

Chaire PcEn Pluralisme culturel et Éthique du numérique

1. A method for measuring content prominence on Netflix
2. Algorithmic panic: European contents on the Netflix homepage

Director:

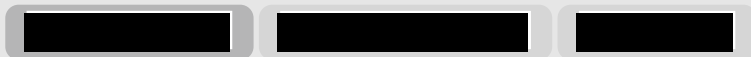


For more information, please visit pcen.fr

21/01/2022

A method for measuring content prominence on Netflix

Detailed walkthrough of a novel method using single-board computers to simulate human interaction, permitting *ceteris paribus* analysis of prominence strategies.



Partners



At the Chaire PcEn, we have placed content prominence at the heart of our research program on cultural diversity. Whereas traditional approaches to measuring cultural diversity involved a fixed and limited number of works, we adopt the view that streaming services today give access to such a large selection of contents that those established methods are no longer relevant. Indeed, what is the use of measuring diversity in a catalogue that no one user ever sees in its entirety? Surely the more relevant space within which to study cultural diversity is the user's screen and the subset of works from the catalogue that appear on it.

The present document outlines a new method for measuring content prominence on Netflix in a controlled environment. Using bots to mimic human behaviour on the platform, the researcher retains total control over what information is sent to the service and is able to automatically retrieve data pertaining to individual contents' placement on the service.

Context

Studies that look at the make-up of streaming services' catalogues are commonplace, and researchers tend to draw conclusions about these services' content strategies based on the availability or unavailability of certain works. While this approach does have its merits (bird's eye view of a content offer, insight into an acquisitions strategy...), it also overlooks a crucial aspect of non-linear content offers: the salience of some types of contents over others on the user's screen.

Prominence on a linear television schedule is governed only by time; the most prominent programs are those that are broadcast during peak-audience time slots. However, in a non-linear context, prominence becomes a multi-faceted concept. Prominence in space simply corresponds to the position of each title on the page, the most prominent being those placed in the most noticeable spots. If we want to identify the most prominent titles over a period of time, then we must look at the length of time each title has stayed in a prominent area. Finally, in the case of personalised VoD services such as Netflix or Amazon Prime Video, prominent titles can vary based on each user's past behaviour. This means that prominence on those services can only be inferred on a per-user basis.

In short, prominence as understood in a linear television context answers only to the question “when?”, whereas non-linear prominence is concerned with “where?”, “when?” and “for whom?”.

Summary explanation

For small-scale observations (one or few users during a short period of time), it is possible to keep a manual log of the prominent contents that appear on-screen. However, large-scale analyses require some form of automation in the data collection process. What’s more, VoD services typically do not give access to that kind of data, if any at all.

Our method for measuring content prominence involves computer-simulated web browsing. We use a script library called Watir for the Ruby programming language, which enables us to automate the browsing process by scheduling interactions with the browser (clicking on a button, scrolling a page etc.) and thus simulate human behaviour. When applied to a VoD service interface, automated browsing can be used to plan viewing sessions by scheduling login and logout times and selecting what content is to be “watched” and for how long. In essence, a Watir script browsing Netflix is a pre-programmed bot.

Theoretically-speaking, any number of bots can be created and run simultaneously, and each bot can be assigned a specific viewing schedule. We can also instruct each bot to store any relevant data, such as the contents of the homepage, in a shared database.

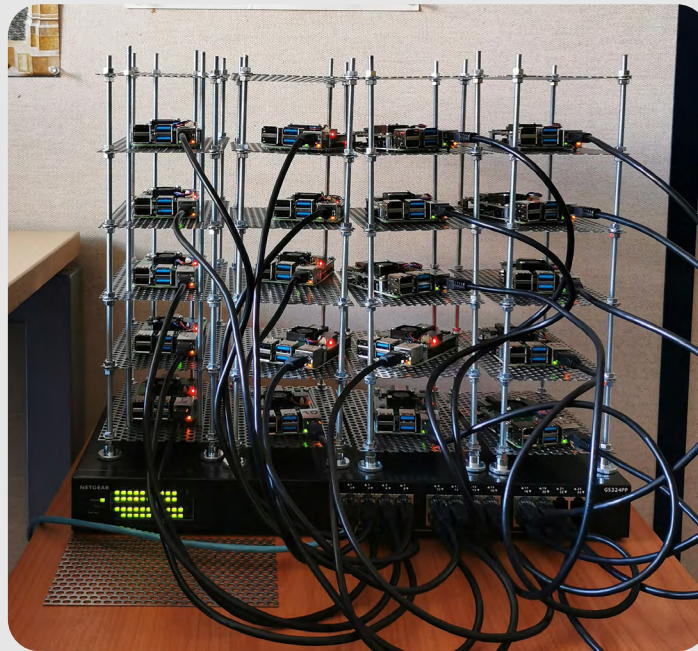
In order to implement this method on a scale where enough observations can be made, we rely on a custom-built cluster of Raspberry Pi Model 4B units.¹ A [Raspberry Pi](#) is a small and compact single-board computer that is ideal for completing simple, repetitive tasks. Each “Pi” is responsible for executing the commands of a single bot.

