

La Charte Éthique et Big Data : pour des ressources pour le TAL (enfin !) traçables et pérennes

Karën Fort^{1,2,3} Alain Couillault^{4,5}

(1) Université de Lorraine

(2) LORIA 54500 Vandœuvre-lès-Nancy

(3) ATALA

(4) Université de La Rochelle

(5) APROGED

`karen.fort@loria.fr, alain.couillault@univ-lr.fr`

RÉSUMÉ

La charte Ethique & Big Data a été conçue à l'initiative de l'ATALA¹, de l'AFCP², de l'APROGED³ et de CAP DIGITAL⁴, au sein d'un groupe de travail mixte réunissant d'autres partenaires académiques et industriels (tels que le CERSA-CNRS, Digital Ethics, Eptica-Lingway, le cabinet Itéanu ou ELRA/ELDA). Elle se donne comme objectif de fournir des garanties concernant la traçabilité des données (notamment des ressources langagières), leur qualité et leur impact sur l'emploi. Cette charte a été adoptée par Cap Digital (co-rédacteur). Nous avons également proposé à la DGLFLF et à l'ANR de l'utiliser. Elle est aujourd'hui disponible sous forme de wiki⁵, de fichier pdf et il en existe une version en anglais. La charte est décrite en détails dans (Couillault et Fort, 2013).

ABSTRACT

The Ethics & Big Data Charter : for tractable and lasting NLP resources

The Ethics & Big Data Charter was designed by ATALA, AFCP, APROGED and CAP DIGITAL, in a working group including other academic and industrial partners (such as CERSA-CNRS, Digital Ethics, Eptica-Lingway, Itéanu office or ELRA/ELDA). Its aims at ensuring the traceability and quality of the data (including language resources), how they are produced and their impact on working conditions. This charter has been adopted by Cap Digital (co-writer). We also proposed it to DGLFLF and ANR. As of today, it is available as a wiki, a pdf file and an English version⁶. The charter is detailed in (Couillault et Fort, 2013).

MOTS-CLÉS : éthique, big data, ressources langagières.

KEYWORDS: ethics, big data, language resources.

1. <http://www.atala.org/>

2. <http://www.afcp-parole.org/>

3. <http://www.aproged.org/>

4. <http://www.capdigital.com/>

5. <http://wiki.ethique-big-data.org>

6. http://wiki.ethique-big-data.org/index.php?title=Accueil#English_Version

Références

COUILLAUT, A. et FORT, K. (2013). Charte éthique et big data : parce que mon corpus le vaut bien ! *In Actes de colloque international Corpus et Outils en Linguistique, Langues et Parole : Statuts, Usages et Mésusages*, Strasbourg, France.