing to adopt a news recommender that promotes news diversity.

This study therefore first aims to give insights into the audiences' perceptions of different types of news recommenders and their underlying news selection mechanisms. To do so, we distinguish between three news selection mechanisms: content-based similarity, collaborative similarity and content-based diversity. The first two are the most dominant mechanisms used by news organizations. They primarily strive for similarity, by respectively looking at the user's previous news use and the news use of similar user profiles (Karimi et al., 2018). In contrast, content-based diversity aims to increase the diversity of the news content. As such, it offers news articles of which the user has read too little. Using this distinction, we aim to give a more fine-grained overview of how algorithmic news selection is perceived by news audiences.

Second, by using content-based diversity as a design mechanism, this study also aims to map the potential of public service algorithms in a news media environment. In recent years, several public service media in Europe (e.g., BBC, NPO, VRT) have explored the use of algorithmic news recommender (Fields et al., 2018; Takken, 2018), yet none of them have used them as a primary tool to provide audiences with news (Van den Bulck & Moe, 2018). This study may give public service media valuable audience knowledge to support them to embrace content-based diversity in their news offer, as well as give them suggestions on how their news recommenders should be designed.

## Literature review

We start our literature review by discussing previous studies on algorithmic aversion and algorithmic appreciation, two terms increasingly used in research on automated decision-making and algorithmic perceptions. Next, we discuss which types of news recommenders exist and what the idea of a content-based diversity in news recommendation entails. To understand why users prefer certain news selection principles more than another, we end our literature review with a comprehensive discussion on the factors that have been found important in literature.

### *Algorithmic aversion and appreciation*

In recent years, research on automated decision making increasingly extended their focus towards the use and perceptions of algorithms (Burton et al., 2020). As such, scholars do not only focus on the societal risks and consequences of algorithms but also examine how and when audiences accept or reject recommendations. In general, studies on the use and perceptions of algorithms show two important, yet somehow contradicting results. On the one hand, there is a dominant body of empirical research on the concept of *algorithmic aversion* (e.g., Burton et al., 2020; Castelo et al., 2019; Dietvorst et al., 2018), supporting the idea that human decision-makers are less likely to choose an algorithm over a human recommender. On the other hand, there is a recent and growing body of empirical studies on the concept of *algorithmic appreciation* (e.g., Araujo et al., 2020; Logg, 2017; Logg et al., 2019), promoting the idea that audiences prefer algorithms over human recommendations.

On both sides, several theories have been raised to explain why people averse or prefer algorithms compared to humans. In research on algorithmic appreciation, for instance, two theories are argued to be important: *theory of machine* and *theory of machine heuristics* (Araujo et al., 2020; Logg, 2017; Logg et al., 2019). The theory of machine deals with how people think how algorithms work. It argues that people have certain ideas about algorithms, including the input, process and output of an algorithm, even when they do not know how they work (i.e., 'folk theories'). These ideas, in turn, might influence how people respond to algorithms (e.g., Bucher, 2017; Eslami et al., 2016). The theory of machine heuristics, then, suggests that algorithmic interfaces seem to be less attributed to human traits (Sundar, 2008). Subsequently, people consider their decisions to be objective and free of ideological biases, which in turn has a positive impact on how these decisions are evaluated. Conversely, interfaces with a more anthropomorphic look will have a lesser degree of objectivity (Araujo et al., 2020).

Both terms – algorithmic appreciation and aversion - are currently exclusively used to compare the perceptual differences between human recommendations and algorithmic recommendations. However, as Möller et al. (2018) argue, algorithms cannot be seen as a single type or entity. In a news media environment, different types of news algorithms exist, each using a different mechanism to select, sort and rank news articles. A case in point is the study of Thurman et al. (2019) where two types of news selection mechanisms are distinguished: news selection based on individuals' peers' consumption (i.e., collaborative news selection) and news selection based on individuals' consumption behavior (i.e., content-based news selection). They found that that this distinction leads to different evaluations according to the source of the data, with user tracking preferred over peer filtering. Building on this work, we broaden the notion of 'algorithmic appreciation' and use it to compare how audiences evaluate different types of news recommenders. As such, we distinguish *between* different news selection mechanisms, rather than between humans and algorithms. This distinction allows us to give a more fine-grained overview of how news recommenders and their news selection mechanisms are perceived.

### *News recommenders and their selection mechanisms*

To understand the different types of news recommenders, it is important to note that news recommenders have several design attributes that determine how news articles are selected, sorted and ranked. First, as previously discussed, news recommenders can make use of various data sources to make recommendations. These data sources may range from the previous consumption behavior of individuals (i.e., content-based news selection), to previous consumption behavior of others such as friends or people with a similar reading behavior (i.e., collaborative news selection). Most commonly, researchers use a combination of both approaches (i.e., hybrid news selection, Karimi et al., 2018). Second, news recommender can make use of a several (combinations of) measures to assess and re-rank the outcome. These measures include, but are not limited to, the "freshness" of an article (i.e., recency), the difference between present and past experiences (i.e., novelty) or the similarity of content (i.e., accuracy). Traditionally, accuracy measures are used to predict the most relevant news item for individuals (Karimi et al., 2018).

Despite the popularity of news recommenders that strive for similarity, several concerns have been raised on the societal implications of accuracy measures. In particular, when it comes to news diversity, news recommenders that use accuracy measures are perceived as a societal risk as they might hinder well-informed citizenry (Zuiderveen Borgesius et al., 2016). Ideally, in a democratic society, people need to come across opinions that differ from their own opinions, so they can develop themselves fully (i.e., 'marketplace of ideas') (Napoli, 2006). As news recommenders are argued to primarily select news based on similarity (Joris et al., 2019), people might be increasingly exposed to the topics or viewpoints that are relevant for them. Subsequently, people would encounter fewer opinions and subjects, which may hinder the democratic opinion-forming process in society (Viķe-Freiberga et al., 2013; Zuiderveen Borgesius et al., 2016).

To tackle this challenge, scholars suggest to use *exposure diversity* as an evaluation measure and design choice in the development of news recommenders (Helberger, 2019). On a conceptual level, exposure diversity refers to the content that an individual actually selects, as opposed to all the content that is available (McQuail, 1992). It is thus concerned with the extent to which the consumed news content is diverse in terms of various content dimensions such as demographics, topics, ideas, viewpoints, etc. From a normative point of view, exposure diversity can be understood as a necessary condition for human progress and a well-functioning democracy (Helberger; 2011). It is not an end in itself, but a mean to engage users to inform themselves, to read different viewpoints and to broaden their perspectives (Helberger, 2012).

More specifically, diversity-based news recommenders focus on what might make a person read in a more diverse way. This differs with current practices in which accuracy measures are used or in which diversity is used to increase the internal differences within the set of recommended items (see e.g.,

Castells et al., 2015). Instead, diversity-based news recommenders aim to achieve diversity *within* the reading behavior of individuals and provide news articles of which the user has read too little. As such, their primary role is to compensate the biases in people's news selection choices. As previous research has shown that people generally expose themselves to news articles that are in line with their main interests, views and behaviors (i.e., selective exposure) (D'Alessio & Allen, 2002; Hart et al., 2019), we argue that diversity-based news recommenders primarily deliver news articles that do not align with someone's interests or opinions.

Collaborative elements such as a friend's news consumption behavior could also be used as a data source in diversity-based news recommenders. However, in this study, we focus on the most frequently used recommendation techniques. As such, we distinguish between three news selection mechanisms in this study: content-based similarity, content-based diversity and collaborative similarity. Considering this typology, we may formulate the following research question (**RQ1**): how do audiences evaluate the following news selection mechanisms: content-based similarity, content-based diversity and collaborative similarity?

Few studies have focused on the relationship between algorithmic news perception and news selection mechanisms such as diversity. In journalism studies, one Dutch study showed that audiences' evaluations regarding news personalization depend on the expected output of news recommendation (Bodó et al., 2019). In particular, if people expect that news recommenders will deliver them news in a more diverse way, they value personalization more, the study shows. These results also align with studies in the computer sciences (e.g., Ekstrand et al., 2014), which suggest that an increase in the diversity of recommendations can lead to a better quality perception.

#### *User characteristics*

On an individual level, several characteristics may explain how people differently evaluate the discussed news selection mechanisms. These characteristics may range from news use and news attitudes, such as news interest and mobile use (Fletcher & Nielsen, 2019; Thurman et al., 2019), to socio-demographic characteristics, including age, gender and level of education (e.g., Bodó et al., 2019; Fletcher & Nielsen, 2019; Thurman et al., 2019).

Although these studies show that a broad range of characteristics is relevant, these scholars primarily focus on traditional news use characteristics. However, as news recommenders are technologies that provide people with a selective range of news articles, technological attitudes such as technology optimism or cognitional attitudes such as news information overload might also explain how people evaluate news recommenders and their underlying news selection mechanisms. Therefore, we broaden the user characteristics to explain the audiences' evaluations from news use with more technological and cognitional characteristics. In the following paragraphs, we will elaborate on these characteristics. This also leads us to the second research question (**RQ2**): which user characteristics explain differences in audiences' evaluations of different news selection mechanisms?

A first characteristic to consider is the concept of technology optimism, which refers to the belief that technology offers people increased control, flexibility and efficiency (Parasuraman & Colby, 2015). As Sundar and Marathe (2010) state, people with such a positive view on technology are likely to have strong self-efficacy and clear outcome expectations, both good predictors of technology use. Applying this to a news recommender context, in which news selection is automated and adapted to how people want to receive news articles, we expect that people who are optimistic about technology also appreciate news recommenders more, independent from the underlying news selection mechanism.

A second characteristic is news information overload, referring to the psychological stress and negative emotions induced by the large amount and speed of news production nowadays (Lee et al., 2017), such as being stressed or being overwhelmed. To cope with these feelings, people may search for news selection strategies that filter the news offer, which allows them to save cognitive resources or operate

more efficiently (Pentina & Tarafdar, 2014; Schmitt et al., 2018). These strategies may include the habitual use of certain types of news sources (Yuan, 2011), or the use of news recommenders to filter out relevant news messages (Aljukhadar et al., 2012; Beam & Kosicki, 2014; Dylko et al., 2017). As such, we expect that people who experience news information overload more may evaluate any type of news recommender more positively.

Moreover, we expect that this effect will be stronger for content-based similarity and collaborative similarity. As these news selection types allow people to be exposed to news similar to their interests, people may use them as a way to receive attitude-consistent messages or circumvent counter-attitudinal messages, similar to how people selectively use media sources and content (Knobloch-Westerwick, 2014). As such, news recommenders could be used to converge choice, values and beliefs, to dissolve conflicting perceptions in their media use (Festinger, 1957), and hence to reduce dissonance before an information encounter.

A third characteristic is the concept of online news participation, or the act of commenting and sharing news articles online. As Fletcher and Park (2017) explain, these actions may have several motivations, from a desire to voice opinions on matters of public concern (Nagar, 2011), to balancing discussions, and expressing emotions or opinions (Stroud et al., 2016). What most of these motivations have in common is a willingness to participate in civic activities and to feel connected to their local communities. Subsequently, we may argue that people who participate online may be more favorable to the idea of a shared public sphere where people discuss news articles that are of the most public concern. In the context of news recommendations, this idea closely relates to how collaborative filtering works. Based on the popularity of a news article and the user's social network, this type of news recommender will select news articles that align with the interests of a community. Therefore, we expect that people who participate with news articles online have higher a higher preference for collaborative mechanisms, compared to other types of news recommenders.

## Methods

As the literature study has shown, literature on algorithmic perceptions generally neglect the existence of different types of news recommenders as well as the importance of technological and cognitional attitudes in explaining audiences' evaluations. To understand how audiences evaluate different news recommenders, a large-scale survey was conducted based on the different news selection mechanisms that underlie these news recommenders. In this section, we present the methods we used to conduct this survey.

### *Sampling and procedure*

In November 2019, a large-scale survey was launched in Ghent among 4,995 citizens. Ghent is a densely populated city in Flanders, the northern and Dutch-speaking part of Belgium. It is the second-largest city in Flanders and covers a population of approximately 250,000 residents. Flanders has a relatively small and concentrated media market that is characterized with a democratic corporatist system (Hallin & Mancini, 2004). There are four large news media companies that supply a significant part of the audience: DPG Media, Mediahuis, Roularta and VRT (Picone & Donders, 2020). Audiences in Flanders generally have high levels of trust in traditional news sources (Newman et al., 2020). 61% say they use social media at least once a month to follow the news (Vandendriessche & De Marez, 2019).

The inclusion criteria to participate in this study were: (1) being older than 18 and (2) not living in an institutional setting (e.g., a home for the elderly or prison) and (3) not been recently contacted by other researchers from Ghent University. All information required was derived from the municipal registry administered by Ghent, except for the last criterion. Based on these criteria, we drew an implicit stratified sample ( $n = 4,995$ ) from the municipal registry. This sample was stratified by four variables: age (i.e., 18–24, 25–34, 35–44, 45–54, 55–64, 64–74, 75+), sex (i.e., male, female), current

nationality (i.e., domestic nationality, foreign nationality) and neighborhood (i.e., 25 neighborhoods). Appendix 1 shows an overview of the sample. The sampling procedure and data protocol were approved by the data protection officers of both Ghent University and Ghent as well as the city council. Information on this approval can be found in the protocol attached (see Appendix 2).

To invite citizens to participate, two invitation letters were sent out in two consecutive weeks (see Appendix 3). To increase response-rate, we followed the guidelines described by Dillman, Smyth, and Christian (2014) (e.g., personalized salutations, signature). Both letters were sent on Tuesday and were assumed to be received on Thursday or Friday, close to the weekend. The self-administrated questionnaire could only be filled in online, by making use of a unique code attached to the letter. People who were not able to go online were recommended to go to the local library, where they can make use of free Wi-Fi and PC's connected to the web. To stimulate people to participate, we provided 50 €10-vouchers as incentives.

To derive a representative response sample and to understand non-response, we drew a random sample of 200 persons who did not participate and were significantly underrepresented in our response sample. This was especially needed for persons with a foreign nationality. To stimulate them to participate, we conducted 152 home visits where we invited them personally to participate. In Appendix 4 and 5, the results of these home visits are presented as well as a final overview of the respondents' characteristics. The final response rate of this survey was 20.74% (n = 1036) before data cleaning and 18.88% after data cleaning (n = 943).

### *Measures*

The survey used a 5-point Likert scale, where 1 is equal to strongly disagree and 5 is equal to strongly agree (except where noted). Appendix 6 summarizes all variables, including questions texts and response samples per item.

### **Dependent variables**

For our dependent variables, the evaluation of the news selection mechanisms, we developed three scales based on the three selection mechanisms we previously discussed in our literature review: content-based similarity, collaborative similarity and content-based diversity. On each of these scales, respondents had to indicate to what degree they think that they are good ways to select news articles for him/her.

*Content-based similarity* - We used items that are concerned with previous consumption behavior (e.g., 'having stories automatically selected for me based on what I have consumed') and respondents' interests (e.g., 'having stories automatically selected for me based on my interests').

