2023 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)

Discovering ideological structures in representation learning spaces in recommender systems on social media data

Tim Faverjon
Médialab (Sciences Po)
Learning Planet Institute (Université Paris Cité)
tim.faverjon@sciencespo.fr

Pedro Ramaciotti
Complex Systems Institute (Paris Ile-de-France CNRS)
Sciences Po médialab & LPI (Université Paris Cité)

***Abstract*—Recommender systems in social platforms attract attention in part because of their potential impact over political phenomena, such as polarization or fragmentation of online communities. These research topics are also important because of the need for understanding systemic effects in view of upcoming risk-oriented AI regulation in the EU and the US. A common approach leverages outcomes of recommendations to audit recommender systems. A different approach is that of explainability, seeking to render recommendation mechanisms intelligible to humans, potentially enabling both auditing and actionable design tools. This second approach is particularly challenging in the context of online systems of political opinions because of the intrinsic unobservability of opinions. In this article we leverage multi-dimensional political opinion estimation of large online populations (along a left-right dimension but also along other political dimensions) to investigate latent spaces in representation learning computed by recommender systems. We train a recommender based on ubiquitous collaborative filtering principles using data on content sharing on Twitter by a large population, evaluating accuracy and extracting a latent space representation leveraged by the recommender. On the other hand, we leverage multi-dimensional political opinion inference to position users in political spaces representing their opinions. We then show for the first time the relation between latent representations leveraged by a recommender system and the spatial representation of users. We show that some dimensions learned by the recommender capture ideological positions of users, bridging politics and algorithmics in our social and algorithmic system, opening a path towards political explainability of AI.**

***Index Terms*—Recommender systems, political ideology, algorithm audit, algorithm explainability**

ASONAM '23, November 6-9, 2023, Kusadasi, Turkey

ACM ISBN 979-8-4007-0409-3/23/11…$15.00
https://doi.org/10.1145/3625007.3627336

I. Introduction

Social platforms play an important role in political debate in several countries, in information dissemination and consumption, and in shaping individual perceptions and political opinions [1]. Social platforms have also multiplied the sources of information available, making recommender and information retrieval systems indispensable to navigate the large and growing set of contents [2], [3]. Recommender systems in social media have also become central to platforms in their efforts to deliver satisfactory user experiences, and in driving metrics related to their underlying advertising business model [4], [5]. These developments, together with the scope of the massification of social media, have raised concerns about the social and political impacts of recommender systems [6]–[8]. These concerns are diverse, including for example the spread of misinformation, polarization and fragmentation of online communities, or mass exposure to harmful algorithmic biases [9].

To address these concerns, algorithm auditing and algorithm explainability methods have emerged as two relevant tools for evaluating, but also for understanding how algorithms might be impacting online social systems. Algorithm auditing involves the systematic assessment of outcomes of recommendations: *i.e.*, what was recommended to whom. Characterizing users and items, auditing involves measuring properties of recommendations such as biases [10], diversity [11], novelty [12], or their ability to contribute to polarized states in these large online social systems [13]. Algorithm explainability, on the other hand, focuses on producing human-understandable explanations of processes deciding a recommendation, allowing an examiner to understand the steps and the factors that explain that an item is recommended over another to a given user. By understanding the underlying mechanisms, undesired outcomes can be identified and potentially avoided allowing for platform designers to make informed decisions, address concerns, and enhance the transparency and fairness of content recommendations.

Several application cases have recently seen important advances regarding explainability (see [14] for a survey on the state of the different techniques for explainability, and [15] specifically for explainability of recommender systems). However, algorithm explainability in the context of political opinions is especially challenging because of their inherent unobservability. Algorithm explainability in the context of political opinions would allow for addressing questions, *e.g.*, such as: To what degree is item i being recommended to user u because of the political preferences p of user u? (even if political preferences p were not used during training). While several socio-demographic attributes are often inferred or volunteered by users in platforms, political opinions are hard to determine. More importantly, political opinions are harder to conceptualize, to model, and to measure, let alone to infer. While the US setting lend itself naturally to binary classifications (*e.g.*, Democrat- or Republican-leaning) or one-dimensional continuous ideological scales ranging from liberals to conservatives, it is known that, in general, different national settings require additional ideological and issue dimensions to explain political opinions [16] and observed online behavior [17]. In national settings with comparatively less issue alignment, more dimensions (such as dimensions modeling attitudes immigration, religious principles, or income redistribution) might be needed, and might even be more important than traditional left-right political dimensions [18].

