

Description du projet

Datagotchi : Les élections fédérales canadiennes de 2023

La Chaire de leadership en enseignement des sciences sociales numériques (CLESSN) souhaite lancer une nouvelle version améliorée de la plateforme web Datagotchi. Comme pour les éditions des élections fédérales canadiennes de 2021 et des élections générales québécoises de 2022, l'application permettra aux électeurs de découvrir comment leurs caractéristiques liées à leur mode de vie peuvent être utilisées par l'équipe de recherche pour prédire leur choix de vote lors des prochaines élections fédérales canadiennes de 2023.

État des connaissances et théorie

L'ère numérique a apporté des changements significatifs à la recherche en sciences sociales. Les chercheurs ont désormais accès à des données numériques massives, provenant de sources telles que les médias sociaux, les transactions financières, les capteurs et les dispositifs portables. Ces données peuvent être collectées et traitées plus rapidement que jamais grâce à l'utilisation de technologies telles que l'analyse de données en temps réel, l'apprentissage automatique et l'intelligence artificielle. Cela permet aux chercheurs de découvrir de nouveaux modèles de comportement, de prévoir les tendances futures et de mieux comprendre les phénomènes sociaux.

Le projet Datagotchi vise à répondre à une problématique théorique précise en mettant les données au service des sciences sociales et de la recherche en opinion publique. Le pouvoir prédictif et explicatif des variables sociodémographiques traditionnelles — telles que la religion, le revenu et l'éducation — étant en baisse, une restructuration de la méthodologie de recherche est essentielle. Les données numériques massives permettent désormais d'étudier divers phénomènes avec une précision temporelle et géographique granulaire (Lazer & Radford, 2017). Dans cette optique, certaines études ont examiné des réseaux d'interaction impliquant des millions de personnes sur Twitter et sur Yahoo pour faire des prédictions électorales autrefois impensables (State, Park, Weber, & Macy, 2015; De Beauchamp, 2017). Ces recherches prometteuses illustrent dans quelle mesure les données massives numériques permettent la construction d'une nouvelle « science de la société » (Lazer & Radford, 2017, p.25) qui étudie la société comme une entité entière composée de sous-systèmes et d'individus connectés de manière dynamique selon leurs « manières d'être » et leur localisation géospatiale.

Les études quantitatives en comportements politiques utilisent fréquemment des données de sondages publics et électoraux pour étudier les comportements et les attitudes des individus. Cependant, les variables sociodémographiques traditionnelles telles que la religion, le revenu et l'éducation ont une capacité explicative et prédictive en baisse, ce qui nécessite l'intégration de marqueurs de socialisation diversifiés pour une meilleure compréhension de la fragmentation croissante de la sphère culturelle et sociale (Bishop, 2009 ; Bertrand et Kamenica, 2018). Ces variables misant sur le style de vie des individus pourraient permettre aux modèles de prédiction électorale de regagner leur pouvoir explicatif et prédictif. Les possibilités d'accumulation d'un volume important de données permettent de concevoir des analyses sophistiquées avec un pouvoir de précision impressionnant. Les études récentes démontrent que des outils de mesure plus raffinés sont capables de cerner les attitudes latentes, contrairement aux premiers travaux empiriques en opinion publique qui questionnaient l'existence même des attitudes politiques individuelles (Converse, 1964 ; Zaller, 1992 ; Stimson, 2018). Ainsi, les avancées dans la collecte et l'analyse de données permettent une meilleure compréhension des transformations sociales profondes de notre époque (Lamont et al., 1996 ; Fischer et Mattson, 2009).

Des outils ont été développés et validés pour mesurer les attitudes sur les enjeux éthiques et moraux (Ansolabehere, Rodden et Snyder, 2008). De plus, l'étude des valeurs et des croyances profondes joue un rôle crucial dans la recherche contemporaine en opinion publique et en psychologie politique (Feldman, 1988; Inglehart, 1990; Jacoby, 2006). La diminution de la capacité prédictive des modèles traditionnels, combinée aux nouvelles données numériques, a conduit certains chercheurs à s'intéresser aux "lifestyle politics" (Giddens, 1991; Bennett, 1998; Micheletti, 2003). La fragmentation des unités de socialisation, telles que le style de musique, l'allégeance esthétique ou artistique, le mode de vie quotidien, le style vestimentaire, les habitudes alimentaires, etc., pourrait aider à expliquer les valeurs, les préférences, les attitudes et les opinions individuelles (Weiss, 1988; Chaney, 2002; Stolle, Hooghe et Micheletti, 2005). Les travaux de DellaPosta, Shi et Macy (2015) et de Mutz et Rao (2018) sur les raisons expliquant l'attachement des progressistes américains aux cafés latte en sont un exemple. Cependant, ces variables sociales en constante évolution sont peu explorées dans les études en opinion publique et en psychologie. Leur intégration promet des avancées scientifiques, théoriques et empiriques prometteuses.