In practice, we use our Pi cluster to test certain hypotheses. For example, if we were to test the hypothesis that European content is prominently placed on the Netflix homepage, we could instruct 10 of our bots to watch European contents for 3 hours daily over a week, and another 10 of our bots to watch non-European contents at the same rate. Each day, we would collect the homepage from every bot’s Netflix profile and insert it into a shared database. When the

¹ There are two obstacles to conducting our experiment on the cloud using virtual machines instead of our own physical devices. The first is Netflix blocking IP addresses connected to data centres and the second is expense-related, as renting 20 server instances equipped with graphics cards would in the long run become far more costly than building our own cluster.

experiment is over, we can compare the make-up of every homepage, right down to the placement of every single title, and observe the prominence of European contents, all other things being equal.

Thus being able to control every aspect of a bot's behaviour on the platform is crucial to answering precise research questions that require *ceteris paribus* conditions.



Our cluster of Raspberry Pi's

Technical explanation

The following is a more in-depth walkthrough of our methodology and of the technologies it relies on.

Pi cluster

Our cluster of 20 Raspberry Pi's is powered via a PoE+ (power over ethernet) ethernet switch that provides both power and an internet connection to each device. The Pi's are fitted with a PoE+ hat enabling them to be powered via an ethernet cable. An SD card on each Pi is configured with a custom version of Raspberry Pi OS with specific settings (language, timezone etc.), pre-installed code libraries (Ruby, Watir etc.) and unique credentials (device name and ID, Netflix account and IoT hub credentials). On boot, a *systemd* service automatically executes a series of commands: pull the latest version

of our code from our GitHub repo and run a Ruby script that connects to our Scaleway-hosted MQTT broker and awaits further instructions.

MQTT and MQTT broker

Our Raspberry Pi's are at the heart of the system, as they execute the main function of our setup: browsing Netflix and "watching" content on the service. However, communication to and from the Pi's is also crucial. They need to be told exactly how to behave and where to store data. For this, we rely on an IoT (Internet of Things) messaging protocol called MQTT. The MQTT protocol allows us to quickly send messages as packets of data to and from the Pi's. All messages pass through an MQTT broker that acts as a post office, directing each published message to its intended recipient. We use a cloud-hosted MQTT broker provided by our partner [Scaleway](https://www.scaleway.com/en/iot-hub/), which allows us to publish and receive messages from anywhere.² What's more, this cloud-hosted MQTT broker makes it possible for us to communicate with our Pi's from a webpage, which brings us to our next essential component: the monitoring interface.

² For more info, see: <https://www.scaleway.com/en/iot-hub/>

Monitoring interface

We built a web-based monitoring interface from scratch to act as a kind of control tower from which we can oversee our whole fleet of Pi's. Using Javascript's "MQTT over websockets" capabilities, we can connect our webpage to our MQTT broker and use it to control and monitor all devices.

With the MQTT protocol, information is transferred on topics, to which a device can subscribe and publish on. In our case, each device subscribes to its relevant MQTT topics and, in turn, each section of our monitoring interface publishes and receives messages on those same topics.

Upon opening the webpage, a connection is established with the MQTT broker. Each Pi has its own dedicated section on the page, on which live information pertaining to its status, schedule and viewing sessions is displayed. Effectively, the monitoring interface reads and publishes MQTT messages on a variety of topics for each device: 'status', 'schedule', 'sessions', 'homepages', 'watched contents'...

The image shows a monitoring interface with two device profiles. The first profile, 'pcen-001', is in a 'running' state. It has a profile name 'wTpvCV5l', an agenda '81439253,70035036,81278699', a description 'Japanese', and is set to run every day at 18:00 for 3 hours, 30 days. It started on 20/01/2022 at 21:01. The message log shows a retained message from 20/01/2022 at 21:22 with ID 74220.717440732. The second profile, 'pcen-002', is partially visible and shows similar fields.

The flow of data

1. Content database

Our starting point is a MySQL database containing the entirety of the Netflix catalogue in France, along with associated metadata sourced from IMDb. This database is the result of monthly parsings from the platform itself. Each title is matched with its corresponding page on IMDb using a custom-built algorithm.