In this article we use newly developed methods for estimating multi-dimensional political opinions of online populations. Doing so, we bridge for the first time 1) the configuration of political opinions of users en the one hand, with 2) the learned representations leveraged by recommender systems on the other, allowing for algorithmic political explainability: *i.e.*, how do political opinions of users determine what is recommended to them? We consider a large Twitter population of nearly 360k users for which we obtain an embedding in a two-dimensional political space spanned by the two most significant dimensions for the population, as computed in [19]: a Left-Right dimension, and a dimensions measuring attitudes towards elites and institutions. Next, we collect URLs pointing to web content shared by users of this population on tweets posted by them, and we train and validate a widely used recommendation principle: Non-Negative Matrix Factorization (NMF). We retrieve the learned representation space of the trained NMF recommender and compare the positions of users in this latent space to their political positions in the two-dimensional political space. We find that these political dimensions are recovered by the representation space of the recommender system. We show how to identify the relation between these two spaces and how to measure it, and prove that one of the dimensions of the recommender captures, for example, left-leaning ideologies, while another one captures right-leaning ideologies.

These results provide an actionable path towards treating concerns related to recommender systems in social settings. Algorithm political explainability (*i.e.*, knowing which political attitude or ideology played what role in producing a given recommendation) holds the promise of both, auditing what algorithms do, but also opens relevant ways in which algorithms could be modified. Once dimensions relating to particular political attitudes are identified in the representation space of a recommender system, the possibility exists for implementing constrained learning methods to manage what algorithms can leverage in terms of political preferences of users. Future methods leveraging this principle offer new ways for platforms to moderate undesired outcomes when needed, particularly in the face of upcoming regulation regarding systemic consequences of AI.

II. Data

To propose our method for analyzing the degree to which learned representations computed by a recommender system relate to political dimensions underlying the training data, we will use two datasets. The first dataset comes from previous work [19] in which a large population of Twitter users were positioned in a multi-dimensional political opinion space based on how they followed political figures. The second dataset contains URLs of web content shared by Twitter users on the platform, and will be used to train and test a recommender system.

A. Multi-mimensional political estimates for Twitter population

To obtain a large Twitter population with known multi-dimensional political estimates we leverage the dataset produced in [19]. This dataset contains a list of 368.831 Twitter users in France with positions in a geometrical political space, in which dimensions capture political positions towards different issues or ideologies computed with data collected in 2019. We retain the two most important political dimensions, described as those that have the most explicative power for describing the social graph (*i.e.*, follower network): 1) a Left-Right dimensions, and 2) an Anti-elite salience dimension measuring mistrust in elites and institutions. The reader is referred to the acknowledgment section (acknowledgement section hidden during double-blind review) for learning how we obtained this dataset and for the legal deposit and ethics board approval, and to [19] for more details and for benchmarks validating these ideological and political positions. These two political dimensions are endowed by construction with reference points. For the Left-Right dimension, position 0 marks the leftmost position for political parties, 5 marks the political center, and 10 marks the rightmost position for political parties. Individual users can be to the left position 0 (*i.e.*, be more leftist than the leftmost political party) or to the right of position 10 (*i.e.*, be more rightist than the rightmost political party). For the Anti-elite salience dimension, position 0 indicates granting no importance to anti-establishment and anti-elite rhetoric, while 10 indicates granting great importance to this. These reference points are also conceived for parties, with individual users being allowed positions outside the 0 to 10 segment. Spatial references are calibrated with respect to party positions for the purpose or readability and calibration of the political embedding [19, Section 5] (see [20] for a description of the political survey used for calibration). Fig. 1 show the spatial distribution of the 368.831 users of the acquired dataset along the two political dimensions. The distribution shows a clear bias of the sample, towards right-wing ideologies and high negative attitudes towards elites and institutions, which will be taken into account in our analysis.