Objectifs

Les objectifs centraux du projet demeurent les mêmes que dans les éditions précédentes, bien que l'interface visuelle et l'expérience utilisateur seront améliorées.

1. L'objectif scientifique de Datagotchi est d'explorer la relation entre les habitudes de vie et les attitudes et comportements politiques. Les chercheurs utilisent cet outil de collecte de données confidentielles pour mieux comprendre comment les habitudes de consommation, de loisirs, de préférences musicales et de films peuvent influencer le choix de vote, qui reste peu exploré dans la recherche actuelle.
2. Datagotchi a également pour objectif de démocratiser les outils de ciblage utilisés par les partis politiques et de sensibiliser les citoyens aux enjeux démocratiques et éthiques soulevés par l'utilisation de l'IA en politique. La plateforme interactive cherche à informer les utilisateurs de la façon dont les données numériques sont utilisées à des fins électorales et des enjeux éthiques qui en découlent. Elle se veut ainsi un outil de conscientisation et d'éveil face à l'utilisation à grande échelle des données personnelles par les partis politiques.
3. Enfin, Datagotchi vise également à démocratiser l'analyse statistique en construisant une plateforme visuellement raffinée et intuitive pour les utilisateurs de tous horizons, y compris des citoyens d'âge scolaire, des chercheurs, des stratèges ou des décideurs publics. La plateforme sera attrayante pour un large public en raison de sa facilité d'utilisation et de la rigueur scientifique de ses analyses. En outre, elle permettra de vulgariser les connaissances scientifiques et de transférer des connaissances à un large public.

Méthodologie

Infrastructure :

La nouvelle version de Datagotchi sera basée sur la même infrastructure que la précédente, mais avec des améliorations visuelles telles que l'ajout d'animations, une amélioration de l'application pour la version "Desktop", et quelques changements pour simplifier son utilisation, tels que des modifications du tableau de bord pour rendre la manipulation des données plus intuitive pour l'utilisateur.

Activités de collecte :

Une nouvelle étude pilote sera réalisée à partir du panel de Leger Opinion (LÉO) pour la sélection des questions à ajouter au sondage de la plateforme.

Analyse des données :

Différents modèles statistiques, tels que la régression logistique et les modèles de forêts d'arbres décisionnels, seront comparés pour déterminer celui qui permettra d'obtenir les prédictions de vote les plus fiables et performantes. Les analyses préliminaires seront réalisées à partir de données collectées par les sondages pilotes, et un algorithme d'apprentissage sera entraîné sur ces données pour faire des prédictions sur de nouvelles données. Les utilisateurs pourront fournir des commentaires sur la précision des prédictions, ce qui contribuera à améliorer l'algorithme pour les autres utilisateurs.

Conservation des données :

Les données seront conservées dans la banque de données existante de Datagotchi sous la responsabilité du professeur Yannick Dufresne, et destinée aux participants ayant consenti à cette banque. La gestion et la conservation des données des participants qui ont donné leur consentement se feront conformément à la politique de gestion de la banque de données de Datagotchi (N° d'approbation : 2021-065 CG).

Retombées

Le développement de la recherche sur les fondements et la structure de l'opinion publique et des comportements politiques est d'une importance cruciale. Non seulement cette recherche aura un impact sur la sphère académique, en contribuant de manière significative au champ de la science politique, de la psychologie et des sciences sociales en général, mais elle aura également des répercussions publiques importantes. En effet, en cherchant à améliorer la capacité et la précision des modèles quantitatifs de prédiction en science politique, les résultats de ce projet seront bénéfiques pour les décideurs publics, les acteurs politiques, les médias et le grand public. Grâce à la diffusion de la plateforme et à la vulgarisation des résultats de recherche, ce projet permettra aux citoyens de mieux comprendre les structures culturelles et institutionnelles qui les entourent au quotidien. Cette entreprise vise à démocratiser l'utilisation d'outils de ciblage par les partis politiques, en fournissant des analyses vulgarisées et rigoureuses, et à sensibiliser les citoyens aux enjeux soulevés par l'ère numérique.