2. Scheduling

In order to obtain a list of titles that a bot will be assigned to watch, we query the database specifying our intended filters. For example, we can pull out a list of horror series produced between 2000 and 2010. We then paste this list into the 'Agenda' box of a device's section on the monitoring webpage. Additional options allow us to set the frequency and duration of the bot's viewing sessions. Finally,

hitting the ‘Send schedule’ button triggers the publishing of several MQTT messages containing the schedule as specified above, which the device receives, “unpacks” and integrates into its script. Once the schedule is received, a ‘Start’ button appears which, when pressed, sends an MQTT message to the device telling it to start the experiment.

3. On the Pi's

The script on each Pi is written in the Ruby programming language, and uses several Ruby gems: ‘Watir Webdriver’ to automate the browsing process, ‘Rufus scheduler’ to schedule tasks, and the ‘MQTT’ gem to establish a connection to our MQTT broker and thus enable our code to publish and read MQTT messages. At the start of the experiment, each Pi is instructed to create a new profile on its Netflix account and immediately parse the homepage. This is, in a way, a “virgin” homepage because the recommender system does not yet know the user’s preferences. Once this first homepage parsing is complete, the script pauses while it waits for its first scheduled viewing session.

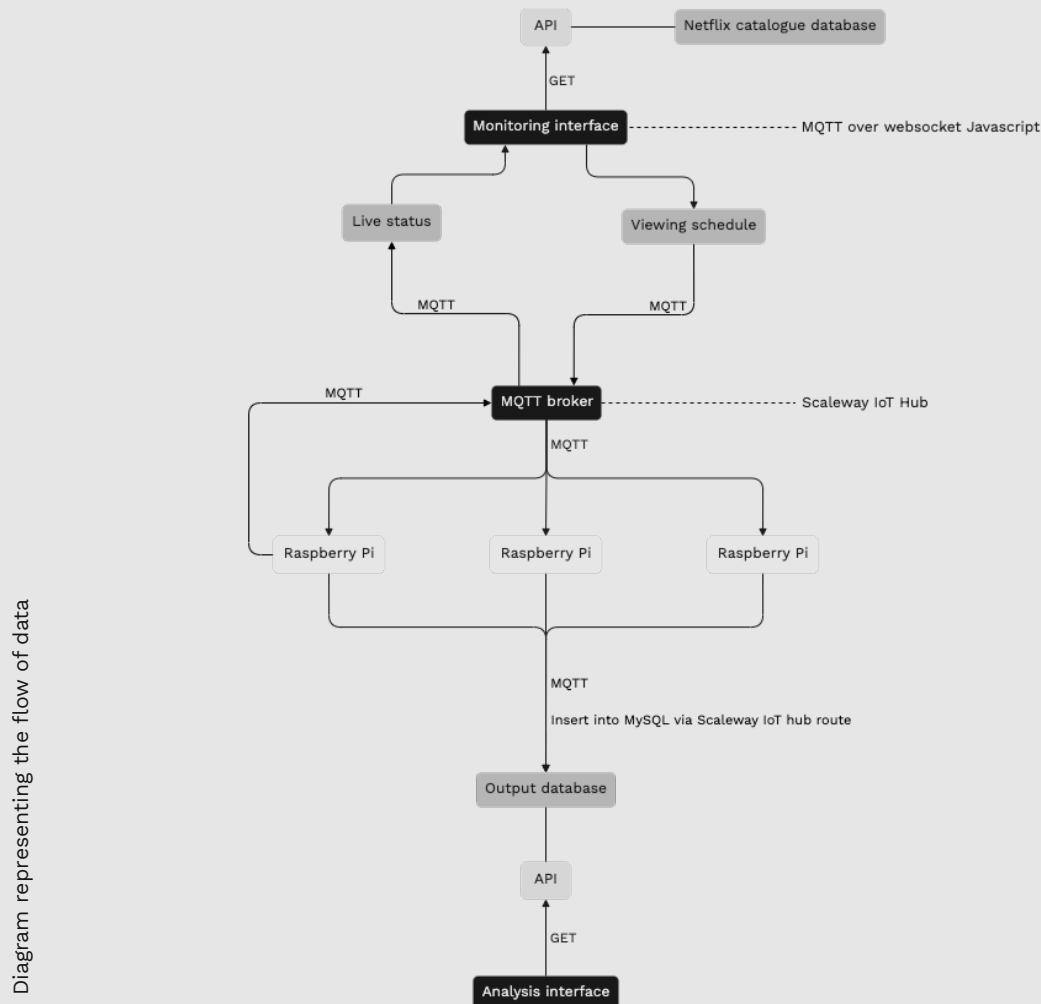
Thanks to the Widevine DRM library available since March 2021 on the Raspberry Pi, our Chromium browser can play DRM-protected videos such as those on Netflix. On top of that, the videos can be played in headless mode, meaning that no screen is actually needed.