*Collaborative similarity* - We used items that focus on the user's social environment (e.g., having stories automatically selected for me based on what my friends have consumed is a good way to get news) and the community's interests (e.g., 'having stories automatically selected for me based on what is currently popular).

*Content-based diversity* - We used items that relate with the outcomes of a news recommender that is being used by people who primarily expose themselves to news articles that are in line with their main interests, views and behaviors (D'Alessio & Allen, 2002; Hart et al., 2019). Subsequently, as a diversity-based news recommender aims to achieve diversity *within* the individual's reading behavior, the news recommender will primarily deliver counter-attitudinal stories in terms of viewpoints and topics. Taking this into account, we used items that focus on diversity in counter-attitudinal viewpoints (e.g., 'having stories selected for me based on the viewpoints I do not agree with') and topics (e.g., 'having stories selected for me based on the topics I do not usually read').

### **Independent variables**

For our independent variables, the user characteristics, we conducted a review of existing scales. Taking into account the length of the scales, the respondents' attention span and the scales' empirical validity, we made several operational choices to limit the survey length.

*Trust* - We used the scale of Yale et al. (2015), which focuses on three dimensions of trust: balance, honesty and currency (e.g., to what degree do you think that the news of a news organization is balanced).

*News information overload* - We extended the scale of Lee et al. (2017) and focused on the respondents' cognitive and emotional state when being confronted with the amount and speed of news production nowadays (e.g., 'I often felt overwhelmed about the large amount of daily news').

*Technology optimism* - We used the scale developed by Parasuraman and Colby (2015) that maps the benefits that are related to technology use in general (e.g., new technologies contribute to a better quality of life).

*Concerns towards news personalization* - We built on the study by Thurman et al. (2019) who distinguished three concerns: concerns related to privacy, concerns related to missing important information and concerns related to missing challenging viewpoints. Respondents were asked to what degree they think that news personalization affects each of these aspects (e.g., I worry that more personalized news may mean that I miss out on challenging viewpoints).

*News interests* - We used an index developed by Van Damme et al. (2019) which sums up several news themes such as politics, sports or celebrity news. Respondents were asked to give a score between 0 (= no interest) and 10 (= very interested) to each of these news themes.

*News use* - We asked respondents how often they use different kinds of news sources (i.e., television, radio, news website). Respondents could choose between four frequencies: never, monthly, weekly and daily. To have a more detailed look at their news consumption, we also asked respondents how much time (in minutes) per day they have regularly spent on reading news.

*Foreign news use* - We used a binary category (1 = used a foreign news source last year), as was for paying behavior (1 = having paid for online news last year).

*Socio-demographic variables* - We also collected multiple socio-demographic variables: age, gender, education, employment status and nationality. For age and gender, we asked the respondents' year of birth and sex. For education, employment status and nationality, we used a list of categories that aligns with the current living situation in Flanders. From these categories, we subsequently derived the level of education, current working status and nationality in Flanders.

### *Data analysis*

In the first phase of data analysis, we performed a descriptive analysis on all variables. In this analysis, we retrieved all univariate statistical parameters and inspected whether there were items with little variation or skewed means (see Appendix 6). In this phase, we also checked the construct validity of all latent constructs by conducting a confirmatory factor analysis (CFA) (DeVellis, 2016; Farrell, 2010) (see Appendix 8). Second, we performed a latent class analysis (LCA) to classify news users with similar news interest, preferences or use into distinct groups of people. Subsequently, we conducted a linear and hierarchical regression analysis. A hierarchical analysis allowed us to specify the order in which predictors are entered in the regression model (Bryman, 2015). In this study, they were entered based on their importance in predicting the outcome (low to high). As such, we first entered socio-demographic variables, then the news use variables, and finally the variables related to news and technology attitudes. In addition, we also conducted a residual analysis to check whether assumptions such as linearity, homoscedasticity, and normality were met (Kutner et al., 2005). All assumptions were confirmed.

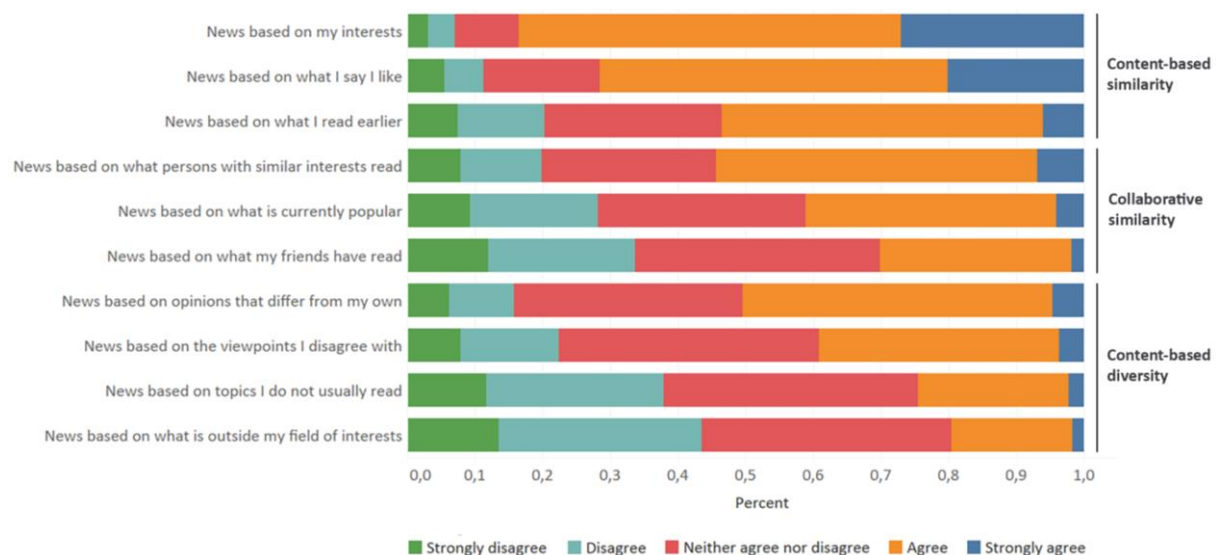
## Results

### *Descriptive analysis: Audiences' evaluation of different news selection mechanisms*

To understand how audiences evaluate the different news selection mechanisms (RQ1), we conducted a descriptive analysis of all survey items (see Figure 1). Figure 1 shows that people generally have a higher preference for news selection mechanisms underlying content-based similarity (items 1-3) than for mechanisms expressing collaborative similarity (items 4-6) or content-based diversity (items 7-10). This preference was particularly high for news selection based on the users' interest (item 1) or the users' explicit preferences (item 2); respectively 83.46% and 71.48% of the respondent (strongly) agree that these two news selection mechanisms are good ways to select news articles. Here, we also note the lowest percentages of respondents strongly disagreeing, disagreeing or neither agreeing nor disagreeing and the highest mean scores (respectively 4.00 and 3.75) (see Appendix 6). Analysis of the mean scores (see appendix 7) confirm that the mean scores of these two items are significantly different from all other mean scores.

Looking at the differences between the items of collaborative similarity and content-based diversity, we see a slightly higher preference for items covering collaborative similarity, such as news selection based on people with similar interests or news selection based on popularity. However, this preference does not hold for all items related to content-based diversity. In particular, items 7 and 8 are on par evaluated than for items used to measure collaborative similarity. This is also reflected in the mean scores that are not significantly different from each other (see Appendix 7). This result might be explained by the fact that the first two items (items 7 and 8) of content-based diversity focus on diversity in opinions and viewpoints and the last two items (items 9 and 10) relate to diversity in topics and interests. Our results show that there is a significant difference in evaluation between these news selection mechanisms.

When we look at the other side of the bar charts, we see the percentages of those who (strongly) disagree with the different news selection mechanisms. Most remarkable here are the percentages for item 7 which represents 'news based on opinions that differ from my own'. In contrast to other items covering content-based diversity, a relatively low percentage (15.80%) is observed.



**Figure 1.** Stacked bar chart to measure the audiences' evaluation of different news selection mechanisms: content-based similarity (items 1-3), collaborative similarity (items 4-6) and content-based similarity (items 7-10).



*Cluster analysis: profiling audiences by their evaluation of different news selection mechanisms*

To have a deeper understanding of how audiences evaluate different news selection mechanisms, we also conducted a latent class cluster analysis. This analysis enabled us to identify distinct groups of people who similarly evaluate the different news selection mechanisms (items 1-10). Based on the different LC-models and their model fit information (see Appendix 9), we chose for the 4-cluster model which has a p-value < .05 and the best fit in relation to the number of clusters to use (i.e., elbow-method).

In this 4-cluster model, we identified four groups with each similar evaluations of the news selection mechanisms. The first group is the “*optimistic news selector*” (32.47%), referring to the group’s broad acceptance of various news selection mechanisms. People in this group like to receive news based on their interests, but are also interested to encounter opinions that differ from their own. News based on what their friends have read is not their most favorite way to select news articles, but they do not reject this news selection mechanism. The second group is the “*choosy news selector*” (31.02%) who is particularly interested in their own news interests. They are not in favor to receive news topics they do not usually read. If they want some diversity in their news diet, they prefer to receive different opinions that differ from their own. The third group is the “*broad-minded news selector*” (27.91%) who is also open for any news selection mechanism, but, in contrast to the first and the second cluster, they have a very high preference for content-based diversity. In particular, they highly value viewpoint diversity. The fourth group is the “*averse news selector*” (8.60%) who is generally not in favor for any news selection mechanism. If one mechanism should be picked, they choose to receive news based on their interests.

**Table 1.** Mean scores of the 4-cluster model: audiences’ evaluation of different news selection mechanisms

	Cluster 1: optimistic news selector (32.47%)	Cluster 2: choosy news selector (31.02%)	Cluster3: broad-minded news selector (27.91%)	Cluster4: averse news selector (8.60%)
News based on my interests	3.67	4.45	4.33	2.57
News based on what I say I like	3.28	4.32	4.14	2.21
News based on what I read earlier	3.05	3.67	3.70	1.81
News based on what persons with similar interests read	2.91	3.79	3.84	1.65
News based on what is currently popular	2.87	3.34	3.44	1.72
News based on what my friends have read	2.63	3.04	3.37	1.46
News based on opinions that differ from my own	3.40	3.16	3.96	1.66
News based on the viewpoints I disagree with	3.14	2.92	3.81	1.60
News based on topics I do not usually read	3.06	2.21	3.47	1.46
News based on what is outside my field of interests	3.01	2.02	3.28	1.40

*Linear and hierarchical regression analysis: Explaining audiences’ evaluation by user characteristics*

This study also raised the question which user characteristics explain differences in audiences’ evaluations of different news selection mechanisms (RQ2). To answer this question, we first conducted a latent class cluster analysis to identify a typology of news audiences, based on their shared news use and news interests (see Appendix 10-13 for model fit information and mean scores). These typologies are subsequently used in our linear regression analysis to discover how they and other user characteristics explain the audiences’ evaluation towards different news selection mechanisms. The results of these analyses are presented in Table 2, which shows the regressions’ standardized coefficients and their associated model statistics (e.g., adjusted R<sup>2</sup>).

Starting with the first regressions on *content-based similarity*, the results show eleven user characteristics that have been found significant in the final model (i.e., regression (3)). The most prominent predictors of the audiences' evaluation towards content-based similarity are age ( $\beta = -0.246, p < .05$ ), nationality ( $\beta = 0.206, p < .05$ ), news interests ( $\beta = -0.136, p < .05$ ;  $\beta = -0.116, p < .05$ ;  $\beta = -0.082, p < .05$ ) and concerns regarding challenging viewpoints ( $\beta = -0.134, p < .05$ ). As such, the results show that those with a domestic nationality and people with a broad range of news interests (cluster 1) score significantly higher in their evaluation of content-based similarity, compared to foreigners and the Low interested (cluster 4). However, we also see an interaction effect of nationality and the use of foreign news sources ( $\beta = 0.103, p < .05$ ), indicating that people with a foreign nationality who use foreign news sources evaluate content-based similarity better than people who do not have a foreign nationality, or do not use foreign news sources.

Moving on to the results on *collaborative similarity*, we also see nine significant characteristics in the final model. The most prominent characteristics to predict users' evaluation of collaborative similarity are age ( $\beta = -0.232, p < .05$ ), news interests ( $\beta = -0.189, p < .05$ ;  $\beta = -0.073, p < .05$ ;  $\beta = -0.181, p < .05$ ), news information overload ( $\beta = 0.166, p < .05$ ) and online news participation ( $\beta = 0.154, p < .05$ ). Although these characteristics show some similarities with the predictors of content-based similarity, several differences can be noted. More specifically, online news participation and technology optimism ( $\beta = 0.131, p < .05$ ) were in our sample not found to be predicting content-based similarity, yet both are significantly and positively associated with collaborative similarity. Moreover, a difference can be found in the relationships' strength of news information overload. Here, we see a stronger relationship with collaborative similarity.

Then, looking to the final regression model of *content-based diversity*, we found eight significant characteristics. The most prominent predictors for a positive evaluation of content-based diversity are concerns regarding missing challenging viewpoints ( $\beta = 0.198, p < .05$ ), news interests ( $\beta = -0.113, p < .05$ ;  $\beta = -0.013, p < .05$ ) and news information overload ( $\beta = 0.127, p < .05$ ). In contrast to the similarity-based mechanisms, we find different predictors for diversity-based mechanisms. For instance, the use of foreign news sources did not score significant in the previous regressions, but is for this dependent variable an interesting predictor. Moreover, we also observe other types of statistical relationships. A case in point is the concerns regarding missing challenging viewpoints that are, on the one hand, negatively associated with content-based similarity and, on the other hand, positively associated with content-based diversity. Combining these two results, this means that the more people have concerns about missing challenging viewpoints, the more negative they are towards content-based similarity, and the more positive they are towards content-based diversity. However, we also see some similarities. For instance, the independent variable 'news information overload' is a significant predictor for all dependent variables. This result indicates that the more people experience a news information overload, the better any type of algorithmic news selection is perceived.

Finally, our analysis also shows that the technological and cognitional variables have the largest explanatory power compared to socio-demographic or news use variables. In particular, looking to the adjusted  $R^2$  (see Table 2), which explains how much variance is explained by each block, we see that the highest adjusted  $R^2$  for each regression model is situated in the third and last block.