B. Media sharing data

To train our recommender system we use user-item information available from Twitter. We randomly select 40.000 of the 368.831 users in the first dataset and collect their last 3.200 tweets leading to the time of collection of the first dataset. The accounts of 50 users were no longer available for collection. The need for selecting a random subsample stems

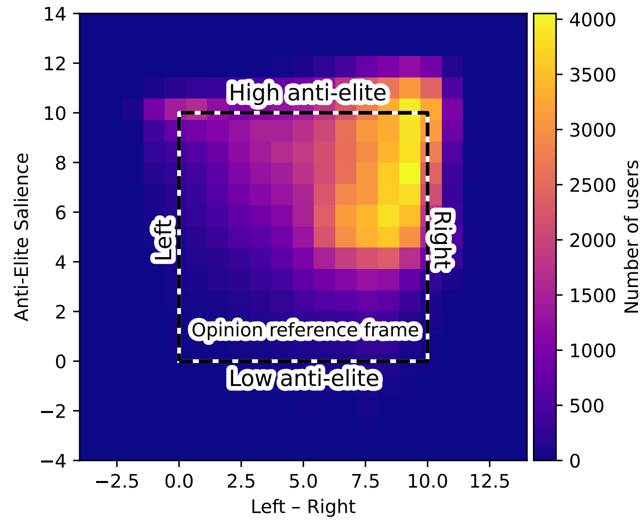

Fig. 1. Spatial distribution of the 368.831 Twitter users of the acquired dataset along the two political dimensions that are most explicative of political follower network in the data: Left-Right and Anti-elite salience. Reference positions of 0 and 10, available with the obtained data, are displayed as a $[0, 10]^2$ region.

from the cost of collection of tweets for all users in the first dataset. We keep only tweets that contain an URL pointing to a webpage, resulting in 23.534.803 tweets with URLs from 426.014 unique web domains (including, *e.g.*, main media outlets in France, e-commerce sites, and known blogs and forum sites). We will use these web domains as a set of items $\mathcal{I}$ chosen by users $\mathcal{U}$ (with known political coordinates) to train and validate a web domain recommender. We further filter this dataset by deleting tweets containing URLs pointing to social platforms such as twitter.com, facebook.com, linkedin.com, youtube.com, or instagram.com, after having disambiguated shortened URLs. We argue that these domains do not provide useful signals, as a link towards Twitter will point to content of a specific user or media outlet, which our data does not identify. Additionally, we remove from this dataset domains that have been shared less than 10 times so as to ensure enough implicit expressions of preference per item. This processing results in a dataset of 29.373 users sharing URLs from 32.639 items (*i.e.*, the domains) on 3.277.738 posts (*i.e.*, tweets).

III. Training a Media Recommender System on Twitter Data

Formally, we consider that a user $u \in \mathcal{U}$ can choose each item $i \in \mathcal{I}$ a number of times r_{ui}. We cast our recommendation problem as one of predicting future choices based on past ones, *i.e.*, $\mathcal{K} \subset (\mathcal{U} \times \mathcal{I} \times \mathbb{R}^{+})$. We consider web domains to be the items to be recommended to users. An ubiquitous recommendation principle – collaborative filtering – relies on collective choices to predict future behavior, by creating a representation of past choices on which to evaluate and leverage similarities between users and items. We take a generalist recommender, the NMF algorithm [21] focused on collaborative filtering – to the exclusion of content filtering – that relies on spatial representations of users and items. This choice presents the additional interest of generalizing our explainability method potentially to many other recommender systems [22].

Let $R \in \mathbb{R}_{+}^{n \times m}$, with $m = |\mathcal{I}|$ and $n = |\mathcal{U}|$, be our observation matrix, where, for all $(u, i, r_{ui}) \in \mathcal{K}$ and $R_{ui} = r_{ui}$. Let $k \in \mathbb{N}_{+}$ be the dimension of the learned representation space ($k < min(n, m)$). The NMF recommendation problem is prescribed as finding matrices $P \in \mathbb{R}_{+}^{n \times k}$ and $Q \in \mathbb{R}_{+}^{m \times k}$ of positive values, that minimize $\|R - PQ^{T}\|_{2}$. We employ the following regularized implementation [23]: find $P \in \mathbb{R}_{+}^{n \times k}$ and $Q \in \mathbb{R}_{+}^{m \times k}$ such that they minimize the functional expression :

$$
\begin{aligned}
ll\mathcal{L}(P,Q) = \quad & \tfrac{1}{2}||R - PQ^{T}||_{2}^{2} \\
& +\alpha_{P} l_{1} k ||P||_{1} + \alpha_{Q} l_{1} k ||Q||_{1} \\
& +\tfrac{1}{2}\alpha_{P}(1-l_{1})k||P||_{2}^{2} + \tfrac{1}{2}\alpha_{Q}(1-l_{1})k||Q||_{2}^{2},
\end{aligned}
$$

