

Un modèle conceptuel de narration de données

Faten El Outa*, Patrick Marcel*, Veronika Peralta*

* Laboratoire LIFAT, Université de Tours
prenom.nom@univ-tours.fr

Résumé. La narration de données est l'activité de production de narrations étayée par des faits extraits de l'analyse de données, à l'aide de visualisations potentiellement interactives. Malgré l'intérêt croissant de plusieurs communautés, il n'y a pas de définition consensuelle ni de modèle formel de narration de données. Dans cet article, nous proposons un modèle conceptuel de narration de données. Il décrit tout le processus de narration : la définition d'un objectif d'analyse, la collecte et l'exploration de données, la découverte de faits marquants, l'extraction des principaux messages, leur structuration pour construire l'intrigue et la présentation finale par des moyens visuels.

1 Introduction

La narration est considérée comme l'une des plus anciennes professions et un pilier de la communication d'informations comme moyen éducatif. Elle concerne l'usage de techniques pour communiquer une histoire à une audience. Plus récemment, la narration de données (*i.e.* narrer une histoire avec des visualisations de données) (Hullman et al., 2013) reçoit un intérêt croissant dans plusieurs communautés telles que le journalisme, les affaires, le e-gouvernement et la science des données. Elle est décrite comme l'activité de production de narrations, étayée par des faits, extraits de l'analyse de données, à l'aide de visualisations potentiellement interactives (Carpendale et al., 2016). Plus précisément, ces narrations peuvent être vues comme des séquences ordonnées d'étapes, chacune pouvant contenir des mots, images, visualisations, audio, vidéo, etc., basés sur des données (Kosara et Mackinlay, 2013).

En dehors de ces considérations générales, et à notre connaissance, il n'existe pas de définition consensuelle ni de modèle formel de la narration de données. Bien que la narration des données ait essentiellement attiré l'attention de la communauté visualisation (Carpendale et al., 2016; Kosara et Mackinlay, 2013), nous affirmons : (i) qu'une approche plus globale est nécessaire dans des domaines tels que la visualisation, la gestion des données, l'exploration de données et l'apprentissage automatique, et (ii) que la modélisation conceptuelle du domaine devrait orienter les recherches futures, afin d'aider à comprendre les narrations de données. Une telle modélisation des aspects et des choix de conception dans le domaine de la narration des données, fournie aux concepteurs de systèmes et d'algorithmes, permettra de faciliter l'exploration