**Table 2.** Linear and hierarchical regression analysis: standardized regression coefficients

Dependent variables		Content-based similarity			Collaborative similarity			Content-based diversity		
Regression models		(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)
Socio-demographic variables	Age	-0.227*	-0.222*	-0.246*	-0.243*	-0.223*	-0.232*	-0.068	-0.059	-0.016
	Gender (male)	-0.026	-0.050	-0.043	0.047	0.008	0.021	-0.061	-0.056	-0.085*
	Nationality (foreigner)	0.107*	0.185*	0.206*	0.027	0.065	0.071	0.055	0.056	0.024
	Nationality parents (nationals)	-0.003	0.004	0.007	-0.031	-0.024	-0.026	0.033	0.024	0.026
	Higher education (no)	0.064	0.076*	0.072*	-0.019	0.021	0.009	-0.027	-0.068	-0.076*
	Active employment (no)	-0.055	-0.068	-0.053	-0.111*	-0.118*	-0.098*	-0.009	-0.014	-0.022
News use and interest	News interest 2: political interested		-0.163*	-0.136*		-0.212*	-0.189*		-0.039	-0.067
	News interest 3: moderate interested		-0.078*	-0.082*		-0.072*	-0.073*		-0.126*	-0.113*
	News interest 4: low interested		-0.137*	-0.116*		-0.203*	-0.181*		-0.112*	-0.103*
	News use 2: moderate news omnivore		0.017	0.016		0.038	0.037		0.064	0.065
	News use 3: highly traditionalist		-0.010	0.033		0.043	0.085		-0.128	-0.183
	News use 4: moderate traditionalist		-0.008	0.021		0.006	0.044		-0.061	-0.083
	News use (in minutes)		0.010	0.015		0.025	0.038		0.066	0.075*
	Use of foreign news sources (no)		-0.037	-0.004		-0.023	0.010		0.081*	0.076*
	Interaction effect: nationality		0.101*	0.103*		0.041	0.029		0.012	-0.005
	Having paid for online news (no)		0.069*	0.067*		0.006	-0.004		0.054	0.041
	Mobile news use (no)		0.006	-0.020		0.040	0.007		-0.006	0.019
	Online news use (no)		0.038	0.051		0.129	0.132		-0.122	-0.158
Online news participation (no)		0.045	0.021		0.185*	0.154*		0.046	0.043	
Technological and cognitional variables	News information overload			0.111*			0.166*			0.127*
	Technology optimism			0.057			0.131*			-0.049
	Trust in news media			0.112*			0.087*			0.029
	Concerns: privacy			-0.073*			-0.050			-0.060
	Concerns: important information			0.011			0.006			0.016
	Concerns: challenging viewpoints			-0.134*			-0.076			0.198*
Observations		937	937	937	937	937	937	937	937	937
R <sup>2</sup>		0.052	0.099	0.145	0.047	0.151	0.201	0.009	0.059	0.111
Adjusted R <sup>2</sup>		0.046	0.081	0.121	0.041	0.133	0.179	0.003	0.038	0.086
Residual Std. Error		0.976	0.958	0.937	0.976	0.928	0.904	0.994	0.976	0.951
F Statistic		8.388*	5.268*	6.085*	7.588*	8.456*	9.052*	1.449*	2.910*	4.494*

Note: 1. coefficient is significantly different from zero (\*) when p-value < .05;

2. in case of dummy variables (i.e., nominal variables with two categories), reference category is placed between brackets;

3. for cluster variables (i.e., news use and news interest) the following categories are used as reference category: news omnivore and all-round interested

4. for full overview categories cluster variables: see Appendix 12 and 13;

5. five missing values for socio-demographic variables resulting in 937 observations.

## Discussion

In this article, we investigated which type of news recommender people prefer and why they do so. As such, it aimed to explain how audiences evaluate different news selection mechanisms that underlie news recommenders. In contrast to previous scholarly work on audience evaluations, our work distinguishes *between* different news recommenders (rather than recommender versus human/editorial recommendation, see e.g., Araujo et al., 2020; Thurman et al., 2019). We therefore reviewed the existing literature on algorithmic perceptions and explored the use of algorithms in the news environment. Based on this review, we distinguished between three news selection mechanisms, namely between content-based similarity, content-based diversity and collaborative similarity. Recently, diversity-based news algorithms are gaining scholarly attention recommenders (e.g., Helberger, 2019; Helberger et al., 2018), yet insights in how audiences perceive them remained unclear. We outlined several user characteristics that might explain why audiences prefer certain news selection mechanisms more than another. Subsequently, to map these evaluations and user characteristics, we conducted a survey in Ghent among 943 citizens. Based on the results of this survey, some conclusions can be made.

To start, we found that audiences generally prefer content-based similarity over collaborative similarity and content-based diversity. This preference was clearly visible in that relatively high percentages were observed of people who (strongly) agreed with survey items covering content-based similarity. For news organizations, this result confirms the commercial rationale that underlies their current news recommenders, that is, provide people with news in line with their interests. By using content-based similarity as a dominant news selection mechanism, they indeed give audiences the news they want. However, this result also indicates that, when the choice of news selection mechanism is given to the people, they have a tendency to prefer news articles that interests them and thus possibly circumvent news articles they are not familiar with. This implies the risk of being selectively exposed and favor information that reinforces pre-existing views, similar to the risks of news consumption in a non-algorithmic context (Knobloch-Westerwick, 2014).

With these risks in mind, policymakers as well as news organizations should consider implementing ways to reduce the risks of similarity-based news selection mechanisms. These may include features such as the 'serendipity'-concept where news articles are occasionally recommended by chance to keep audiences surprised and engaged (De Pessemier et al., 2014; Schönbach, 2007). A more profound idea to mitigate these risks lies in the idea of content-based diversity in which news recommenders are optimized to make individuals read in a more diverse way. Although our analysis points out that this design is not preferred by news audiences, it shows that there is still a considerable amount of people that are not reluctant to this design mechanism. This is especially true for news selection based on the heterogeneity of viewpoints in the news. This result should stimulate universities as well as public service media to develop a public service algorithm that should ideally focuses on the heterogeneity of viewpoints rather than topics.

However, to address a larger audience, diversity-based news recommenders should take into account the appealing features that underlie similarity-based news recommenders. As such, we plea for personalized news diversity. This means that news recommenders should not randomly select news articles, but should take the audiences' news interests and use into account when selecting news articles. For instance, for people who dislike politics news or foreign affairs, news recommender should not simply start their recommendations with these topics. Instead, they should find touch-points in the person's news use or interest, and subsequently guide him or her into a more diverse news diet. This may start, for instance, with news articles about education or economics, which may lie in the closer area of someone's interests, but significantly differs from political news. As such, news recommenders may slowly guide audiences into a more diverse news diet. We call this approach

'personalized diversity' in which the aim is to make news consumption more diverse by using the personalization features that underlie similarity-based news recommenders.

Moreover, and linked to the latter, more insights on the role of UX or interface design in news recommendations would be valuable. In particular, by using simple design strategies such as the use of (dis)like- or read later-buttons (e.g., Schnabel et al., 2020) or flexible preference settings (e.g., Harambam et al., 2019) news recommenders can give users more control which, in turn, may also lead to more explicit feedback of what users want. As such, news recommenders can find more easily an optimal balance of recommending articles that are just diverse enough to achieve the desired effect, without causing irritation or other back-fire effects.

Furthermore, when considering why audiences prefer certain news selection mechanics more than another, the results of our study show a range of user characteristics that have been found significant. Most of these results align with previous studies on content-based and collaborative news selection. For instance, Fletcher and Nielsen (2019) too found that approval for algorithmic selection is stronger amongst younger people and that those with a high level of interest in soft news topics are more comfortable with news selection based on what their peers have consumed. Similar to this result, we also identified effects of age, news interest and several other news use and socio-demographic variables. However, in addition to these characteristics, our study showed that several user characteristics related to the technological and cognitive nature of algorithmic news selection are on par, and even more important, to explain how audience evaluations differ. In particular, our analysis shows that concepts such as news information overload or technology optimism have a larger explanatory power compared to socio-demographic or news use variables. Especially news information overload is a powerful predictor of how algorithmic news selection is perceived. This means that the more people feel overwhelmed by the amount of news, the more they approve algorithmic news selection in general. This result ties into the idea of news recommenders as an individual filter for the growing abundance of online information (Aljukhadar et al., 2012; Beam & Kosicki, 2014; Dylko et al., 2017).

#### *Limitations and future research*

As our study focused on attitudes towards different news selection mechanisms, it is important to note that the actual behavior of audiences might not correspond with their expressed attitudes. In particular, when using a diversity-based news recommender, audiences might be confronted with the cognitive and affective difficulties related to receiving counter-attitudinal news. As such, they might not select the recommended news articles and draw back to their selective news habits. It is therefore important to conduct further research on how audiences actually react to the different news selection mechanisms and how this differs with their expressed attitudes. This will give us more insights into the effects of news recommenders on audiences' reading behavior as well as on the cognitive and affective difficulties related to the formation of a diverse news diet.

Furthermore, we would also like to point attention to the conceptualization of content-based diversity. In particular, we noticed that content-based diversity can be interpreted in various ways, similar to what we see with the general term 'news diversity' (Joris et al., 2020). In most studies on news recommenders, diversity refers to the internal differences within the set of recommendation items (Castells et al., 2015), as a feature that makes sure that the similarity among items in a recommendation set is minimized (Karimi et al., 2018). This conceptualization differs with our understanding that focuses on the differences within the user's reading behavior (i.e., exposure diversity) and the selection of news articles of which the user has read too little. Scholars should be aware for such slight, but important interpretation differences as they may impact how the news recommender system's output looks like, and subsequently, how audiences perceive these systems. Qualitative research may also provide important contributions on the latter area, to understand how audiences interpret and give meaning to news recommenders and their news selection principles.

Attention should also be paid to how content-based diversity is operationalized. In our study, we looked at the output of a diversity-based news recommender and the biases in people's news selection choices. More specifically, we look at what type of news content people generally read too little and how a diversity-based news recommender would compensate these biases in order to achieve exposure diversity. As previous research has found that people generally select pro-attitudinal information over counter-attitudinal information (D'Alessio & Allen, 2002; Hart et al., 2019), we expected that a diversity-based news recommender would primarily provide counter-attitudinal stories in terms of viewpoints and topics. However, as Hart et al. (2019) also explains, in some circumstances (e.g., debate one issue), people may search for counter-attitudinal content with the objective of providing counterarguments against it. In such cases, diversity-based news recommenders may also provide pro-attitudinal information.

The latter example shows that the lines between pro-attitudinal and counter-attitudinal content may be less distinct than we assumed in our study. However, at this point, we also want to stress the importance of accepting counter-attitudinal information in the context of exposure diversity. In particular, when people are open to receive viewpoints they do not agree, they will also be likely to accept the cognitive dissonance that comes along with reading counter-attitudinal information. As cognitive dissonances are argued to be one of the most important hindrances that prevent people to expose themselves to diverse information (Knobloch-Westerwick, 2014), we strongly believe that people who are open to receiving counter-attitudinal stories will also favor the promotion of exposure diversity by means of algorithmic news selection.

As a last limitation, we want to note that our population sample is focused on citizens living in Ghent. Although Ghent is a relatively large city in Belgium, the findings of this study cannot be generalized to people living in other cities in Belgium or other countries in Europe as the cultural context may significantly differ and may affect how people perceive and consume news.

## **Research details**

1. Acknowledgements: We would like to thank Judith Vermeulen (Ghent University), prof. dr. Eva Lievens (Ghent University), Hanne Elsen (DPO Ghent University) and Seppe Vansteelant (DPO City of Ghent) for their contributions to our study protocol in which all aspects regarding data protection and privacy were covered. In addition, we also thank Regina Debusschere (student Ghent University) who helped us to conduct home visits and to stimulate people with a foreign nationality to participate in our study.
2. Funding: This work was supported by Ghent University under Grant BOFGOA2018000601.
3. Disclosure statement: No potential conflict of interest was reported by the authors.
4. Data availability statement: The data that support the findings of this study are available from the corresponding author, upon reasonable request.

## Bibliography

- Aljukhadar, M., Senecal, S., & Daoust, C.-E. (2012). Using recommendation agents to cope with information overload. *International Journal of Electronic Commerce*, 17(2), 41-70.
- Araujo, T., Helberger, N., Kruijkemeier, S., & de Vreese, C. H. (2020). In AI we trust? Perceptions about automated decision-making by artificial intelligence. *AI & SOCIETY*.
- Beam, M. A., & Kosicki, G. M. (2014). Personalized news portals: Filtering systems and increased news exposure. *Journalism & Mass Communication Quarterly*, 91(1), 59-77.
- Bodó, B. (2019). Selling News to Audiences—A Qualitative Inquiry into the Emerging Logics of Algorithmic News Personalization in European Quality News Media. *Digital Journalism*, 7(8), 1054-1075.
- Bodó, B., Helberger, N., Eskens, S., & Möller, J. (2019). Interested in Diversity: The role of user attitudes, algorithmic feedback loops, and policy in news personalization. *Digital Journalism*, 7(2), 206-229.
- Bryman, A. (2015). *Social research methods*. Oxford: Oxford University Press.
- Bucher, T. (2017). The algorithmic imaginary: exploring the ordinary affects of Facebook algorithms. *Information, Communication & Society*, 20(1), 30-44
- Burton, J. W., Stein, M. K., & Jensen, T. B. (2020). A systematic review of algorithm aversion in augmented decision making. *Journal of Behavioral Decision Making*, 33(2), 220-239.
- Castells, P., Hurley, N. J., & Vargas, S. (2015). Novelty and diversity in recommender systems. In *Recommender systems handbook* (p. 881-918): Springer.
- Castelo, N., Bos, M. W., & Lehmann, D. R. (2019). Task-dependent algorithm aversion. *Journal of Marketing Research*, 56(5), 809-825.
- Chen, Y., Wu, C., Xie, M., & Guo, X. (2011). Solving the sparsity problem in recommender systems using association retrieval. *Journal of computers*, 6(9), 1896-1902.
- D'Alessio, D., & Allen, M. (2002). Selective exposure and dissonance after decisions. *Psychological Reports*, 91(2), 527-532.
- De Pessemier, T., Dooms, S., & Martens, L. (2014). Context-aware recommendations through context and activity recognition in a mobile environment. *Multimedia Tools and Applications*, 72(3), 2925-2948.
- DeVellis, R. F. (2016). *Scale development: Theory and applications* (Vol. 26): Sage publications.
- Dietvorst, B. J., Simmons, J. P., & Massey, C. (2018). Overcoming algorithm aversion: People will use imperfect algorithms if they can (even slightly) modify them. *Management Science*, 64(3), 1155-1170.
- Dillman, D. A., Smyth, J. D., & Christian, L. M. (2014). *Internet, phone, mail, and mixed-mode surveys: the tailored design method*: John Wiley & Sons.
- Dylko, I., Dolgov, I., Hoffman, W., Eckhart, N., Molina, M., & Aaziz, O. (2017). The dark side of technology: An experimental investigation of the influence of customizability technology on online political selective exposure. *Computers in Human Behavior*, 73, 181-190.
- Ekstrand, M. D., Harper, F. M., Willemsen, M. C., & Konstan, J. A. (2014). User perception of differences in recommender algorithms. Paper presented at the Proceedings of the 8th ACM Conference on Recommender systems, Foster City, Silicon Valley, California, USA.
- Eslami, M., Karahalios, K., Sandvig, C., Vaccaro, K., Rickman, A., Hamilton, K., & Kirlik, A. (2016). First I "like" it, then I hide it: Folk Theories of Social Feeds. Paper presented at the Proceedings of the 2016 CHI conference on human factors in computing systems.
- Farrell, A. M. (2010). Insufficient discriminant validity: A comment on Bove, Pervan, Beatty, and Shiu (2009). *Journal of business research*, 63(3), 324-327.
- Festinger, L. (1957). *A theory of cognitive dissonance* (Vol. 2): Stanford university press.
- Fields, B., Jones, R., & Cowlshaw, T. (2018). *The case for public service recommender algorithms*. Paper presented at the FATREC Workshop.