where $\alpha_P,\ \alpha_Q,\ l_1$ are parameters respectively controlling regularization on P and Q. We further modified the problem by log-scaling the observed data in R, considering instead a modified observation matrix $\tilde{R}$ with values $\tilde{r}_{ui} = 1 + 0.98 \cdot log(1 + 2r_{ui})$, to account for the long-tailed distribution of number shares per domain (value 0.98 allows for our empirical observations $\tilde{r}_{ui}$ to be between 1 and 10). Abusing notation we will retain the notation of log-scaled values as r_{ui}. We solve the optimization problem with the Multiplicative Update solver [24], a gradient descent with adaptive learning rate that forces the non-negativity of the solution. We initialize our algorithm with a Non-negative Double Singular Value Decomposition method as proposed by [25], to improve time of convergence and accuracy. We then optimize the hyper-parameters with a Particle Swarm Optimization (PSO) [26] algorithm on the $Hits@10$ accuracy metric, defined below. Once embeddings P and Q have been computed, predictions are computed as $\hat{R} = PQ^{T}$.

To train our recommender and to test its performance we use the $Hits@10$ metrics [27]. We build from our dataset $\mathcal{K}$ a train set $\mathcal{K}^{\text{train}}$ with 80% of tuples in $\mathcal{K}$ and a test set with the rest $\mathcal{K}^{\text{test}} = \mathcal{K} \setminus \mathcal{K}^{\text{train}}$ of (u, i, r_{ui}) observations. After training with $\mathcal{K}^{train}$, we compute one recommendation for each user based on the K items with highest prediction score that are also previously not chosen, and compare these with $(\mathcal{U} \times \mathcal{I} \times \mathbb{R}_{+}) \setminus \mathcal{K}^{train}$ using the $Hits@K$ metric.

Let us call $\mathcal{T} := (\mathcal{U} \times \mathcal{I} \times \mathbb{R}^{+}) \setminus \mathcal{K}^{train}$, with $\mathcal{T} = \mathcal{T}^{u} \times \mathcal{T}^{i} \times \mathcal{T}^{r}$. $\mathcal{T}$ is the set of possible tuples user, item, and number of observations that are not in the train set. Let $u \in \mathcal{T}^{u}$ be a user, and $K \in \mathbb{N}_{+}^{*}$ a positive integer. The set of the K best predicted items for u among $\mathcal{T}$ is : $Top@K(u, \mathcal{T}) = \left\{i \in \mathcal{T}^{i}\ s.t.\ \left|\left\{j \in \mathcal{T}^{i} s.t. \hat{r}_{uj} \geq \hat{r}_{ui}\right\}\right| < K\right\}$. We also define the set of the K most observed items by user u that are not in the train set: $Pref@K(u, \mathcal{T}) = \left\{i \in \mathcal{T}^{i}\ s.t.\ \left|\left\{j \in \mathcal{T}^{i}\ s.t. r_{uj} \geq r_{ui}\right\}\right| < K\right\}$. We define the metric $Hits@K$ as the proportion of best items guessed for u and for all users:

$$Hits@K(u, \mathcal{T}) = \frac{|Top@K(u, \mathcal{T}) \cap Pref@K(u, \mathcal{T})|}{K}$$

and,

$$Hits@K(\mathcal{T}) = \sum_{u \in \mathcal{T}^{u}} \frac{Hits@K(u, \mathcal{T})}{|\mathcal{T}^{u}|}.$$

This metric is in fact the F1-score of the classification problem of finding the top K best items for each users. We assess the quality of our recommender by comparing it to random guess and to accuracies reported in the literature. For $K = 10$ the random guess performance of this task is $Hits@10((\mathcal{U} \times \mathcal{I} \times \mathbb{R}^{+}) \setminus \mathcal{K}^{\text{train}}) \approx 10^{-4}$. After training our NMF with optimized hyper-parameters using PSO we obtain a value of $Hits@10((\mathcal{U} \times \mathcal{I} \times \mathbb{R}_{+}) \setminus \mathcal{K}^{\text{train}}) = 0.35$ on our dataset, with $k = 12$ latent dimensions of representation in our embedding. This result compares positively to the estimated random guess and is in line with reported performance of real-world systems [28].

