

# Stunning Doodle: un outil pour la visualisation et l'analyse conjointe de graphes de connaissances et leurs plongements

A. Ettore, A. Bobasheva, F. Michel, C. Faron

Université Côte d'Azur, CNRS, Inria, I3S, France

{aettore, abobasheva, fmichel, faron}@i3s.unice.fr

## Résumé

*Ces dernières années, l'utilisation croissante des graphes de connaissances dans des domaines nouveaux et variés nécessite de rendre ces ressources accessibles et compréhensibles par des utilisateurs aux profils de plus en plus divers. De plus, l'application de méthodes d'apprentissage automatique sur des plongements de graphes de connaissances renforce encore la visibilité de ce type de représentation, mais soulève un nouveau problème de compréhension et d'interprétabilité de ces plongements. Dans ce travail, nous montrons comment des techniques de visualisation peuvent être utilisées pour explorer et interpréter conjointement les graphes de connaissances et les plongements de graphes.*

## Mots-clés

*Plongements de graphes de connaissances, Visualisation.*

## Abstract

*In recent years, the growing application of Knowledge Graphs to new and diverse domains has created the need to make these resources accessible and understandable by users with increasingly diverse backgrounds. Moreover, the emerging use of Knowledge Graph Embeddings as input features of Machine Learning methods has given even more visibility to this kind of representation, but raising the new issue of understandability and interpretability of such embeddings. In this paper, we show how visualization techniques can be used to jointly explore and interpret both Knowledge Graphs and Graph Embeddings.*

## Keywords

*Knowledge Graph Embeddings, Visualization.*

## 1 Introduction

Au cours de la dernière décennie, l'adoption des graphes de connaissances (KGs) dans de multiples domaines n'a cessé de croître de sorte que de plus en plus de projets s'appuient sur ce type de représentation pour stocker leurs données en en conservant toute la richesse sémantique. L'une des raisons de leur succès croissant est la possibilité de leur appliquer des méthodes d'apprentissage automatique en utilisant une représentation à faible dimension de ces KGs, obtenue par le processus de plongement de graphes. Néanmoins, il n'est pas aisé de comprendre les informations cap-

turées par les plongements de graphes (GEs). En effet, les GEs sont calculés à l'aide de techniques d'apprentissage automatique, des "boîtes noires" qui traduisent chaque élément du graphe en un vecteur de faible dimension. Même si le processus algorithmique de calcul des plongements est bien compris, une relation entre les caractéristiques et le rôle de l'élément du graphe et sa représentation vectorielle dans l'espace de plongement ne peut être établie avec certitude. En d'autres termes, il n'est pas aisé de répondre à certaines questions, notamment :

- Que représentent mes plongements ?
- Comment sont-ils liés à la structure et à la sémantique de mon KG ?
- Comment puis-je améliorer mes plongements pour qu'ils soient mieux adaptés à mon application downstream ?

Récemment, plusieurs efforts de recherche ont été faits dans cette direction pour commencer à donner du sens aux informations capturées par les GEs. Certaines approches mettent en œuvre des stratégies d'explication pour des modèles de plongement spécifiques [3]; tandis que d'autres proposent des méthodes pour vérifier si un élément de connaissance spécifique représenté dans un KG est réellement encodé et capturé par ses GEs [1].

Dans [2], nous avons abordé cette question d'un point de vue différent. Nous pensons que l'information portée par les GEs pourrait être explorée et dévoilée grâce à l'utilisation de techniques de visualisation qui favoriseraient la découverte de la relation logique entre le graphe et ses plongements. Notre objectif est donc de fournir un outil de visualisation permettant l'analyse et le décodage des informations capturées par les KGEs en dévoilant la relation entre, d'une part, la structure et la sémantique du KG et, d'autre part, les KGEs générés à partir de celui-ci. Nous présentons *Stunning Doodle* [2], un outil conçu pour la visualisation de KGs basés sur RDF et de leur GEs. *Stunning Doodle* fournit une visualisation du graphe à analyser, offrant un aperçu riche de la structure et de la sémantique des données. Cette visualisation est ensuite enrichie en connectant les sommets du graphe avec les GEs correspondants à analyser.

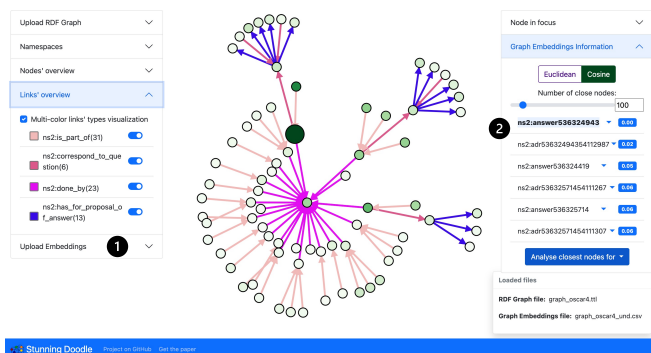


FIGURE 1 – *Stunning Doodle* montrant sommets proches dans l’espace de plongement.