- Fletcher, R., & Nielsen, R. K. (2019). Generalised scepticism: how people navigate news on social media. *Information, Communication & Society*, 22(12), 1751-1769.
- Fletcher, R., & Park, S. (2017). The impact of trust in the news media on online news consumption and participation. *Digital Journalism*, 5(10), 1281-1299.
- Good, N., Schafer, J. B., Konstan, J. A., Borchers, A., Sarwar, B., Herlocker, J., & Riedl, J. (1999). Combining collaborative filtering with personal agents for better recommendations. *AAAI/IAAI*, 439.
- Haim, M., Graefe, A., & Brosius, H.-B. (2018). Burst of the filter bubble? Effects of personalization on the diversity of Google News. *Digital Journalism*, 6(3), 330-343.
- Hallin, D. C., & Mancini, P. (2004). Comparing media systems: Three models of media and politics.
- Harambam, J., Bountouridis, D., Makhortykh, M., & Van Hoboken, J. (2019, September). Designing for the better by taking users into account: a qualitative evaluation of user control mechanisms in (news) recommender systems. In *Proceedings of the 13th ACM Conference on Recommender Systems* (p. 69-77).
- Hart, W., Albarracín, D., Eagly, A. H., Brechan, I., Lindberg, M. J., & Merrill, L. (2009). Feeling validated versus being correct: a meta-analysis of selective exposure to information. *Psychological bulletin*, 135(4), 555.
- Helberger, N. (2011). Diversity by design. *Journal of information policy*, 1, 441-469.
- Helberger, N. (2012). Exposure diversity as a policy goal. *Journal of Media Law*, 4(1), 65-92.
- Helberger, N. (2019). On the democratic role of news recommenders. *Digital Journalism*, 1-20.
- Helberger, N., Karppinen, K., & D'Acunto, L. (2018). Exposure diversity as a design principle for recommender systems. *Information, Communication & Society*, 21(2), 191-207.
- Jones, B., & Jones, R. (2019). Public Service Chatbots: Automating Conversation with BBC News. *Digital Journalism*, 7(8), 1032-1053.
- Joris, G., De Grove, F., Van Damme, K., & De Marez, L. (2020). News diversity reconsidered : a systematic literature review unraveling the diversity in conceptualizations. *Journalism Studies*, 21(13), 1893–1912.
- Joris, G., Colruyt, C., Vermeulen, J., Vercoutere, S., De Grove, F., Van Damme, K., ... & Martens, L. (2019, August). News Diversity and Recommendation Systems: Setting the Interdisciplinary Scene. In *Proceedings of the IFIP International Summer School on Privacy and Identity Management* (p. 90-105).
- Karimi, M., Jannach, D., & Jugovac, M. (2018). News recommender systems – Survey and roads ahead. *Information Processing & Management*, 54(6), 1203-1227.
- Knobloch-Westerwick, S. (2014). *Choice and preference in media use: Advances in selective exposure theory and research*: Routledge.
- Kutner, M. H., Nachtsheim, C. J., Neter, J., & Li, W. (2005). *Applied linear statistical models* (Vol. 5): McGraw-Hill Irwin New York.
- Laganà, F., Elcheroth, G., Penic, S., Kleiner, B., & Fasel, N. (2013). National minorities and their representation in social surveys: which practices make a difference? *Quality & Quantity*, 47(3), 1287-1314.
- Lee, S. K., Lindsey, N. J., & Kim, K. S. (2017). The effects of news consumption via social media and news information overload on perceptions of journalistic norms and practices. *Computers in Human Behavior*, 75, 254-263.
- Liang, T.-P., Lai, H.-J., & Ku, Y.-C. (2006). Personalized content recommendation and user satisfaction: Theoretical synthesis and empirical findings. *Journal of Management Information Systems*, 23(3), 45-70.
- Liu, J., Dolan, P., & Pedersen, E. R. (2010). *Personalized news recommendation based on click behavior*. Paper presented at the Proceedings of the 15th international conference on Intelligent user interfaces.
- Logg, J. M. (2017). Theory of machine: When do people rely on algorithms? *Harvard Business School working paper series# 17-086*.



- Logg, J. M., Minson, J. A., & Moore, D. A. (2019). Algorithm appreciation: People prefer algorithmic to human judgment. *Organizational Behavior and Human Decision Processes*, 151, 90-103.
- Möller, J., Trilling, D., Helberger, N., & van Es, B. (2018). Do not blame it on the algorithm: an empirical assessment of multiple recommender systems and their impact on content diversity. *Information, Communication & Society*, 21(7), 959-977.
- Nagar, N. a. (2011). The loud public: The case of user comments in online news media (Doctoral dissertation). Retrieved from <https://isoj.org/wp-content/uploads/2018/01/Nagar09.pdf>
- Napoli, P. M. (2006). The Marketplace of Ideas Metaphor In Communications Regulation. *Journal of Communication*, 49(4), 151-169.
- Newman, N., Fletcher, R., Schulz, A., Andi, S. & Nielsen, R. (2020). *Reuters institute digital news report 2020*. Reuters Institute for the Study of Journalism.
- Parasuraman, A., & Colby, C. L. (2015). An updated and streamlined technology readiness index: TRI 2.0. *Journal of service research*, 18(1), 59-74.
- Pentina, I., & Tarafdar, M. (2014). From “information” to “knowing”: Exploring the role of social media in contemporary news consumption. *Computers in Human Behavior*, 35, 211-223.
- Picone, I., & Donders, K. (2020). Reach or trust optimisation? A citizen trust analysis in the Flemish public broadcaster VRT. *Media and Communication*, 8(3), 348-358.
- Saxe, R., & Kanwisher, N. (2003). People thinking about thinking people: the role of the temporoparietal junction in “theory of mind”. *Neuroimage*, 19(4), 1835-1842.
- Schmitt, J. B., Debbelt, C. A., & Schneider, F. M. (2018). Too much information? Predictors of information overload in the context of online news exposure. *Information, Communication & Society*, 21(8), 1151-1167.
- Schnabel, T., Amershi, S., Bennett, P. N., Bailey, P., & Joachims, T. (2020, July). The Impact of More Transparent Interfaces on Behavior in Personalized Recommendation. In *Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval* (p. 991-1000).
- Schönbach, K. (2007). ‘The own in the foreign’: reliable surprise-an important function of the media? *Media, Culture & Society*, 29(2), 344-353.
- Stroud, N. J., Van Duyn, E., & Peacock, C. (2016). News commenters and news comment readers. *Microsoft Word-Egaging News Project*, 1-21.
- Sundar, S. S. (2008). *The MAIN model: A heuristic approach to understanding technology effects on credibility*: MacArthur Foundation Digital Media and Learning Initiative.
- Sundar, S. S., & Marathe, S. S. (2010). Personalization versus customization: The importance of agency, privacy, and power usage. *Human Communication Research*, 36(3), 298-322.
- Takken, W. (2018). NPO wil 'bubbel van de kijker opblazen'. *NRC*. Retrieved from <https://www.nrc.nl/nieuws/2018/10/08/npo-wil-bubbel-van-de-kijker-opblazen-a2376340>
- Thurman, N., Möller, J., Helberger, N., & Trilling, D. (2019). My friends, editors, algorithms, and I: Examining audience attitudes to news selection. *Digital Journalism*, 1-23.
- Van den Bulck, H., & Moe, H. (2018). Public service media, universality and personalisation through algorithms: mapping strategies and exploring dilemmas. *Media, Culture & Society*, 40(6), 875-892.
- Kristin Van Damme, Anissa All, Lieven De Marez & Sarah Van Leuven (2019) 360° Video Journalism: Experimental Study on the Effect of Immersion on News Experience and Distant Suffering, *Journalism Studies*, 20(14), 2053-2076.
- Vandendriessche, K., & De Marez, L. (2019). *Imec.digimeter 2019: Digitale mediatrends in Vlaanderen*. Gent: imec.
- Vīķe-Freiberga, V., Däubler-Gmelin, H., Hammersley, B., & Maduro, L. M. P. P. (2013). A free and pluralistic media to sustain European democracy. *Report, High Level Group on Media Freedom and Media Pluralism*.
- Yale, R. N., Jensen, J. D., Carcioppolo, N., Sun, Y., & Liu, M. (2015). Examining first-and second-order factor structures for news credibility. *Communication Methods and Measures*, 9(3), 152-169.

- Yuan, E. (2011). News consumption across multiple media platforms: A repertoire approach. *Information, Communication & Society*, 14(7), 998-1016.
- Zuiderveen Borgesius, F., Trilling, D., Möller, J., Bodó, B., de Vreese, C. H., & Helberger, N. (2016). Should we worry about filter bubbles? *Internet Policy Review*, 5(1), 1 - 16.

## Appendices

### Appendix 1. Sample characteristics (absolute and relative statistics)

	Foreign nationality			Domestic nationality			Total			
	M	F	Total	M	F	Total	M	F	Unknown	Total
18-24	82(1.64%)	82(1.64%)	164(3.28%)	182(3.64%)	173(3.46%)	355(7.11%)	264(5.29%)	255(5.11%)	0(0.00%)	519(10.39%)
25-34	216(4.32%)	201(4.02%)	417(8.35%)	381(7.63%)	379(7.59%)	760(15.22%)	597(11.95%)	580(11.61%)	0(0.00%)	1,177(23.56%)
35-44	179(3.58%)	183(3.66%)	362(7.25%)	278(5.57%)	261(5.23%)	539(10.79%)	457(9.15%)	444(8.89%)	0(0.00%)	901(18.04%)
45-54	126(2.52%)	116(2.32%)	242(4.84%)	260(5.21%)	239(4.76%)	498(9.97%)	386(7.73%)	354(7.09%)	0(0.00%)	740(14.81%)
55-64	69(1.38%)	55(1.10%)	124(2.48%)	273(5.47%)	282(5.65%)	555(11.11%)	342(6.85%)	337(6.75%)	0(0.00%)	679(13.59%)
65-74	28(0.56%)	25(0.50%)	53(1.06%)	215(4.30%)	244(4.88%)	459(9.19%)	243(4.86%)	269(5.39%)	0(0.00%)	512(10.25%)
75-84	11(0.22%)	9(0.18%)	20(0.40%)	113(2.26%)	195(3.90%)	308(6.17%)	124(2.48%)	204(4.08%)	0(0.00%)	328(6.57%)
85+	2(0.04%)	2(0.04%)	4(0.08%)	50(1.00%)	85(1.70%)	135(2.70%)	52(1.04%)	87(1.74%)	0(0.00%)	139(2.78%)
Total	713(14.27%)	673(13.47%)	1,386(27.75%)	1,752(35.08%)	1,857(37.18%)	3,609(72.25%)	2,465(49.35%)	2,530(50.65%)	0(0.00%)	4,995(100.00%)

## Appendix 2. Study protocol

PROTOCOL VOOR DE ELEKTRONISCHE MEDEDELING VAN PERSOONSgegevens VAN Stad Gent NAAR  
UGent  
in het kader van het onderzoeksproject NewsDNA

5/09/2019

*Dit protocol wordt gesloten conform artikel 8, §1, van het decreet van 18 juli 2008 betreffende het elektronische bestuurlijke gegevensverkeer.*

### TUSSEN

**Stad Gent**, met zetel ten stadhuize, Botermarkt 1, 9000 Gent, vertegenwoordigd door het college van burgemeester en schepenen, voor wie tekenen: de heer Mathias De Clercq, burgemeester, en mevrouw Mieke Hullebroeck, algemeen directeur, in uitvoering van het collegebesluit d.d. 5 september 2019

hierna: "Stad Gent";

### EN

**UNIVERSITEIT GENT**, openbare instelling met rechtspersoonlijkheid, opgericht ingevolge het bijzonder decreet d.d. 26 juni 1991 betreffende de Universiteit Gent en het Universitair Centrum Antwerpen (B.S. 29 juni 1991, zoals later gewijzigd), met bestuurszetel te 9000 Gent, Sint-Pietersnieuwstraat 25, met KBO-nummer 0248.015.142, rechtsgeldig vertegenwoordigd door de Rector, prof. dr. Rik Van de Walle, die Prof. Dr. Frederik de Grove en Prof. Dr. Lieven De Marez aanduidt als verantwoordelijken voor de uitvoering van dit Protocol.

hierna: Universiteit Gent ('UGent');

Stad Gent en UGent worden hieronder ook wel afzonderlijk aangeduid als een "partij" of gezamenlijk als de "partijen";

### NA TE HEBBEN UITEENGEZET

- A. Stad Gent is een lokaal bestuur;
- B. UGent is een openbare instelling met rechtspersoonlijkheid, opgericht ingevolge het bijzonder decreet d.d. 26 juni 1991 betreffende de Universiteit Gent en het Universitair Centrum Antwerpen;
- C. **C1. BESCHRIJVING VAN DE ALGEMENE CONTEXT:**

In het kader van het onderzoeksproject NewsDNA voert de onderzoeksgroep voor *Media, Innovatie en CommunicatieTechnologieën* (MICT-UGent, vakgroep Communicatiewetenschappen van de

faculteit Politieke en Sociale Wetenschappen) onder begeleiding van Prof. Dr. Frederik de Grove en Prof. Dr. Lieven De Marez onderzoek uit naar het nieuwsgebruik en nieuwsvertrouwen bij de Vlaamse bevolking. Het onderzoek heeft een duurtijd van 4 jaar (2018-2022) en beoogt fundamentele inzichten te verwerven in de relatie tussen mens, maatschappijen nieuwsmedia en de rol van digitale technologieën daarin. In samenwerking met andere faculteiten van de Universiteit Gent (i.e. Faculteit Ingenieurswetenschappen en Architectuur vertegenwoordigd door Prof. Dr. Luc Martens, Faculteit Letteren en Wijsbegeerte vertegenwoordigd door Prof. Dr. Veronique Hoste, Faculteit Recht en Criminologie vertegenwoordigd door Prof. Dr. Eva Lievens) stelt NewsDNA zelf ook het ontwikkelen van nieuwe digitale technologieën voorop, en met name een algoritme dat toelaat om diverse nieuwsartikelen aan te bevelen.

In het tweede levensjaar van het onderzoeksproject heeft onderzoeksgroep MICT de intentie om een grootschalige vragenlijst af te nemen in Gent. Het doel van deze vragenlijst is om onder meer een algemeen beeld te krijgen van het nieuwsgebruik en –vertrouwen in Vlaanderen en de nood van een diversiteitsalgoritme in te schatten. Het belang om hierover kennis te vermeerderen heeft zich de afgelopen jaren steeds meer gesteld door de stijgende invloed van digitale/sociale media op bepaalde fundamentele democratische processen zoals het (correct) informeren van de bevolking. Verschillende beleids(ondersteundende) organen, zoals de door de Europese Commissie in het leven geroepen deskundigengroep op hoog niveau inzake nepnieuws en online-desinformatie en het Comité van Ministers van de Raad van Europa, maken hierover ook melding in hun beleidsrapportering en benadrukken de meerwaarde van zulke onderzoeken voor de bevolking en de bredere samenleving.