IV. Explaining the Representation Learning Space with Political Data

We now use our knowledge of the positions of users in the Left-Right and Anti-elite salience political dimensions to compute explanations of dimensions of the space learned by the NMF recommender. Each learned dimension represents a feature of users and items, estimated and leveraged by the recommender. Our explanation method seeks to relate these features to political ideologies measured by our two political dimensions. We call l_i the i-th learned dimensions, with $i = 0, 1, ..., k-1$ (with $k = 12$). To identify learned dimensions that might have a relation with political dimensions, we first measure the correlation between user positions according to these two types of dimensions. Among l_i for $i = 0...k-1$, we find that dimensions l_3 and l_4 hold statistically significant correlations with our political dimensions (Pearson correlation coefficient of respectively 0.24 for l_3 and -0.46 for l_4 with Left-Right with $p-value < 10^{-3}$ for both). We then measure the mutual information of the positions of users on each learned dimension with each of the two political dimensions. We find that the two learned dimensions that have the highest value of mutual information with political ones are the ones with the highest correlation, l_3 and l_4. The mutual information of positions of users on l_3 and Left-Right is 0.09, and the mutual information of positions of users on l_4 and Left-Right is 0.14. For reference, the next highest mutual information value is 0.06. We will now seek to characterize users with salience l_3 and l_4 values in the space subtended by our two political dimensions. Fig. 2 shows how different populations filtered by their values along these learned dimensions l_3 and l_4 distribute spatially in our political space. For l_3 and l_4 we consider a threshold l_3^* and l_4^* for which we will examine the distributions of users with values above those thresholds. For a given threshold, we select the population of users with values above the threshold, compute a kernel density estimation in our two-dimensional political space, and identify the level curve for the probability being equal to 0.5. Fig. 2 shows that, for raising values of thresholds l_3^* and l_4^* for l_3 and l_4, the selected sub-populations are further specialized in particular regions with precise political ideologies. Learned dimension l_3 increasingly identifies, with increasing values, left-leaning anti-elite users, while learned dimension l_4 increasingly identifies, with increasing values, right-leaning anti-elite users.

V. Discussion and conclusions

In this article we have used knowledge on the political opinions of users, represented in geometrical political space, to inspect learned representations computed and leveraged by a recommender system. These representations were computed by the recommender without using political positions of users. This has allowed us, for the first time, to show how and to what degree a recommender system might be learning political ideologies and attitudes, even when these are not variables supplied during training. In other words, our results constitute the first example of political algorithmic explainability. By inspecting dimensions of the latent learned space of our NMF recommender, we identified several dimensions that hold some information contained in the political space representation of users, as measured by mutual information. In particular for the two latent dimensions of the recommender holding the most political information, we were able to show that they correspond to precise political niches in terms of ideologies. The first of these two dimensions was found to be identifying users holding far-left ideologies and high anti-elite sentiments, while the second one was identifying users holding right-wing ideologies and relatively high anti-elite sentiments. While more analyses are needed to fully characterize how and if other latent dimensions of the recommender hold political importance, our results effectively trace a path bridging opinion dynamics and recommender systems explainability, opening the possibility to do political AI explainability and to modify the design of recommenders via constraint learning.

These results also call for further modelization and testing to improve our understanding of how political ideology might translate to activity that the recommender might observe and leverage. In particular, future results will focus on disentangling to what other features these ideologies might be related to, particularly in terms of classic socio-demographic and socio-professional characterizations of users.

Our results are easily generalizable to any recommender that relies on learned representations, making our method adaptable, for example, to encoder architectures. This new field of research may in turn interest several different actors. On the one hand, researchers interested in auditing may now inquire about the degree to which algorithms are leveraging political ideologies and attitudes when investigating the systemic consequences of given recommenders on online social systems. On the other hand, platforms interested in measuring and limiting the political consequences of recommendations might want to implement these methods, as they would allow to posit new optimization problems, maximizing accuracy while constraining the use of political information. While costly in modelization efforts, these constrained problems might help platforms to show compliance in minimizing systemic risks of AI systems in the face of upcoming risk-oriented regulation for algorithms.

