



HAL
open science

Caractérisation des fausses nouvelles Une approche basée sur la modélisation conceptuelle

Nicolas Belloir, Wassila Ouerdane, Oscar Pastor

► **To cite this version:**

Nicolas Belloir, Wassila Ouerdane, Oscar Pastor. Caractérisation des fausses nouvelles Une approche basée sur la modélisation conceptuelle. 41st INFORSID 2023, May 2023, La Rochelle, France. pp.103-104. hal-04453332

HAL Id: hal-04453332

<https://hal.science/hal-04453332v1>

Submitted on 12 Feb 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Caractérisation des fausses nouvelles

Une approche basée sur la modélisation conceptuelle

Nicolas Belloir^{1,2}, Wassila Ouerdane³, Oscar Pastor⁴

1. CReC St-Cyr, Académie Militaire de St Cyr Coëtquidan, Guer, France

2. IRISA, Vannes, France

3. MISC, Centrale Supélec, Université Paris-Saclay, Gif sur Yvette, France

4. PROS Research Group, Universitat Politècnica de València, Spain

Contact : nicolas.belloir@irisa.fr

RÉFÉRENCE DE L'ARTICLE INTERNATIONAL. Cet article est un résumé de: Belloir N., Ouerdane W., Pastor, O.: *Characterizing Fake News: A Conceptual Modeling-based Approach*. 41st International Conference on Conceptual Modeling (ER) 2022 : Hyderabad, India: 115-129

MOTS-CLES : Modélisation conceptuelle, fausses nouvelles, intelligence artificielle explicable

KEYWORDS : Conceptual Modeling, Fake News, Explainable Artificial Intelligences

La désinformation, et en particulier l'un de ses mécanismes les plus voyants, les fausses nouvelles, est devenue un réel problème de société. Être capable d'identifier puis de lutter contre des fausses nouvelles est une nécessité. Dans ce contexte, il ne suffit pas de pouvoir identifier une fausse nouvelle, il faut également pouvoir expliquer la raison qui amène à ce résultat. Même si de nombreuses actions louables telles que le « fact checking » existent, force est de constater que seule l'utilisation de processus automatisés permet de lutter contre la masse des fausses nouvelles.

De nombreuses approches visant à détecter des fausses nouvelles de manière automatique existent, utilisant en majorité des techniques d'intelligence artificielle (IA). Ces dernières, souvent basées sur des approches de « machine learning », sont inégales et forcément dépendantes du jeu de données avec lequel elles sont entraînées. De plus, si ces techniques peuvent dire si oui ou non une nouvelle est une fausse nouvelle, elles ne permettent pas d'expliquer pourquoi.

Dans ce contexte, nous défendons l'idée qu'il est nécessaire de disposer d'un cadre précis permettant de comprendre les concepts inhérents au fonctionnement des fausses nouvelles. L'objectif est de pouvoir s'appuyer sur ce cadre de manière à justifier par la description les mécanismes impliqués permettant de déterminer en quoi telle nouvelle est une fausse nouvelle. En effet, tout système d'information, destiné à traiter des informations sur les fausses nouvelles, doit identifier en détail

quelles sont les entités pertinentes qui caractérisent conceptuellement les différentes dimensions de ces dernières. Plus précisément, nous suivons un processus basé sur l'intelligence artificielle explicable (XAI) proposé par (Spreeuwenberg, 2019) pour faciliter la construction de modèles bien justifiés et explicables pour la détection de fausses nouvelles. Notre objectif est de proposer une approche qui soit compréhensible, fiable et gérable par les humains. Les différentes étapes suggérées par cette approche sont les suivantes : (i) Obtenir une compréhension partagée du domaine, (ii) comprendre la tâche à effectuer et sélectionner le bon périmètre, (iii) collecter les bonnes données et améliorer leur qualité, (iv) sélectionner les techniques d'IA qui donnent des résultats, (v) générer de bonnes explications, et (vi) faire évoluer la solution dans le temps. Notre contribution établit la fondation d'un tel processus en résolvant la première étape, essentielle, qui consiste à obtenir une compréhension commune de ce qu'est une fausse nouvelle à travers un modèle conceptuel. À partir de cette base conceptuelle solide, le reste du processus XAI proposé peut être appliqué de manière fiable. L'explicabilité de notre approche est guidée par le modèle conceptuel qui est conforme au cœur de la contribution : avoir une définition ontologiquement bien fondée de ce qu'est une fausse nouvelle, qui est directement dérivée du modèle conceptuel.

Ainsi, nous avons spécifié un modèle conceptuel caractérisant le concept de fausse nouvelle et ses principales caractéristiques. Dans (Belloir et al, 2022), nous avons identifié ces dernières à partir de la littérature SHS, sans toutefois les spécifier. Parmi celles-ci, la première concerne l'*origine* des fausses nouvelles. Le constat est fait qu'une fausse nouvelle est rarement créée par une personne isolée. Identifier sa chaîne de création permet de comprendre quel objectif elle sert, et donc de mieux la mesurer et la contrer. Deuxièmement, les fausses nouvelles sont généralement construites sur une *déformation de la réalité*. Identifier cette distorsion est un point essentiel pour permettre son explication. Troisièmement, afin de renforcer sa crédibilité, une fausse nouvelle s'appuie souvent sur une « autorité ». Cette dernière peut être réelle ou fictive, et son message peut lui-même être sorti de son contexte ou totalement inventé. Là encore, il est opportun de pouvoir l'identifier. Enfin, les fausses nouvelles sont construites pour influencer l'opinion de leur cible en générant une *charge émotionnelle*. La réaction de la cible l'amène à tirer certaines conclusions et à changer son opinion sur un sujet. Savoir identifier ce mécanisme permet ainsi de le pointer du doigt. Ainsi le modèle conceptuel que nous présentons est articulé autour des trois dimensions principales suivantes : (i) l'identification de l'origine des fausses nouvelles, (ii) la relation entre les faits vrais et faux, et (iii) la cible comprenant l'opinion et l'émotion. Le modèle conceptuel ainsi proposé a été modélisé en UML, en focalisant sur un niveau de détail plus fin et exploitable.

Bibliographie

- Belloir N, Ouerdane W, Pastor O, Frugier E et de Barmon L.A. (2022), A conceptual characterization of fake news: a positioning paper. *Actes de RCIS 2022*. Barcelone.
- Spreeuwenberg, S. (2019): *AIX: Artificial Intelligence needs eXplanation: Why and how transparency increases the success of AI solutions*. LibRT BV, Amsterdam.