## 2 Analyse des GEs avec Stunning Doodle

*Stunning Doodle* a été développé pour combler un manque dans le domaine de l’analyse visuelle des KGEs. Son objectif principal est de fournir aux utilisateurs une visualisation avancée de graphes RDF, enrichie d’informations extraites des GEs générés à partir de ces graphes. Premièrement, *Stunning Doodle* permet aux utilisateurs de visualiser et de naviguer dans les graphes RDF grâce à son système d’exploration de graphe et à l’utilisation de fonctionnalités avancées de filtrage et de personnalisation basées sur la sémantique des sommets et des arêtes.

La fonctionnalité clé fournie par *Stunning Doodle* est la possibilité de visualiser conjointement les KGEs et le KG à partir duquel ils sont générés. Plus précisément, *Stunning Doodle* permet de visualiser, pour chaque sommet, ses sommets les plus proches dans l’espace de plongement, selon la distance euclidienne ou la distance cosinus. La figure 1 montre un exemple d’analyse des KGEs. Après avoir téléchargé un fichier stockant les GEs calculés à partir du KG (1) visualisé, l’utilisateur peut sélectionner un sommet d’intérêt et visualiser ses plus proches voisins dans l’espace de plongement. Les sommets les plus proches sont affichés avec un gradient de couleur qui représente leur distance par rapport au sommet dont le plongement est analysé ; les sommets plus sombres sont plus proches (dans l’espace de plongement) du sommet sélectionné tandis que les sommets plus clairs sont plus éloignés. Si une relation entre un couple de sommets visualisés existe dans le KG, alors l’arête correspondante est directement affichée dans le graphique selon les paramètres de personnalisation sélectionnés. La liste des sommets les plus proches avec leur distance est affichée dans le menu “Graph Embeddings Information” sur le côté droit de la page (2). Outre cette liste, des options supplémentaires permettent à l’utilisateur de choisir la métrique de distance souhaitée (euclidienne ou cosinus dans l’implémentation actuelle) et de personnaliser le nombre de sommets les plus proches à afficher. L’exemple de la figure 1 montre les 100 sommets les plus proches dans l’espace de plongement du sommet `ns1:answer536324943`, selon la distance cosinus. Le sommet actuellement sélectionné

(de plus grande taille) est ici le sommet à partir duquel les distances sont calculées. Il est facilement reconnaissable sur la visualisation par sa couleur sombre due au fait que sa distance par rapport à lui-même est nulle. Son URI est donc le premier élément de la liste des sommets les plus proches. Les différentes nuances de vert de chaque sommet mettent clairement en évidence les sommets les plus proches parmi les 100 visualisés, tandis que les liens montrent comment ils sont connectés dans le KG. Tout en visualisant les voisins d’un sommet dans l’espace de plongement, il est toujours possible d’accéder aux fonctionnalités de navigation dans le KG grâce aux boutons de la section “Node in focus”. Par conséquent, tout sommet affiché peut être développé pour afficher les sommets qui lui sont liés dans le KG, même s’ils ne sont pas proches dans l’espace de plongement. Naturellement, tout nouveau sommet peut être développé à son tour pour visualiser la partie souhaitée du KG. Toutes les arêtes et les sommets affichés dont les plongements ne sont pas proches du sommet initial seront visualisés en fonction des options de personnalisation sélectionnées dans les sections “Nodes’ overview” et “Links’ overview”.

En résumé, *Stunning Doodle* permet à l’utilisateur de comprendre en un coup d’œil quels sommets d’un KG sont considérés comme similaires dans l’espace de plongement, tout en gardant la trace de leurs connexions dans le KG. Cela permet d’avoir un aperçu immédiat des informations capturées par les KGEs, par exemple quels prédicats ont le plus grand impact ou quels modèles de connectivité sont les plus pris en compte pendant le processus de plongement.

## 3 Conclusion

*Stunning Doodle* est une première étape pour combler le manque dans le domaine de l’analyse visuelle des KGEs. Cet outil de visualisation permet de trouver un lien entre le contenu et la structure de n’importe quel KG et ses plongements. Nous avons implémenté un ensemble de fonctionnalités pour faciliter l’exploration et la compréhension de n’importe quel KG et pour analyser les KGEs en connectant les deux, et en donnant du sens aux informations capturées par les KGEs.

## Références

- [1] Antonia Ettore, Anna Bobasheva, Catherine Faron, and Franck Michel. A systematic approach to identify the information captured by Knowledge Graph Embeddings. In *IEEE/WIC/ACM International Joint Conference on Web Intelligence and Intelligent Agent Technology (WI-IAT)*, 2021.
- [2] Antonia Ettore, Anna Bobasheva, Franck Michel, and Catherine Faron. Stunning Doodle : a Tool for Joint Visualization and Analysis of Knowledge Graphs and Graph Embeddings. In *ESWC 2022*, 2022.
- [3] Rex Ying, Dylan Bourgeois, Jiaxuan You, Marinka Zitnik, and Jure Leskovec. GNNExplainer : Generating explanations for graph neural networks. *Advances in neural information processing systems*, 32 :9240, 2019.