Acknowledgments

This work has been funded by the "European Polarisation Observatory" (EPO) of CIVICA Research (co-)funded by EU's Horizon 2020 programme under grant agreement No 101017201, by the Data Intelligence Institute of Paris (diiP), and by the French National Agency for Research (ANR) under grants ANR-19-CE38-0006 "Geometry of Public Issues" (GOPI) and ANR-18-IDEX-0001 "IdEx Université de Paris".

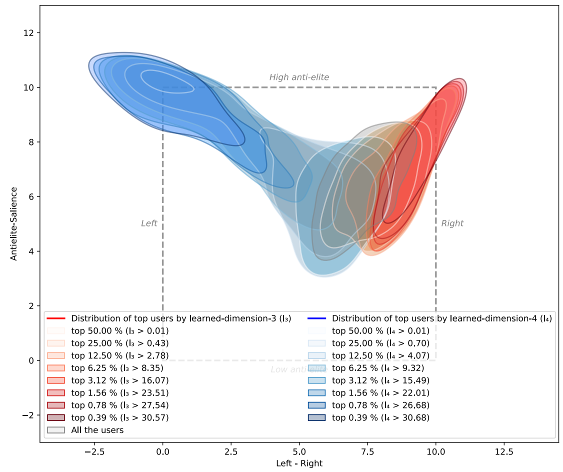

Fig. 2. Conditional distribution of the political attitudes of users according to growing values of positions on the latent space learned by the recommender. We display the two latent dimensions with the highest mutual information with political dimensions, the third l_3 and fourth l_4. For a given threshold value for l_3 or l_4 we compute a two-dimensional KDE and plot the level curve at probability equal 0.5. Users with highest values along these l_3 and l_4 feature dimensions, have salient left- and right-leaning ideologies.

This work is also supported by project Social Media for Democracy SoMe4Dem funded by EU's Horizon programme under grant agreement No 101094752.

Our study did not involve experimentation with human subjects, and all data used is publicly available through Twitter's API. Data declared the 19 March 2020 and 15 July 2021 at the registry of data processing at the *Fondation Nationale de Sciences Politiques* (Sciences Po) in accordance with General Data Protection Regulation 2016/679 (GDPR) and Twitter policy. For further details and the respective legal notice, please visit https://medialab.sciencespo.fr/en/activities/epo/.

REFERENCES

[1] N. Gaumont, M. Panahi, and D. Chavalarias, "Reconstruction of the socio-semantic dynamics of political activist Twitter networks—Method and application to the 2017 French presidential election," *PLOS ONE*, vol. 13, no. 9, p. e0201879, Sep. 2018.

[2] S. Messing and S. J. Westwood, "Selective Exposure in the Age of Social Media: Endorsements Trump Partisan Source Affiliation When Selecting News Online," *Communication Research*, vol. 41, no. 8, pp. 1042–1063, Dec. 2014.

[3] W. L. Bennett and S. Iyengar, "A New Era of Minimal Effects? the Changing Foundations of Political Communication," *Journal of Communication*, vol. 58, no. 4, pp. 707–731, Dec. 2008.

[4] C. A. Gomez-Uribe and N. Hunt, "The Netflix Recommender System: Algorithms, Business Value, and Innovation," *ACM Transactions on Management Information Systems*, vol. 6, no. 4, pp. 13:1–13:19, Dec. 2016.

[5] D. Jannach and M. Jugovac, "Measuring the Business Value of Recommender Systems," *ACM Transactions on Management Information Systems*, vol. 10, no. 4, pp. 16:1–16:23, Dec. 2019.

[6] D. O'Callaghan, D. Greene, M. Conway, J. Carthy, and P. Cunningham, "Down the (White) Rabbit Hole: The Extreme Right and Online Recommender Systems," *Social Science Computer Review*, vol. 33, no. 4, pp. 459–478, Aug. 2015.

[7] A. J. B. Chaney, B. M. Stewart, and B. E. Engelhardt, "How algorithmic confounding in recommendation systems increases homogeneity and decreases utility," in *Proceedings of the 12th ACM Conference on Recommender Systems*, ser. RecSys '18. New York, NY, USA: Association for Computing Machinery, Sep. 2018, pp. 224–232.

[8] M. Haim, A. Graefe, and H.-B. Brosius, "Burst of the Filter Bubble?" *Digital Journalism*, vol. 6, no. 3, pp. 330–343, Mar. 2018.

