

Cartographier les communautés de Twitch par la data

BOUCHAIB, NICOLAS, nicolas.bouchaib@first-link.fr, First Link (Data Scientist), Bordeaux

BAHS, NATANIEL, nataniel.bahs@first-link.fr, First Link (CEO), Bordeaux

Thématique : Analyse des données.

Résumé : Les réseaux sociaux ont développé depuis leur création des dynamiques communautaires en ligne. Ces communautés, définies par leurs usages des plateformes, sont identifiables grâce à la Data Science. First Link a développé en 2021 un nouvel outil de datavisualisation en cartographie qui permet de mettre en image ces dynamiques communautaires sur la plateforme Twitch.

Mots clés : Datavisualisation, Twitch, Communautés, Audiences, Data Analyse, Influence.

1. Introduction

A l'heure du web social, comment comprendre les dynamiques communautaires sur les médias sociaux ? Depuis le premier confinement de Mars 2020, la plateforme de streaming en direct Twitch a vu son nombre d'utilisateurs exploser. Sur ce réseau hautement social, les communautés se forment au fil des jeux et des streamers à succès. En mobilisant sa Suite Sight Web, qui permet de cartographier le web pour en comprendre le fonctionnement, l'Agence First Link a développé un outil inédit d'analyse des flux d'audiences sur la plateforme.

2. Méthodologie

Pour analyser les dynamiques communautaires, nous avons développé en Python un outil d'extraction qui permet de récolter les données issues de Twitch via l'API. Optimisé via du multithreading, cet outil nous permet de scrapper et parser en continu les audiences du Top 100 Twitch afin de former des datasets contenant plusieurs milliards de variables.

Nous envoyons ensuite ces données dans la suite Sight Web, package d'outils propriétaires de l'agence qui permet de nettoyer, uniformiser puis cartographier les données.

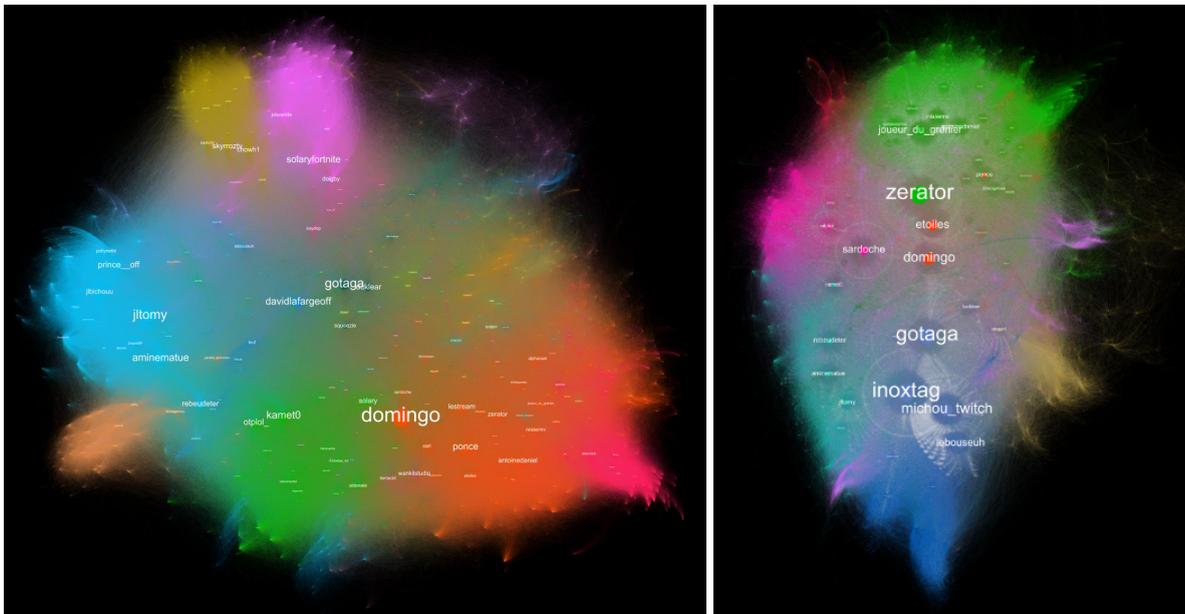
Enfin, nous utilisons le logiciel open source Gephi, qui permet de calculer des indicateurs issus de la théorie des graphes, clusteriser les communautés, spatialiser les données puis d'exporter des rendus de l'ordre de plusieurs GigaPixels.

3. Originalité / perspective

Cette cartographie des dynamiques communautaires est originale à différents niveaux :

- *Son ampleur* : les cartographies sont une datavisualisation de millions de variables à la fois (~2 millions de points, dizaine de millions de liens). Bien que des travaux du même genre aient déjà été réalisés, jamais avec l'ensemble des viewers sur la visualisation.
- *Capacité d'insights* : la cartographie permet de comprendre et de justifier les dynamiques communautaires. Cette connaissance est riche en insights pour les vidéastes, les équipes marketing, les chercheurs en sciences humaines, en analyse de signaux...

Actuellement, les données traitées relèvent du Top 100 Twitch FR. Notre objectif à court terme est d'optimiser les process pour décliner ce genre d'analyse à différentes échelles, notamment sur le Top 100 Monde.



Références

- <https://techno.konbini.com/fr/gaming/z-event-cette-incroyable-infographie-illustre-les-constellations-de-streamers/>
- https://youtu.be/N61_kHXpaFA
- <https://www.cbnews.fr/mobile/image-first-link-analyse-comportement-2000-streamers-twitch-francais-61690>
- <https://www.placeco.fr/article/article/de-twitch-a-la-french-tech-first-link-cartographie-les-reseaux-sociaux-868>
- <https://www.ladn.eu/media-mutants/reseaux-sociaux/lecons-importantes-carte-twitch-francais/>
- <https://www.larevuedudigital.com/le-media-social-twitch-cartographie-par-une-jeune-agence-de-data-intelligence/>