Bibliographie

- Ansolabehere, S., Rodden, J. et Snyder, J. M. (2008). The Strength of Issues : Using Multiple Measures to Gauge Preference Stability, Ideological Constraint, and Issue Voting. *American Political Science Review*, 102(2), 215–232.
- Barberá, P., Wang, N., Bonneau, R., Jost, J. T., Nagler, J., Tucker, J. et González-Bailón, S. (2015). The Critical Periphery in the Growth of Social Protests. *PloS One*, 10(11), 1-15.
- Beauchamp, N. (2017). Predicting and Interpolating State-Level Polls Using Twitter Textual Data. *American Journal of Political Science*, 61(2), 490–503.
- Bennett, W. L. (1998). The Uncivic Culture : Communication, Identity, and the Rise of Lifestyle Politics. *PS : Political Science & Politics*, 31(4), 741–761.
- Bertrand, M. et Kamenica, E. (2018). *Coming Apart ? Cultural Distances in the United States Over Time* (Rapport technique). National Bureau of Economic Research.
- Bishop, B. (2009). *The Big Sort : Why the Clustering of Like-Minded America is Tearing Us Apart*. Houghton Mifflin Harcourt.
- Campbell, Angus. 1960. "Converse, Pe Miller, We and Stokes, de (1960). the American Voter." *Ann Arbor*.
- Chaney, D. C. (2002). From Ways of Life to Lifestyle : Rethinking Culture as Ideology and Sensibility. In *Culture in the Communication Age* (p. 85-98). Routledge.
- Converse, P. E. (1964). The Nature of Belief Systems in Mass Publics. In D. E. Apter (Ed.), *Ideology and Discontent* (pp. 206–261). Free Press of Glencoe.
- Dalton, Russell J. 2007. "Partisan Mobilization, Cognitive Mobilization and the Changing American Electorate." *Electoral Studies* 26 (2). Elsevier: 274–86.
- Dalton, Russell J, and others. 1988. *Citizen Politics in Western Democracies: Public Opinion and Political Parties in the United States, Great Britain, West Germany, and France*. Chatham House Chatham, NJ.
- DellaPosta, D., Shi, Y. et Macy, M. (2015). Why do Liberals Drink Lattes? *American Journal of Sociology*, 120(5), 1473–1511.

Feldman, S. (1988). Structure and Consistency in Public Opinion : The Role of Core Beliefs and Values. *American Journal of Political Science*, 416–440.

Fischer, C. S. et Mattson, G. (2009). Is America Fragmenting ? *Annual Review of Sociology*, 35, 435-455.

Giddens, A. (1991). *Modernity and Self-Identity : Self and Society in the Late Modern Age*. Stanford University Press.

Inglehart, R. (1990). *Culture Shift in Advanced Industrial Society*. Princeton University Press.

Jacoby, W. G. (2006). Value Choices and American Public Opinion. *American Journal of Political Science*, 50(3), 706–723.

Krosnick, Jon A. 1990. "Government Policy and Citizen Passion: A Study of Issue Publics in Contemporary America." *Political Behavior* 12 (1). Springer: 59–92.

Lamont, M., Schmalzbauer, J., Waller, M. et Weber, D. (1996). Cultural and Moral Boundaries in the United States : Structural Position, Geographic Location, and Lifestyle Explanations. *Poetics*, 24(1), 31-56.

Lazarsfeld, Paul Felix, Bernard Berelson, and Hazel Gaudet. 1944. "The People's Choice." Duell, Sloan & Pearce.

Lazer, David, and Jason Radford. 2017. "Data Ex Machina: Introduction to Big Data." *Annual Review of Sociology* 43. Annual Reviews: 19–39.

Micheletti, M. (2003). Shopping with and for Virtues. In *Political Virtue and Shopping* (pp. 149–168). Springer.

Mutz, D. C. et Rao, J. S. (2018). The Real Reason Liberals Drink Lattes. *PS : Political Science & Politics*, 1–6.

State, B., Park, P., Weber, I. et Macy, M. (2015). The Mesh of Civilizations in the Global Network of Digital Communication. *PloS One*, 10(5), e0122543.

Stimson, J. (2018). *Public Opinion in America : Moods, Cycles, and Swings*. Routledge.

Stolle, D., Hooghe, M. et Micheletti, M. (2005). Politics in the Supermarket : Political Consumerism as a Form of Political Participation. *International Political Science Review*, 26(3), 245-269.

Wang, W., Rothschild, D., Goel, S., et Gelman, A. (2015). Forecasting Elections with Non- Representative Polls. *International Journal of Forecasting*, 31(3), 980–991.

Weiss, M. J. (1988). *The Clustering of America*. Harper Collins.

Wood, Wendy, Carl A Kallgren, and Rebecca Mueller Preisler. 1985. "Access to Attitude- Relevant Information in Memory as a Determinant of Persuasion: The Role of Message Attributes." *Journal of Experimental Social Psychology* 21 (1). Elsevier: 73–85.

Zaller, J. et Feldman, S. (1992). A Simple Theory of the Survey Response : Answering Questions versus Revealing Preferences. *American Journal of Political Science*, 36(3), 579- 616. Zaller, J. R. (1992). *The Nature and Origins of Mass Opinion*. Cambridge University Press.