Gezien de beperkingen in tijd en ruimte, wil onderzoeksgroep MICT volgens de klassieke methodes van sociaalwetenschappelijk onderzoek niet de gehele bevolking, maar een beperkte selectie daarvan, die een volledig getrouwe afspiegeling vormt van de samenleving in kwestie, uitnodigen om deel te nemen (per brief of, in geval van een tekort aan respons, naar aanleiding van een huisbezoek) en bevragen. Om dit concreet te verwezenlijken, vraagt onderzoeksgroep MICT aan de Stad Gent een representatieve steekproeftrekking op basis van het bevolkingsregister uit te voeren en haar (MICT) de aldus verkregen gegevens mee te delen. Het gaat concreet om de naam-, contact- en een aantal socio-demografische gegevens van 5.000 personen verspreid over circa 142 buurten in de Stad Gent (zie ook artikel 2 en 3 van dit Protocol).

Deze samenwerking vergroot in grote mate de kans om een doorsnede van de bevolking te bevragen en de conclusies die uit dit onderzoek voortvloeien, te veralgemenen naar een grotere populatie. Voorgaande onderzoeken georganiseerd door MICT zelf in bibliotheek De Krook en op digitale platformen (e.g. Facebook, Twitter) toonden aan dat dit eerder een uitzondering op de regel vormt en dat bepaalde groepen in de samenleving eerder onder- en oververtegenwoordigd zijn in sociaalwetenschappelijk onderzoek. Een samenwerking met de Stad Gent is vanuit dat sociaal-methodologisch perspectief van groot belang.

## **C2. VOORSTELLING VAN HET VERLOOP VAN DE GEPLANDE GEGEVENSSTROOM.**

In een eerste fase van het onderzoek zal er een digitale bestandsoverdracht plaatsvinden van de Stad Gent naar de UGent. De gegevens die tijdens deze bestandsoverdracht worden doorgegeven zijn het resultaat van een representatieve steekproeftrekking die wordt uitgevoerd door de Stad Gent op basis van het bevolkingsregister. Het gaat concreet om de naam-, contact- en socio-demografische gegevens (incl. statistische sector) van 5.000 personen verspreid over circa 142 buurten in de Stad Gent. De steekproefgrootte (N = 5.000) is gebaseerd op een verwachte responsratio van minimaal 10 procent en een minimum steekproefgrootte van 500 personen. De samenstelling van de steekproef

is gebaseerd op de verdeling die in de huidige samenleving en met name Groot-Gent aanwezig is op vlak van geslacht, leeftijd, nationaliteit en opleidingsniveau en in de publieke databanken van de stad Gent wordt weergegeven voor het jaar 2018 (zie bijlage A voor steekproefkader).

Om de digitale gegevensoverdracht optimaal te beveiligen, worden er verschillende organisatorische en technische maatregelen genomen door de UGent. Zo zullen de persoonsgegevens via een beveiligd internetprotocol (sFTP) doorgestuurd worden en opgeslagen worden in een met wachtwoord beveiligde map van Prof. dr. Frederik De Grove op de UGent-server. Deze gegevens worden apart bewaard, los van de socio-demografische gegevens die in een latere fase ook worden verzameld door middel van de vragenlijst. De databases kunnen wel aan elkaar gekoppeld worden door middel van een pseudoniem, dit om (1) de rechten van respondenten te kunnen garanderen (bv. recht op inzage of gegevenswissing) en (2) te kunnen bepalen bij wie huisbezoeken dienen uitgevoerd te worden (zie *infra*).

De persoonsgegevens zullen via een beveiligd internetprotocol (sFTP) doorgestuurd worden tussen Stad Gent en UGent.

- D. De partijen wensen overeenkomstig artikel 8, §1, van het decreet van 18 juli 2008 betreffende het elektronische bestuurlijke gegevensverkeer een protocol te sluiten met betrekking tot de elektronische mededeling van persoonsgegevens. Dat protocol wordt bekendgemaakt op de website van beide partijen.
- F. De functionaris voor gegevensbescherming van Stad Gent heeft op 29 augustus 2019 advies met betrekking tot een ontwerp van dit protocol gegeven.
- G. De functionaris voor gegevensbescherming van UGent heeft op 14 augustus 2019 advies met betrekking tot een ontwerp van dit protocol gegeven.

#### **WORDT OVEREENGEKOMEN WAT VOLGT:**

##### **Artikel 1: Onderwerp**

In dit protocol worden de voorwaarden en modaliteiten van de elektronische mededeling van de persoonsgegevens zoals omschreven in artikel 3 door Stad Gent aan UGent uiteengezet.

##### **Artikel 2: Rechtvaardigingsgronden van zowel de mededeling als de inzameling van de persoonsgegevens**

De beoogde gegevensverwerking door de UGent gebeurt op grond van artikel 6, lid 1, e) AVG en is noodzakelijk voor de vervulling van een taak van algemeen belang.

Dat wetenschappelijk onderzoek verricht door de UGent kan worden beschouwd als "noodzakelijk voor de vervulling van een taak van algemeen belang" blijkt uit het feit dat de UGent:

- kan worden beschouwd als “een instelling met een publieke taak” als vermeld in artikel I.3, 6°, van het Bestuursdecreet van 7 december 2018 (“Bestuursdecreet”)<sup>1</sup>, op basis van de daarin bepaalde criteria. De UGent behoort immers niet tot de Vlaamse overheid (artikel I.3.1° Bestuursdecreet) of tot een lokale overheid (‘Bestuursdecreet’), maar (1) is opgericht met het specifieke doel te voorzien in behoeften van algemeen belang die niet van industriële of commerciële aard zijn\*; (2) bezit rechtspersoonlijkheid op grond van artikel 2, lid 1 van het Bijzonder Decreet betreffende de Universiteit Gent en het Universitair Centrum Antwerpen; en (3) wordt voor meer dan de helft gefinancierd door de Vlaamse overheid.
- \*is opgericht met het specifieke doel te voorzien in behoeften van algemeen belang, en met name (onder andere) het verrichten van wetenschappelijk onderzoek<sup>2</sup>, hetgeen blijkt uit:
  - het arrest nr. 221.642 van de Raad van State van 6 december 2012, waarin werd bevestigd dat de Katholieke Universiteit Leuven kon worden beschouwd als een “bestuursinstantie” op grond van artikel 3, 1°, c) van het intussen door het Bestuursdecreet opgeheven Decreet betreffende de openbaarheid van bestuur van 26 maart 2004 (‘Openbaarheidsdecreet’), en meer bepaald als een rechtspersoon die een taak van algemeen belang behartigt en beslissingen neemt die derden binden. Hetzelfde kan worden gezegd van de Universiteit Gent (die ook al als “bestuursinstantie” kon worden beschouwd op grond van artikel 3 1°, b) van het Openbaarheidsdecreet);
  - de Memorie van Toelichting betreffende het ontwerp van Bestuursdecreet zoals ingediend bij het Vlaams Parlement op 9 juli 2018. Daarin wordt uiteengezet dat aangezien “[e]n universele definitie van het begrip “behoefte van algemeen belang [niet bestaat], het wordt geval per geval ingevuld door de rechtspraak”. Op basis van “[e]nkele voorbeelden uit de casuïstiek van het Hof van Justitie: het ophalen en verwerken van huishoudelijk afval, het beheer van nationale bossen en de opbouw van de bosrijverheid, het beheer van een universiteit, openbare telecommunicatienetwerken, sociale woningbouw, lijkbezorging, het ontwikkelen van onroerende goederen in opdracht van een stad met het oog op terbeschikkingstelling aan technologieondernemingen, enzovoort” werd afgeleid dat “de uitgeoefende activiteit – minstens onrechtstreeks – een ruimere gemeenschap van personen ten goede [komt] en niet uitsluitend [is] gericht op behoeften van welbepaalde personen”. Het verrichten van wetenschappelijk onderzoek in het algemeen en het in dit protocol bedoelde onderzoek specifiek (zie *infra*) voldoet aan deze voorwaarde;
  - het feit dat artikel 4 aangaande de “universitaire zending” van het intussen opgeheven Decreet betreffende de universiteiten in de Vlaamse gemeenschap van 12 juni 1991 bepaalde dat “dat universiteiten, in het belang van de samenleving, terzelfdertijd werkzaam zijn op het gebied van het academisch onderwijs, het wetenschappelijk onderzoek en de wetenschappelijke dienstverlening”. Artikel II.18.,§1 van de Codex Hoger Onderwijs, waarin op heden (onder andere) de zending van universiteiten wordt uiteengezet, bepaalt in dat hogescholen en universiteiten, in het belang van de samenleving, werkzaam zijn op het gebied van hoger onderwijs. Artikel II.18,§2 voorziet dan weer dat universiteiten ook werkzaam zijn op het gebied

<sup>1</sup> Dit wordt bevestigd op de website van de Vlaamse Overheid: <https://overheid.vlaanderen.be/is-uw-organisatie-een-vlaamse-bestuursinstantie>.

<sup>2</sup> De opgelijste bepalingen kunnen fungeren als grondslag voor de verwerking die noodzakelijk is voor de vervulling van een taak van algemeen belang zoals voorgeschreven in overweging (45) van de algemene verordening gegevensbescherming.

van wetenschappelijk onderzoek. Dat niet meer expliciet wordt bepaald dat universiteiten wetenschappelijk onderzoek verrichten "in het belang van de samenleving" kan eerder worden beschouwd als een onnauwkeurigheid van de decreetgever dan een bewuste keuze. De uiteenzetting van de zending van hogescholen en universiteiten in verschillende paragrafen blijkt immers (enkel) te maken te hebben met het feit dat die wel "dezelfde kerntaken [bevat] maar [...] een gedifferentieerde inkleuring [vereist]" (Memorie van Toelichting betreffende het ontwerp van decreet betreffende de herstructurering van het hoger onderwijs in Vlaanderen van 4 maart 2003, opgeheven door de Codex Hoger Onderwijs).

Dat het in dit bedoelde onderzoek specifiek wordt verricht wegens "noodzakelijk voor vervulling van een taak van algemeen belang"<sup>3</sup> blijkt onder meer het opzet van NewsDNA op zich, namelijk: het bevorderen van nieuwsdiversiteit door middel van interdisciplinair onderzoek naar algoritmisch design, personalisatie en het publiek belang. Zoals reeds vermeld (zie *supra*) hebben verschillende beleids(ondersteunende) organen, zoals de door de Europese Commissie in het leven geroepen deskundigengroep op hoog niveau inzake nepnieuws en online-desinformatie en het Comité van Ministers van de Raad van Europa, in hun beleidsrapportering de nood aan dergelijke onderzoeken en de meerwaarde daarvan voor de bevolking en de bredere samenleving reeds benadrukt.

Er dient te worden benadrukt dat in een tweede fase van het onderzoek, met uitzondering van hun 'statistische sector', de socio-demografische gegevens van de deelnemende respondenten ook zullen worden verkregen op basis van de door hen ingevulde vragenlijst, en dus op basis van toestemming overeenkomstig artikel 6, lid 1, a) AVG.

Stad Gent heeft de opgevraagde gegevens oorspronkelijk verzameld voor volgende doeleinden:

- Elke gemeente moet de bevolkingsregisters houden met daarin ingeschreven alle personen die de gemeente als hoofdverblijfplaats hebben. (Wet van 19 juli 1991 betreffende de bevolkingsregisters, de identiteitskaarten, de vreemdelingenkaarten en de verblijfsdocumenten en tot wijziging van de wet van 8 augustus 1983 tot regeling van een Rijksregister van de natuurlijke personen, art. 1)
- Het houden van de registers behoort tot de bevoegdheid van het college van burgemeester en schepenen. (Koninklijk Besluit van 16 juli 1992 betreffende de bevolkingsregisters en het vreemdelingenregister, art. 4)

De UGent zal de opgevraagde naam- en contactgegevens verwerken voor volgende doeleinden (zie ook *infra* 'Artikel 3'):

- Het (her)adresseren aan en het (her)versturen per post van een persoonlijke uitnodiging naar de respondenten. De UGent zal de groep van respondenten (N = 5.000) uitnodigen voor deelname. Deze uitnodiging gebeurt respectievelijk door een uitnodigingsbrief (per post) en een herinneringsbrief (per post) te versturen naar het adres waar de respondent is gedomicilieerd. De respondent zal in deze brief persoonlijk gevraagd worden om via digitale weg deel te nemen aan het onderzoek. Elke respondent krijgt hiervoor een persoonlijke URL ter beschikking die overeenkomt met de unieke code of pseudoniem van de respondent;
- Het gericht laten verlopen van huisbezoeken: personen die niet hebben deelgenomen en waarvan geweten is dat ze behoren tot een groep die ondervertegenwoordigd is in de reële steekproef, worden aan huis bezocht door doctoraatsstudent Glen Joris (MICT). De uitvoering van deze fase is afhankelijk van de mate van (non-)respons die door de UGent

<sup>3</sup> Zie p. 27 van het advies 06/2014 over het begrip "gerechtvaardigd belang van de voor de gegevensverwerking verantwoordelijke" in artikel 7 van Richtlijn 95/46/EG van de Groep Gegevensbescherming Artikel 29.



wordt geregistreerd. Indien bijvoorbeeld bepaalde groepen (bv. oudere mannen met een andere dan de Belgische nationaliteit) uit de steekproef ondervetegenwoordigd zijn in de reële steekproef, is de UGent genoodzaakt om, conform de gangbare methodes in sociaalwetenschappelijk onderzoek, respondenten persoonlijk uit te nodigen. Daartoe hebben naast hun socio-demografische gegevens (zie *infra*) ook hun naam- en contactgegevens opnieuw nodig (koppeling nodig op basis van pseudoniem van database waarin naam- en contactgegevens worden bewaard en database waarin de socio-demografische gegevens worden bewaard).

UGent zal de opgevraagde socio-demografische gegevens verwerken voor volgende doeleinden (zie ook *infra* artikel 3):

- Het nagaan van de representativiteit van de uiteindelijke respondenten (gebruikmakende enkel van de database waarin de socio-demografische gegevens worden bewaard);
- Het analyseren van de non-respons (gebruikmakende enkel van de database waarin de socio-demografische gegevens worden bewaard);
- Het gericht laten verlopen van huisbezoeken (zie ook *supra*): de nood aan huisbezoeken is afhankelijk van de mate van (non-)respons. Om daarin inzicht te verwerven, hebben we de socio-demografische gegevens nodig van zowel de uiteindelijke als de vooropgestelde respondenten ten einde een vergelijking te kunnen maken (koppeling nodig op basis van pseudoniem van database waarin naam- en contactgegevens worden bewaard en database waarin de socio-demografische gegevens worden bewaard);
- Het bijkomend analyseren met het oog op academische valorisatie van het onderzoek. Indien de resultaten worden gepubliceerd in een academisch tijdschrift, zullen die geen gegevens bevatten op basis waarvan de respondenten kunnen worden geïdentificeerd (door pseudonimisering en hercodering van de variabele 'nationaliteit' naar 'belg/niet-belg').