[9] Y. Benkler, R. Faris, and H. Roberts, *Network propaganda: Manipulation, disinformation, and radicalization in American politics*. Oxford University Press, 2018.

[10] E. Bakshy, S. Messing, and L. A. Adamic, "Exposure to ideologically diverse news and opinion on facebook," *Science*, vol. 348, no. 6239, pp. 1130–1132, 2015.

[11] P. Ramaciotti, R. Lamarche-Perrin, R. Fournier-S'Niehotta, R. Poulain, L. Tabourier, and F. Tarissan, "Measuring diversity in heterogeneous information networks," *Theoretical computer science*, vol. 859, pp. 80–115, 2021.

[12] J. Bobadilla, F. Ortega, A. Hernando, and A. Gutiérrez, "Recommender systems survey," *Knowledge-based systems*, vol. 46, pp. 109–132, 2013.

[13] P. Ramaciotti Morales and J.-P. Cointet, "Auditing the effect of social network recommendations on polarization in geometrical ideological spaces," in *Proceedings of the 15th ACM Conference on Recommender Systems*, 2021, pp. 627–632.

[14] R. Marcinkevičs and J. E. Vogt, "Interpretability and Explainability: A Machine Learning Zoo Mini-tour," Dec. 2020.

[15] N. Tintarev and J. Masthoff, "A Survey of Explanations in Recommender Systems," in *2007 IEEE 23rd International Conference on Data Engineering Workshop*, Apr. 2007, pp. 801–810.

[16] R. Bakker, S. Jolly, and J. Polk, "Complexity in the European party space: Exploring dimensionality with experts," *European Union Politics*, vol. 13, no. 2, pp. 219–245, Jun. 2012.

[17] P. Ramaciotti and Z. Vagena, "Embedding social graphs from multiple national settings in common empirical opinion spaces," in *2022 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)*. IEEE, 2022, pp. 60–67.

[18] P. Ramaciotti Morales, J.-P. Cointet, and G. M. Zolotoochin, "Unfolding the dimensionality structure of social networks in ideological embeddings," Aug 2021. [Online]. Available: https://hal.archives-ouvertes.fr/hal-03315759

[19] P. Ramaciotti Morales, J.-P. Cointet, G. Muñoz Zolotoochin, A. Fernández Peralta, G. Iñiguez, and A. Pournaki, "Inferring attitudinal spaces in social networks," *Social Network Analysis and Mining*, vol. 13, no. 1, pp. 1–18, 2023.

[20] R. Bakker, L. Hooghe, S. Jolly, G. Marks, J. Polk, J. Rovny, M. Steenbergen, and M. A. Vachudova, "2019 chapel hill expert survey," *Chapel Hill*, 2020, www.chesdata.eu.

[21] Y. Koren, "Factorization meets the neighborhood: a multifaceted collaborative filtering model," in *Proceedings of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining*, 2008, pp. 426–434.

[22] J. Bobadilla, F. Ortega, A. Hernando, and A. Gutiérrez, "Recommender systems survey," *Knowledge-Based Systems*, vol. 46, pp. 109–132, Jul. 2013.

[23] A. Cichocki and A.-H. Phan, "Fast Local Algorithms for Large Scale Nonnegative Matrix and Tensor Factorizations," *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences*, vol. E92.A, no. 3, pp. 708–721, 2009.

[24] D. Lee and H. S. Seung, "Algorithms for Non-negative Matrix Factorization," in *Advances in Neural Information Processing Systems*, vol. 13. MIT Press, 2000.

[25] C. Boutsidis and E. Gallopoulos, "SVD based initialization: A head start for nonnegative matrix factorization," *Pattern Recognition*, vol. 41, no. 4, pp. 1350–1362, Apr. 2008.

[26] J. Kennedy and R. Eberhart, "Particle swarm optimization," in *Proceedings of ICNN'95 - International Conference on Neural Networks*, vol. 4, Nov. 1995, pp. 1942–1948 vol.4.

[27] A. Gunawardana and G. Shani, "A Survey of Accuracy Evaluation Metrics of Recommendation Tasks," *Journal of Machine Learning Research*, pp. 2935–2962, 2009.

[28] W. Hu, M. Fey, M. Zitnik, Y. Dong, H. Ren, B. Liu, M. Catasta, and J. Leskovec, "Open Graph Benchmark: Datasets for Machine Learning on Graphs," Feb. 2021.