Het doeleinde voor de verdere verwerking is wetenschappelijk onderzoek. Op deze verwerkingen geldt een uitzondering volgens artikel 5, 1, b), tweede zinsdeel van de algemene verordening gegevensbescherming: deze verwerkingen worden niet als onverenigbaar beschouwd met de oorspronkelijke doeleinden. Verder is een steekproef uit het bevolkingsregister de enige manier waarop de UGent een representatief staal van de Gentse bevolking kan bevragen.

**Artikel 3: De gevraagde persoonsgegevens en de categorieën en omvang van de gevraagde persoonsgegevens conform het proportionaliteitsbeginsel**

In onderstaande tabel wordt een overzicht gegeven van de verschillende persoonsgegevens die worden meegedeeld, alsook de verantwoording van de proportionaliteit en de bewaartermijn van de gegevens.

Het betreft geen persoonsgegevens als vermeld in artikel 9 en/of 10 van de algemene verordening gegevensbescherming. Indien dat wel het geval is, wordt dit gespecificeerd in onderstaande tabel.

Gegeven 1 <i>het concrete gegeven dat wordt meegedeeld. Als er veel</i>	Contactgegevens: Naam, Voornaam, Adres
--	--

<i>gegevens zijn, kunnen ze in clusters worden vermeld.</i>	
<i>Verantwoording proportionaliteit waarom elk gegeven noodzakelijk is voor het gevraagde doel. Als in artikel 1 en 2 verschillende doelen zijn opgegeven, aangeven voor welk doel het gegeven wordt meegedeeld</i>	Deze contactgegevens zijn noodzakelijk om een persoonlijke uitnodiging per post te versturen naar de respondent. De voor- en achternaam zullen in het bijzonder worden gebruikt om de respondent persoonlijk aan te spreken in de brief en duidelijkheid te verschaffen over wie van het gezin mag deelnemen aan het onderzoek. Het adres wordt tevens gebruikt om het huisbezoek gericht te laten verlopen.
<i>Gegeven 2 het concrete gegeven dat wordt meegedeeld. Als er veel gegevens zijn, kunnen ze in clusters worden vermeld.</i>	Socio-demografische gegevens: geslacht, leeftijd, nationaliteit, statistische sector
<i>Verantwoording proportionaliteit waarom elk gegeven noodzakelijk is voor het gevraagde doel. Als in artikel 1 en 2 verschillende doelen zijn opgegeven, aangeven voor welk doel het gegeven wordt meegedeeld</i>	Alle socio-demografische gegevens die zijn vermeld, zijn noodzakelijk om na te gaan of de uiteindelijke steekproef die de vragenlijst invult, representatief is. Indien dat niet het geval is, kunnen de onderzoekers een bepaalde groep in de bevolking (bv. hooggeschoolde mannen boven de 60 jaren) extra stimuleren (via een huisbezoek) om deel te nemen aan het onderzoek. De gegevens zijn ook nodig om inzicht te verwerven in de non-respons. Ten slotte kunnen bijkomende analyses zijn vereist in het licht van een beoogde publicatie van de resultaten in een academisch tijdschrift (academische valorisatie).

De contactgegevens zullen door de UGent gedurende 12 maanden bewaard worden. Deze bewaartermijn kan worden verantwoord gezien het feit dat de procedure van dataverzameling maximaal 12 maanden in beslag kan nemen. Deze procedure kent verschillende fases die zowel de respondent als de onderzoekers de nodige tijd en ruimte geven om de vragenlijsten in te (laten) vullen. De contactgegevens van respondenten die na het (her)ontvangen van de uitnodigingsbrief de UGent contacteren om aan te geven dat ze niet wensen deel te nemen en bijgevolg geen nieuwe uitnodiging en/of huisbezoek wensen, zullen onmiddellijk worden verwijderd bij ontvangst van een dergelijk bericht.

De socio-demografische gegevens zullen door de UGent gedurende 36 maanden bewaard worden. Deze bewaartermijn kan worden verantwoord gezien het feit dat het onderzoek academisch gevaloriseerd moet worden door middel van een publicatie in een academisch tijdschrift. Indien er toekomstige data-analyses gevraagd worden tijdens dit proces, zullen de socio-demografische gegevens inderdaad noodzakelijk zijn. De socio-demografische gegevens van respondenten die na het (her)ontvangen van de uitnodigingsbrief de UGent contacteren om aan te geven dat ze niet wensen deel te nemen en bijgevolg geen nieuwe uitnodiging en/of huisbezoek wensen, zullen, evenals hun pseudoniem worden bewaard. Aangezien hun contactgegevens wel zullen worden verwijderd zal wat hen betreft het ongedaan maken van de pseudonimisering (met het oog op het opnieuw uitnodigen en verrichten van huisbezoeken indien nodig) niet meer mogelijk zijn. De verdere bewaring van socio-demografische gegevens is noodzakelijk om de non-respons in kaart te kunnen brengen.

#### **Artikel 4: De categorieën van ontvangers en derden die mogelijks de gegevens eveneens verkrijgen**

De UGent zal de meegedeelde persoonsgegevens in het kader van de in artikel 2 vooropgestelde finaliteiten kunnen meedelen aan volgende categorie(ën) van ontvangers:

- EasyPost zal door de UGent als verwerker worden aangesteld voor het (her)adresseren aan en het (her)versturen per post van een persoonlijke uitnodiging naar de respondenten. Concreet zal de UGent (verwerkingsverantwoordelijke) zowel de brieven als de contactgegevens van de respondenten op elektronische wijze meedelen aan EasyPost (verwerker). Deze laatste kan vervolgens de brieven printen, in enveloppen steken, adresseren en ten slotte laten versturen. Daartoe zal een gegevensverwerkingsovereenkomst (Data Processing Agreement) worden gesloten tussen de UGent en EasyPost overeenkomstig Artikel 28 AVG;
- Volgende diensten van UGent zullen toegang hebben tot de gevraagde persoonsgegevens:
  - o Faculteit Politieke en Sociale wetenschappen, Vakgroep Communicatiewetenschappen, onderzoekers betrokken bij het onderzoeksproject NewsDNA (Prof. Dr. Lieven De Marez, Prof. Dr. Frederik De Grove, Kristin Van Damme, Glen Joris);
  - o Faculteit Letteren en Wijsbegeerte, Vakgroep Vertalen, Tolken en Communicatie, onderzoekers betrokken bij het onderzoeksproject NewsDNA (Prof. Dr. Veronique Hoste, dr. Orphée De Clerq, dr. Cynthia Van Héé);
  - o Faculteit Recht en Criminologie, Vakgroep Recht en Technologie, onderzoekers betrokken bij het onderzoeksproject NewsDNA (Prof. Dr. Eva Lievens, Judith Vermeulen).

Enkel personen die omwille van hun functieprofiel deze informatie nodig hebben voor de uitvoering van hun werk, krijgen toegang tot de informatie.

Elke eventuele mededeling van de gevraagde persoonsgegevens door de UGent moet voorafgaandelijk aan Stad Gent worden gemeld en moet uiteraard in overeenstemming zijn met de relevante wet- en regelgeving inzake de bescherming van natuurlijke personen bij de verwerking van persoonsgegevens. Dat betekent onder meer dat de UGent waar vereist een protocol sluit voor de mededeling van de gevraagde gegevens.

### **Artikel 5. Periodiciteit van de mededeling en de duur van de mededeling**

De persoonsgegevens zullen EENMALIG worden opgevraagd en medegedeeld. De verwerking zal, in overeenstemming met de voorgestelde bewaartermijn, maximaal 36 maanden rekenend vanaf de start van het onderzoek in beslag nemen.

### **Artikel 6: Beveiligingsmaatregelen**

Volgende maatregelen worden getroffen ter beveiliging van de mededeling van de persoonsgegevens, vermeld in artikel 2:

- De dataoverdracht gebeurt door middel van sFTP in CSV-formaat; een internetprotocol dat een beveiligde bestandsoverdracht tussen verschillende computers mogelijk maakt;
- De in dit protocol vernoemde onderzoekers betrokken bij het onderzoeksproject NewsDNA zijn contractueel verplicht de aan de UGent van kracht zijnde reglementen en voorschriften na te leven. Relevant in deze context is meer de specifiek de UGent generieke gedragscode voor de verwerking van persoonsgegevens en vertrouwelijke informatie. De UGent onderschrijft bovendien de European Code for Research Integrity of ALLEA-code.

De UGent treft ten minste volgende organisatorische en technische beveiligingsmaatregelen ter beveiliging van de ontvangen persoonsgegevens bij verdere verwerking:

- Bewaring van persoonsgegevens op UGent-servers die door de centrale diensten van de UGent professioneel worden beveiligd en geüpdatet aan de op dit moment huidige beveiligingsstandaarden;
- Gebruik van unieke codes of pseudoniemen om de identificatie van de respondent te beperken tot het minimale;
- Toegangsrestrictie van de datasets door middel van de Windows logingegevens van de betrokken UGent onderzoekers en een wachtwoord dat mondeling aan elkaar wordt overgedragen ter kennisname als dubbele beveiligingslaag;
- UGent verwijdert alle contactgegevens na verloop van 12 maanden en de socio-demografische gegevens (incl. statistische sectoren) na verloop van 36 maanden (zie ook artikel 3) en stuurt hiervan een bevestiging naar [privacy@stad.gent](mailto:privacy@stad.gent).

De UGent moet kunnen aantonen dat de in dit artikel opgesomde maatregelen werden getroffen. Op eenvoudig verzoek van Stad Gent moet de UGent hiervan aan Stad Gent het bewijs overmaken.

In het geval de UGent voor de verwerking van persoonsgegevens die het voorwerp zijn van voorliggend protocol, beroep doet op een verwerker (of meerdere verwerkers), doet de UGent uitsluitend beroep op verwerkers die afdoende garanties met betrekking tot het toepassen van passende technische en organisatorische maatregelen bieden opdat de verwerking aan de vereisten van de algemene verordening gegevensbescherming voldoet en de bescherming van de rechten van de betrokkene is gewaarborgd. De UGent sluit in voorkomend geval met alle verwerkers een verwerkersovereenkomst in overeenstemming met artikel 28 van de algemene verordening gegevensbescherming. Partijen bezorgen elkaar een overzicht van de verwerkers die de gevraagde gegevens verwerken, en actualiseren dit overzicht zo nodig.

### **Artikel 7: Kwaliteit van de persoonsgegevens**

Zodra de UGent één of meerdere foutieve, onnauwkeurige, onvolledige, ontbrekende, verouderde of overbodige gegevens in de persoonsgegevens, vermeld in artikel 3, vaststelt (al dan niet op basis van een mededeling van de betrokkene, meldt zij dat onmiddellijk aan Stad Gent die na onderzoek van de voornoemde vaststellingen de gepaste maatregelen treft zodra mogelijk en de UGent daarvan vervolgens op de hoogte brengt.

### **Artikel 8: Sanctie bij niet-naleving**

Onverminderd haar recht om een schadevergoeding te vorderen en in afwijking van artikel 5, 2°, kan Stad Gent dit protocol middels eenvoudige kennisgeving en zonder voorafgaandelijke ingebrekestelling eenzijdig beëindigen indien de UGent deze persoonsgegevens verwerkt in strijd met hetgeen bepaald is in dit protocol, met de algemene verordening gegevensbescherming of met andere relevante wet- of regelgeving inzake de bescherming van natuurlijke personen bij de verwerking van persoonsgegevens.

### **Artikel 9: Meldingsplichten**

De partijen engageren zich in het licht van artikel 33 van de algemene verordening gegevensbescherming om elkaar zonder onredelijke vertraging op de hoogte te stellen van elk gegevenslek dat zich voordoet betreffende de meegedeelde gegevens met impact op beide partijen en in voorkomend geval onmiddellijk gezamenlijk te overleggen teneinde alle maatregelen te nemen om de gevolgen van het gegevenslek te beperken en te herstellen. De partijen verschaffen elkaar alle informatie die ze nuttig of nodig achten om de beveiligingsmaatregelen te optimaliseren.

De UGent brengt Stad Gent onmiddellijk op de hoogte van wijzigingen van wetgeving met impact op voorliggend protocol, zoals de finaliteit, proportionaliteit, frequentie, duurtijd enz. en in voorkomend geval van wijzigingen omtrent de verwerkers.

### **Artikel 10: Toepasselijk recht en geschillenbeslechting**

Dit protocol wordt beheerst door het Belgisch recht.

Alle geschillen die voortvloeien uit of verband houden met dit protocol worden beslecht door de bevoegde rechtbank in het gerechtelijk arrondissement Oost-Vlaanderen, afdeling Gent.

### **Artikel 11: Inwerkingtreding en opzegging**

Dit protocol treedt in werking op \_\_\_\_\_.

Partijen kunnen dit protocol schriftelijk opzeggen mits inachtneming van een opzegtermijn van [X maanden].

Het protocol eindigt van rechtswege na afloop van de in artikel 5 van dit protocol bedoelde termijn.  
Het protocol eindigt tevens van rechtswege wanneer er geen rechtsgrond meer bestaat voor de gevraagde doorgifte van persoonsgegevens.

Opgemaakt te Gent, op 5 september 2019, in evenveel exemplaren als dat er partijen zijn.

Voor de Stad Gent,




Mieke Hullebroeck  
Algemeen Directeur



Mathias De Clercq  
Burgemeester

Voor de UGent,

~~NAMENS DE RECTOR~~

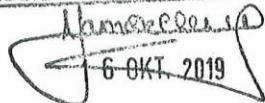


dr. Dirk De Craemer  
afdelingshoofd Onderzoekscóördinatie  
Universiteit Gent

Prof. dr. Rik Van de Walle

Rector Universiteit Gent

GEZIEN EN GOEDGEKEURD



6 OKT. 2019

YANNICK DE CLERCQ  
REGERINGSCOMMISSARIS

**Bijlage A: relatieve verdeling Gentse bevolking naar nationaliteit, geslacht, leeftijd en opleiding**

<b>Variabele</b>	<b>Categorieën</b>	<b>Relatief</b>
Nationaliteit	Belgisch	85,94%
	Buitenlandse nationaliteit	14,06%
Geslacht	Man	49,83%
	Vrouw	50,17%
Leeftijd	15 t/m 19 jaar	5,75%
	20 t/m 24 jaar	7,56%
	25 t/m 29 jaar	12,23%
	30 t/m 34 jaar	10,22%
	35 t/m 39 jaar	9,06%
	40 t/m 44 jaar	7,94%
	45 t/m 49 jaar	7,31%
	50 t/m 54 jaar	7,07%
	55 t/m 59 jaar	6,86%
	60 t/m 64 jaar	6,17%
	65 t/m 69 jaar	5,28%
	70 t/m 74 jaar	4,60%
	75+	9,95%
Opleiding	Lager of geen diploma	12,09%
	Lager secundair	10,31%
	Hoger secundair	41,48%
	Hoger niet-universitair, korte type / professioneel gerichte opleidingen	19,92%
	Hoger niet-universitair onderwijs lange type / Universitair onderwijs / Academische bachelor of master	16,20%

Opm.: aangezien de relatieve verdeling van opleidingsniveau niet terug te vinden is in de publieke databanken van de Stad Gent, is er voor de verdeling van het opleidingsniveau beroep gedaan op de verdeling die aanwezig is in de Belgische populatie.

### Appendix 3. Invitation letter



VAKGROEP COMMUNICATIEWETENSCHAPPEN  
ONDERZOEKSGROEP VOOR MEDIA, INNOVATIE  
EN COMMUNICATIETECHNOLOGIEËN

<voornaam> <achternaam>  
<straat> <huisnummer> <busnummer>  
<postcode> <gemeente>, België

Universiteit Gent, Campus Aula  
Korte Meer 11  
9000 GENT

DATUM  
7 november 2019

JOUW PERSOONLIJKE CODE  
<identifier>

#### Betreft: deelname nieuwsonderzoek Universiteit Gent

Beste <voornaam>,

De Universiteit Gent voert samen met de stad Gent een grootschalig onderzoek uit naar het nieuwsgebruik van de Gentenaar en heeft **jouw hulp** daarbij nodig!

#### Wat?

Door deel te nemen aan dit wetenschappelijk onderzoek hopen we in kaart te brengen of en hoe de Gentenaar nieuws raadpleegt en **hoe die in de toekomst nieuws zou willen ontvangen**. Deze vragen zijn belangrijk omdat ze ons inzicht kunnen geven in hoe we nieuwe technologieën voor nieuws moeten ontwikkelen. Ook indien je geen nieuws leest, zijn we geïnteresseerd in jouw mening.

#### Hoe deelnemen?

Surf naar [www.deelname.ugent.be](http://www.deelname.ugent.be) en vul jouw persoonlijke code in. Deze code vind je in het blauwe kader hierboven terug. Het invullen van de vragen duurt een tiental minuten. Geen internet? In elke bibliotheek van de Stad Gent kan je gratis op het web. We verloten **50 cadeaubonnen ter waarde van €10** onder de deelnemers die de vragenlijst volledig invullen. Deze cadeaubonnen kan je gebruiken bij één van de 152 deelnemende handelaars in Gent.

#### Hoe komen wij aan jouw naam en adres?

Door middel van een steekproeftrekking kregen wij toegang tot jouw gegevens. Indien je meer informatie wil over het onderzoek, surf dan naar [www.nieuwsonderzoek.ugent.be](http://www.nieuwsonderzoek.ugent.be). Op de achterzijde van deze brief kan je ook de privacy voorwaarden terugvinden die jou meer uitleg geven over hoe we jouw gegevens verwerken.

#### Nog vragen?

Aarzel dan zeker niet om contact op te nemen via [nieuwsonderzoek@ugent.be](mailto:nieuwsonderzoek@ugent.be).

Met vriendelijke groeten,

Prof. Frederik De Grove  
Coördinator nieuwsonderzoek UGent



Waarom krijg jij deze brief?  
Lees de voorwaarden op de keerzijde



### **Wij verkregen jouw gegevens van de Stad Gent**

Op basis van een representatieve steekproeftrekking uit het bevolkingsregister hebben wij jouw naam-, contact- en een aantal demografische gegevens verkregen van de Stad Gent. De Stad Gent heeft jouw gegevens oorspronkelijk verzameld voor het houden van een bevolkingsregister en baseert zich op artikel 5, lid 1, b) van de Algemene Verordening Gegevensbescherming ('AVG' of GDPR in het Engels), dat bepaalt dat de verdere verwerking met het oog op wetenschappelijk onderzoek niet als onverenigbaar met de oorspronkelijke doeleinden wordt beschouwd, om ons jouw gegevens te bezorgen. We zijn dus in het bezit van jouw persoonsgegevens, maar doen er alles aan om hier correct mee om te gaan.

### **Voor welke doeleinden gebruiken wij jouw gegevens?**

Jouw naam- en contactgegevens gebruiken wij om jou per post uit te nodigen om deel te nemen aan ons onderzoek. Concreet doen wij hiervoor beroep op drukkerij Manu-mail te Lebbeke, die zal optreden als verwerker. Ingeval van non-respons sturen wij jou mogelijks een tweede uitnodiging en/of komen wij bij jou langs om jouw deelname alsnog mogelijk te maken. Jouw demografische gegevens (leeftijd, geslacht en nationaliteit) hebben wij nodig om 1) de representativiteit van de uiteindelijke respondenten na te gaan, 2) een analyse te maken van de non-respons, 3) de huisbezoeken gericht te kunnen laten verlopen en 4) bijkomende analyses te verrichten met oog op academische valorisatie van het onderzoek.

### **Waarom mogen wij jouw gegevens verwerken?**

De gegevensverwerking gebeurt op grond van artikel 6, lid 1, e) AVG aangezien wetenschappelijk onderzoek verricht door de UGent kan worden beschouwd als "noodzakelijk voor de vervulling van een taak van algemeen belang". Dat blijkt onder meer ook uit het opzet van het onderzoeksproject #NewsDNA dat onderzoek doet naar hoe digitale media een invloed uitoefenen op bepaalde fundamentele democratische processen zoals het (correct) informeren van de bevolking. Verschillende beleids(ondersteunende) organen hebben recentelijk in hun beleidsrapportering de nood aan dergelijke onderzoeken en de meerwaarde daarvan voor de bevolking en de bredere samenleving reeds benadrukt.

### **Wie heeft toegang tot jouw gegevens?**

Indien het voor de verwerking en de doeleinden die werden opgesteld nodig blijkt, kunnen de volgende personen toegang krijgen tot de gevraagde persoonsgegevens: onderzoekers betrokken bij het onderzoeksproject #NewsDNA en aangesloten bij de onderzoeksgroep voor Media, Innovatie en CommunicatieTechnologieën (imec-mict-UGent), Language and Translation Technology Team (LT3-UGent), Wireless, Acoustics, Environment & Expert Systems (imec-waves-UGent) of Recht & Technologie (L&T-UGent).

### **We verwijderen jouw gegevens zo snel mogelijk!**

Jouw naam- en contactgegevens zullen door de UGent gedurende 12 maanden bewaard worden. Deze bewaartermijn kan worden verantwoord gezien het feit dat de procedure van dataverzameling maximaal 12 maanden in beslag kan nemen. Je kan ons echter contacteren om aan te geven dat je niet wenst deel te nemen en bijgevolg geen nieuwe uitnodiging en/of huisbezoek wenst, in welk geval wij jouw naam- en contactgegevens onmiddellijk zullen verwijderen. Jouw demografische gegevens zullen door de UGent gedurende 36 maanden bewaard worden. Deze bewaartermijn kan worden verantwoord gezien het feit dat het onderzoek academisch gevaloriseerd moet worden door middel van een publicatie in een academisch tijdschrift.

### **Jouw rechten**

Graag informeren wij jou ook over de rechten die je hebt ten aanzien van de door ons verwerkte persoonsgegevens. Je hebt met name:

- 1) **Recht op inzage:** je kan vragen of er persoonsgegevens over jou worden verwerkt en inzage verkrijgen van die persoonsgegevens evenals van de verwerkingsdoeleinden, de betrokken categorieën persoonsgegevens, de (categorieën) van ontvangers, de bewaartermijn, het bestaan van de in 2), 3), 4) en 6) opgesomde rechten en alle beschikbare informatie over de bron van jouw gegevens. In deze verklaring kan je deze informatie echter reeds grotendeels terugvinden;
- 2) **Recht op rectificatie:** indien jouw gegevens niet correct zijn, kan je ons vragen deze te corrigeren of aan te vullen;
- 3) **Recht op gegevenswissing of "vergetelheid":** in een aantal gevallen en omstandigheden kan je jouw persoonsgegevens (zonder onredelijke vertraging) laten wissen. Met betrekking tot de verwerking in de context van dit onderzoek kan je dit vragen, indien: a) jouw gegevens niet langer noodzakelijk zijn voor de doeleinden waarvoor wij ze hebben verzameld; b) wij jouw gegevens op onrechtmatige wijze zouden gebruiken. Op dit recht gelden in het kader van wetenschappelijk onderzoek uitzonderingen voor zover de uitoefening van dit recht de verwezenlijking van de doeleinden van die verwerking onmogelijk dreigt te maken of ernstig in het gedrang dreigt te brengen;
- 4) **Recht op beperking van de verwerking:** indien aan bepaalde criteria is voldaan, kan je ons vragen (tijdelijk) te stoppen met de verwerking (met uitzondering van het opslaan) van jouw persoonsgegevens. Dit recht bestaat meer bepaald: a) gedurende de periode die wij nodig hebben om de juistheid van de door ons verwerkte gegevens te controleren, indien je die betwist; b) indien jouw gegevens onrechtmatig verwerkt zouden zijn, maar je zich verzet tegen de wissing zoals bedoeld in 3); c) wanneer wij jouw gegevens niet meer nodig hebben, maar je daarentegen wel met oog op de instelling, uitoefening of onderbouwing van een rechtsvordering;
- 5) **Recht om gegevens over te dragen:** je hebt het recht om jouw persoonsgegevens over te dragen in een gemakkelijk leesbaar en hanteerbaar formaat aan een andere verwerkingsverantwoordelijke;
- 6) **Recht om klacht in te dienen:** je hebt het recht om klacht in te dienen bij de Vlaamse Toezichtcommissie voor de verwerking van persoonsgegevens indien je van mening bent dat de verwerking door ons van jouw persoonlijke informatie inbreuk maakt op de AVG.

### **Nog vragen?**

Indien je nog vragen hebt of uw rechten, zoals hierboven uiteengezet, wenst uit te oefenen, kan je een brief of mail sturen naar de coördinator van de nieuwsonderzoek: Prof. Frederik De Grove, Korte Meer 11, 9000 Gent, België, [nieuwsonderzoek@ugent.be](mailto:nieuwsonderzoek@ugent.be). De coördinator voor gegevensbescherming van de UGent is bereikbaar via [privacy@ugent.be](mailto:privacy@ugent.be).

\*De Vlaamse Toezichtcommissie voor de verwerking van persoonsgegevens: Koning Albert II-laan 15, 1210 Brussel, T +32 (0)2 553 50 47, [contact@toezichtcommissie.be](mailto:contact@toezichtcommissie.be).

#### Appendix 4. Results home visits (absolute statistics)

Result	Number of people
Participated	17
Not at home	74
Refused	61
Total	152

**Appendix 5. Final response sample (absolute and relative statistics)**

	Foreign nationality			Domestic nationality			Total			
	M	F	Total	M	F	Total	M	F	Unknown	Total
18-24	7(0.74%)	7(0.74%)	14(1.48%)	22(2.33%)	41(4.35%)	63(6.68%)	29(3.08%)	48(5.09%)	0(0.00%)	77(8.17%)
25-34	14(1.48%)	15(1.59%)	29(3.08%)	64(6.79%)	80(8.48%)	144(15.27%)	78(8.27%)	95(10.07%)	0(0.00%)	173(18.35%)
35-44	7(0.74%)	9(0.95%)	16(1.70%)	57(6.04%)	55(5.83%)	112(11.88%)	64(6.79%)	64(6.79%)	1(0.11%)	129(13.68%)
45-54	9(0.95%)	6(0.64%)	15(1.59%)	56(5.94%)	59(6.26%)	115(12.20%)	65(6.89%)	65(6.89%)	0(0.00%)	130(13.79%)
55-64	5(0.53%)	4(0.42%)	9(0.95%)	90(9.54%)	92(9.76%)	182(19.30%)	95(10.07%)	96(10.18%)	0(0.00%)	191(20.25%)
65-74	0(0.00%)	5(0.53%)	5(0.53%)	84(8.91%)	86(9.12%)	170(18.03%)	84(8.91%)	91(9.65%)	0(0.00%)	175(18.56%)
75-84	2(0.21%)	0(0.00%)	2(0.21%)	21(2.23%)	29(3.08%)	50(5.30%)	23(2.44%)	29(3.08%)	1(0.11%)	53(5.62%)
85+	0(0.00%)	0(0.00%)	0(0.00%)	9(0.95%)	6(0.64%)	15(1.59%)	9(0.95%)	6(0.64%)	0(0.00%)	15(1.59%)
<b>Total</b>	<b>44(4.67%)</b>	<b>46(4.88%)</b>	<b>90(9.54%)</b>	<b>403(42.74%)</b>	<b>448(47.51)</b>	<b>851(90.24%)</b>	<b>447(47.40%)</b>	<b>494(52.39%)</b>	<b>2(0.21%)</b>	<b>943(100.00%)</b>

## Appendix 6. Descriptive statistics

Variabele	Question text	Mean/ mode	Standard deviation	Variance	Range	N
News selection mechanisms	News based on my interests	4.00	0.89	0.80	1-5	943
	News based on what I say I like	3.75	1.02	1.04	1-5	943
	News based on what I read earlier	3.32	1.02	1.04	1-5	943
	News based on what persons with similar interests read	3.33	1.04	1.07	1-5	943
	News based on what is currently popular	3.08	1.04	1.09	1-5	943
	News based on what my friends have read	2.86	1.02	1.03	1-5	943
	News based on opinions that differ from my own	3.33	0.94	0.88	1-5	943
	News based on the viewpoints I disagree with	3.13	0.97	0.94	1-5	943
	News based on topics I do not usually read	2.77	1.00	0.99	1-5	943
	News based on what is outside my field of interests	2.64	0.98	0.95	1-5	943
News information overload	I often felt overwhelmed about the large amount of daily news	2.96	1.10	1.21	1-5	943
	I give up following the news due to the large amount of news	2.38	1.04	1.07	1-5	943
	I often felt that there was more news than I could process	3.32	1.13	1.28	1-5	943
	I often doubt whether I do not miss out the most important news of the day due to the large amount of news	2.80	1.03	1.07	1-5	943
	I often do not know where to start due to the large amount of news	2.82	1.04	1.09	1-5	943
	I often felt stressed about the large speed of news coverage	2.65	1.07	1.14	1-5	943
Vertrouwen	Imbalanced - Balanced	3.30	0.92	0.85	1-5	943
	Subjective - Objective	3.16	0.95	0.91	1-5	943
	Not report the whole story - Report the whole story	3.03	0.99	0.99	1-5	943
	Inaccurate - Accurate	3.31	0.93	0.86	1-5	943
	Not honest - Honest	3.49	0.85	0.73	1-5	943
	Not believable - Believable	3.63	0.84	0.71	1-5	943
	Not up-to-date - Up-to-date	3.88	0.89	0.78	1-5	943
	Not current - Current	3.85	0.86	0.75	1-5	943
	Not timely - Timely	3.83	0.91	0.83	1-5	943
Technology optimism	New technologies contribute to a better quality of life	3.43	0.95	0.90	1-5	943
	Technology gives me more freedom of mobility	3.47	0.98	0.96	1-5	943
	Technology gives people more control over their daily lives	3.19	0.96	0.91	1-5	943
	Technology makes me more productive in my personal life	3.19	1.02	1.04	1-5	943
Concerns about news personalization	I worry that more personalized news may mean that my privacy is placed at greater risk	3.33	1.07	1.14	1-5	943
	I worry that more personalized news may mean that I miss out on important information	3.62	1.01	1.02	1-5	943
	I worry that more personalized news may mean that I miss out on challenging viewpoints	3.70	0.99	0.98	1-5	943

Note: 1. all questions and variables were in the questionnaire clarified with an introduction text and several examples. To keep this table clear, we removed these texts. The complete questionnaire is available from the corresponding author, upon reasonable request. 2. Items marked with an asterisk were characterized with little variation or skewed means. These items were removed from further analysis.

## Appendix 6. Descriptive statistics (continued)

Variabele	Question text	Mean/mode	Standard deviation	Variance	Range	N
News interest	Societal issues such as news about environment or economy	7.70	1.90	3.61	0-10	943
	Moviestars, royalties and celebrities	3.81	2.63	6.92	0-10	943
	Politics and politicians	6.29	2.70	7.28	0-10	943
	Sport	4.21	3.22	10.39	0-10	943
	Human interest	5.47	2.59	6.70	0-10	943
	Crime, accidents and disasters	5.75	2.57	6.59	0-10	943
	Music, film and other cultural activities	5.98	2.52	6.34	0-10	943
	Fashion, food, nightlife and lifestyle news	4.76	2.68	7.19	0-10	943
News use	How many minutes did you yesterday spend to the consumption of news?	64.21	46.15	6.83	n/a	943
Age	What is your age?	51	17.85	318.78	n/a	943
News use	National television	"Daily"	n/a	n/a	1-4	943
	Regional television	"Never"	n/a	n/a	1-4	943
	Radio news	"Daily"	n/a	n/a	1-4	943
	Newspaper	"Daily"	n/a	n/a	1-4	943
	News website/application of traditional news organizations	"Never"	n/a	n/a	1-4	943
	Alternative news websites*	"Never"	n/a	n/a	1-4	943
	Digital news letter	"Never"	n/a	n/a	1-4	943
	News on social media	"Never"	n/a	n/a	1-4	943
	News via Google News or other news aggregators*	"Never"	n/a	n/a	1-4	943
	News via search machines	"Never"	n/a	n/a	1-4	943
Use of foreign news sources	Did you last year used a foreign news source?	"Yes"	n/a	n/a	n/a	943
Paid news	Did you last year paid once to read news articles online?	"No"	n/a	n/a	n/a	943
Smartphone use	What is your main way of accessing online news?	"Smartphone"	n/a	n/a	n/a	943
Online news	Do you read news online?	"Yes"	n/a	n/a	n/a	943
Online news participation	Did you liked, shared or commented upon a news article last week?	"Yes "	n/a	n/a	n/a	943
Gender	What gender is on your identity card?	"Female"	n/a	n/a	n/a	942
Education	What is your highest level of education?	"Master "	n/a	n/a	n/a	942
Employment	What is your current occupation status?	"Full time"	n/a	n/a	n/a	940
Nationality	What is your nationality?	"Domestic nationality"	n/a	n/a	n/a	942
Nationality parents	Wat is the nationality of your parents?	"Domestic nationality"	n/a	n/a	n/a	943

Note: 1. all questions and variables were in the questionnaire clarified with an introduction text and several examples. To keep this table clear, we removed these texts. The complete questionnaire is available from the corresponding author, upon reasonable request. 2. Items marked with an asterisk were characterized with little variation. These items were removed from further analysis.

**Appendix 7. Mean difference scores of audiences' evaluations of different news selection mechanisms**

	1 - News based on my interests	2 - News based on what I say I like	3 - News based on what I read earlier	4 - News based on what persons with similar interests read	5 - News based on what is currently popular	6 - News based on what my friends have read	7 - News based on opinions that differ from my own	8 - News based on the viewpoints I disagree with	9 - News based on topics I do not usually read	10 - News based on what is outside my field of interests
1 - News based on my interests	-	-0.255***	-0.685***	-0.670***	-0.928***	-1.140***	-0.673***	-0.878***	-1.232***	-1.363***
2 - News based on what I say I like	-	-	-0.431***	-0.416***	-0.673***	-0.885***	-0.419***	-0.624***	-0.978***	-1.108***
3 - News based on what I read earlier	-	-	-	0.015	-0.243***	-0.455***	0.012	-0.193***	-0.547***	-0.678***
4 - News based on what persons with similar interests read	-	-	-	-	-0.258***	-0.470***	-0.003	-0.208***	-0.562***	-0.692***
5 - News based on what is currently popular	-	-	-	-	-	-0.212***	0.467***	0.050	-0.304***	-0.435***
6 - News based on what my friends have read	-	-	-	-	-	-	0.255***	0.262***	-0.092*	-0.223***
7 - News based on opinions that differ from my own	-	-	-	-	-	-	-	-0.205***	-0.559***	-0.689***
8 - News based on the viewpoints I disagree with	-	-	-	-	-	-	-	-	-0.354***	-0.485***
9 - News based on topics I do not usually read	-	-	-	-	-	-	-	-	-	-0.130***
10 - News based on what is outside my field of interests	-	-	-	-	-	-	-	-	-	-

Note: mean difference is significant when p-value < .05 (\*). To counter the problem of familywise error rates, the following corrections are calculated for significance: Holm correction: p-value < .01138 (\*\*); Bonferroni correction: p-value < .001111 (\*\*\*)

## Appendix 8. Construct reliability and validity

Latent variable	Observed variable	Cronbach's alpha	Factor loadings	Item variance	Squared multiple correlation
Evaluation: content-based similarity	News based on what I read earlier	0.731	0.515*	0.265	0.507
	News based on my interests		0.758*	0.575	
	News based on what I say I like		0.825*	0.680	
Evaluation: collaborative similarity	News based on what persons with similar interests read	0.688	0.654	0.427	0.435
	News based on what my friends have read		0.756	0.571	
	News based on what is currently popular		0.553	0.306	
Evaluation: content-based diversity	News based on the viewpoints I disagree with	0.778	0.715	0.511	0.468
	News based on opinions that differ from my own		0.737	0.544	
	News based on topics I do not usually read		0.652	0.426	
	News based on what is outside my field of interests		0.626	0.392	
News information overload	I often felt overwhelmed about the large amount of daily news	0.837	0.792	0.627	0.469
	I give up following the news due to the large amount of news		0.593	0.352	
	I often felt that there was more news than I could process		0.639	0.409	
	I often doubt whether I do not miss out the most important news of the day due to the large amount of news		0.628	0.394	
	I often do not know where to start due to the large amount of news		0.785	0.617	
	I often felt stressed about the large speed of news coverage		0.645	0.416	
Technology optimism	New technologies contribute to a better quality of life	0.808	0.677	0.458	0.515
	Technology gives me more freedom of mobility		0.693	0.481	
	Technology gives people more control over their daily lives		0.766	0.586	
	Technology makes me more productive in my personal life		0.731	0.534	
Trust	Imbalanced - Balanced	0.890	0.732	0.535	0.449
	Subjective - Objective		0.655	0.428	
	Not report the whole story - Report the whole story		0.685	0.469	
	Inaccurate - Accurate		0.737	0.543	
	Not honest - Honest		0.812	0.659	
	Not believable - Believable		0.784	0.614	
	Not up-to-date - Up-to-date		0.653	0.426	
	Not current - Current		0.596	0.355	
	Not timely - Timely		0.557	0.311	

Note: Thresholds used to consider removing items (DeVellis, 2016): Cronbach's alpha: > .7; Factor loadings (CFA): > .5; Item variance (R-square): > .4; Squared multiple correlation: > .4; based on these thresholds, no items were removed.

**Appendix 9.** Model fit information for LC-clusters ‘evaluation of different news selection mechanisms’ with 1-6 classes

#-cluster	LL	BIC(LL)	AIC(LL)	AIC3(LL)	Npar	L <sup>2</sup>	df	p-value	Class.Err.
1-Cluster	-12481.75	25237.46	25043.49	25083.49	40.00	12668.55	903.00	.00	.00
2-Cluster	-11752.83	23854.96	23607.66	23658.66	51.00	11210.72	892.00	.00	.01
3-Cluster	-11568.80	23562.25	23261.61	23323.61	62.00	10842.66	881.00	.00	.10
4-Cluster	-11401.45	23302.89	22948.90	23021.90	73.00	10507.96	870.00	.00	.15
5-Cluster	-11257.57	23090.47	22683.15	22767.15	84.00	10220.20	859.00	.00	.15
6-Cluster	-11176.06	23002.78	22542.12	22637.12	95.00	10057.17	848.00	.00	.14

Note: when comparing models: the lower the BIC and AIC, the better the model fits with the data. Based on the different LC-models and their model fit information, we chose for the 4-cluster model which has a p-value < .05 and the best fit in relation to the number of clusters to use (i.e., elbow-method).

**Appendix 10.** Model fit information for LC-clusters ‘news interest’ with 1-6 classes

	LL	BIC(LL)	AIC(LL)	AIC3(LL)	Npar	L <sup>2</sup>	df	p-value	Class.Err.
1-Cluster	-16757.32	34062.56	33674.64	33754.64	80	20608.39	863	.00	.00
2-Cluster	-16377.30	33364.17	32932.60	33021.60	89	19848.36	854	.00	.09
3-Cluster	-16253.73	33178.66	32703.45	32801.45	98	19601.20	845	.00	.12
4-Cluster	-16138.87	33010.60	32491.75	32598.75	107	19371.50	836	.00	.13
5-Cluster	-16102.09	32998.67	32436.18	32552.18	116	19297.73	827	.00	.17
6-Cluster	-16054.42	32964.97	32358.84	32483.84	125	19202.59	818	.00	.16

Note: when comparing models: the lower the BIC and AIC, the better the model fits with the data. Based on the different LC-models and their model fit information, we chose for the 4-cluster model which has a p-value < .05 and the best fit in relation to the number of clusters to use (i.e., elbow-method).

**Appendix 11.** Model fit information for LC-clusters ‘news use’ with 1-6 classes

	LL	BIC(LL)	AIC(LL)	AIC3(LL)	Npar	L <sup>2</sup>	df	p-value	Class.Err.
1-Cluster	-8827.14	17818.66	17702.28	17726.28	24	6000.38	919	.00	.00
2-Cluster	-8239.38	16704.78	16544.76	16577.76	33	4824.86	910	.00	.01
3-Cluster	-8087.27	16462.21	16258.55	16300.55	42	4520.64	901	.00	.08
4-Cluster	-7983.43	16316.16	16068.86	16119.86	51	4312.96	892	.00	.10
5-Cluster	-7959.41	16329.76	16038.81	16098.81	60	4264.91	883	.00	.15
6-Cluster	-7927.55	16327.69	15993.10	16062.10	69	4201.20	874	.00	.15

Note: when comparing model fit information: the lower the BIC and AIC, the better the model fits with the data. Based on the different LC-models and their model fit information, we chose for the 4-cluster model which has a p-value < .05 and the best model fit in relation to the number of clusters to use (i.e., elbow-method).



## Appendix 12. Mean scores of the 4-cluster model: news interest

News interest topics	Cluster 1: all round interested (43.38%)	Cluster 2: political interested (32.25%)	Cluster 3: moderate interested (18.93%)	Cluster 4: low interested (5.44%)
Societal issues	8.16	8.58	5.71	5.73
Movie stars. royalties and celebrities	5.68	1.97	3.68	0.33
Politics and politicians	6.70	7.78	3.65	3.36
Sport	5.03	3.67	4.25	0.64
Human interest	6.89	4.60	4.79	1.71
Crime. accidents and disasters	7.04	4.79	5.42	2.35
Music. film and other cultural activities	7.07	5.81	4.78	2.41
Fashion, food, nightlife and lifestyle news	6.60	3.31	4.19	0.54

Note: In this model, we identified four types of news interests. The first type is the “all round interested” (43.38%) who is highly interested in all news topics available, including politics, sports and celebrities. The second type is the “political interested” (32.25%) who is particularly interested in societal topics such as politics, economy or health. The third type is the “moderate interested” (18.93%) who is not very interested in news in general, but mainly likes to follow news tragedies such as accidents, protests and disasters. The fourth type is “low interested” (5.44%) who is generally not interested in news.

## Appendix 13. Mean scores of the 4-cluster model: news use

News use channels	Cluster 1: News omnivores (36.30%)	Cluster 2: Moderate news omnivores (29.97%)	Cluster 3: Traditionalists (20.30%)	Cluster 4: Moderate traditionalists (13.43%)
National television	3.83	2.32	3.98	2.56
Regional television	2.45	1.19	2.90	1.36
Radio	3.63	2.86	3.47	3.00
Newspaper	2.79	2.02	3.25	2.12
News website	3.39	3.25	1.01	1.01
Digital news letter	2.48	1.72	1.00	1.00
News on social media	2.20	2.72	1.00	1.01
News via search engines	2.36	2.20	1.01	1.02

Note: In this model, we identified four types of news consumption. The first type is the “news omnivore” (36.30%) who extensively use all available news channel, with television, radio and social media as most favored. A second type is the “moderate news omnivore” (29.97%) who also use various news channels, but less frequently and mainly not news dominant channels such as radio or social media. The third type is the “traditionalist” (20.30%) who consume news on a daily basis via non-digital news channels such as television and radio. The fourth type is the “moderate traditionalist” (13.43%) who occasionally consume news via traditional news channel such as radio and television.

## List of tables, figures and appendices

**Table 1.** Mean scores of the 4-cluster model: audiences' evaluation of different news selection mechanisms

**Table 2.** Linear and hierarchical regression analysis: standardized regression coefficients

**Figure 2.** Stacked bar chart to measure the audiences' evaluation of different news selection mechanisms: content-based similarity (items 1-3), collaborative similarity (items 4-6) and content-based diversity (items 7-10).

**Appendix 1.** Sample characteristics (absolute and relative statistics)

**Appendix 2.** Study protocol

**Appendix 3.** Invitation letter

**Appendix 4.** Results home visits (absolute statistics)

**Appendix 5.** Final response sample (absolute and relative statistics)

**Appendix 6.** Descriptive statistics

**Appendix 7.** Mean difference scores of audiences' evaluations of different news selection mechanisms

**Appendix 8.** Construct reliability and validity

**Appendix 9.** Model fit information for LC-clusters 'evaluation of different news selection mechanisms' with 1-6 classes

**Appendix 10.** Model fit information for LC-clusters 'news interest' with 1-6 classes

**Appendix 11.** Model fit information for LC-clusters 'news use' with 1-6 classes

**Appendix 12.** Mean scores of the 4-cluster model: news interest

**Appendix 13.** Mean scores of the 4-cluster model: news use