Each viewing session has a fixed duration and is followed by a homepage parsing. The next day, each bot simply continues to watch content starting from where it left off the day before.

4. Output database

At the end of each homepage parsing, the data is sent via MQTT, in JSON format, to the broker. When the broker receives a message on the ‘homepages’ topic, a Scaleway IoT Hub route is triggered. An IoT Hub route forwards an MQTT message from the broker to a specified destination. Here, the destination is a MySQL database. In the process, a stored procedure consisting of a number of MySQL queries distribute the contents of the received JSON file to the various tables in the database, while establishing relational connections between them. This “output database” is built in such a way that homepages can be easily reconstructed and compared during analysis.

All the while, regular status updates are sent from the Pi's to the MQTT broker and displayed on the web-hosted monitoring interface.

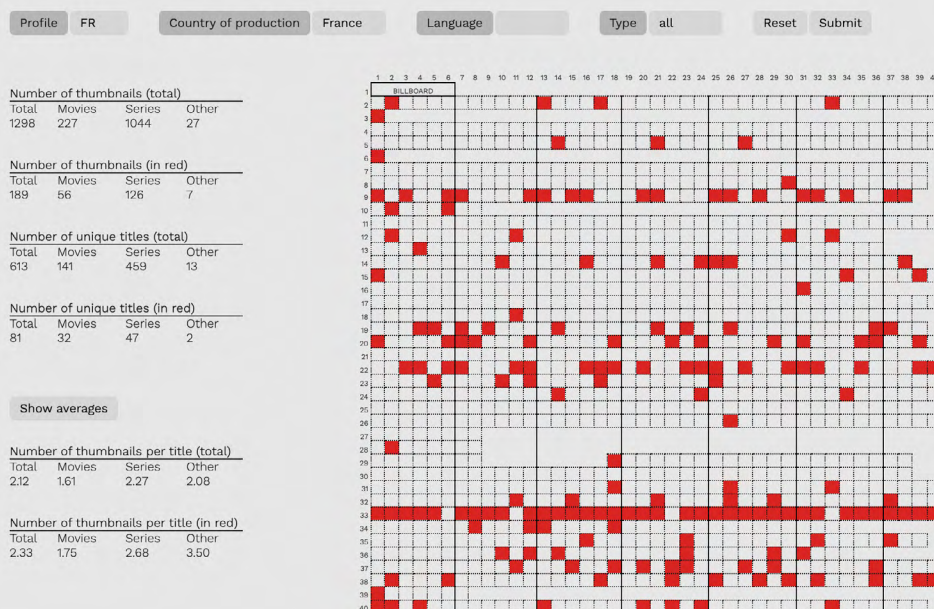


Analysis interface

Once our experiment is over and the data is stored in the output database, we can begin to assess whether our starting hypothesis is confirmed. For this, we can simply run MySQL queries directly on our output database. However, this method can be counterproductive for two reasons. First, the sheer volume of data (one homepage a day for 20 profiles over 10 days amounts to 200 homepages, each with approximately 1400 thumbnails) means that it is difficult to get a sense of the “big picture”. Second, our output data is inherently visual, as it reflects the visual hierarchy of each homepage that we parsed. Conducting analysis on data pertaining to the visual

characteristics of the user interface without some kind of visual representation can indeed be arduous.

This is why we set about building an analytical UI that would allow us to quickly visualise the composition of each homepage. Based on a 40×40 grid representing the entirety of the columns and rows of a typical Netflix homepage, the interface also displays several key metrics, such as the total number of unique titles, the shares of films and series etc. A collection of filters means we can highlight the contents of our choosing on any homepage, and track their presence across several days with the aid of a slider.



Snapshot of the analysis interface (under construction)

After watching only French series for 2 days, there are 189 thumbnails on the homepage advertising content produced by France.

1 day = 3 hours of viewing time from 5pm to 8pm. Days are consecutive.

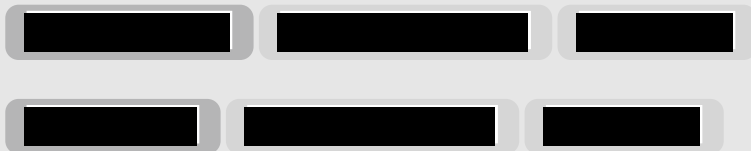
Applicability to other services

Looking beyond our experimental setup's focus on Netflix, its applicability to other services is achievable with only limited effort. The technologies and processes that form the backbone of our methodology on Netflix can be reused, while parts of the monitoring interface, the scripts on each Pi and the output database would have to be reworked. Overall, the structure of the experiment would remain the same. Hence in theory, content prominence strategies can also be studied in various industries, such as music (Spotify, Apple Music...), video (YouTube, TikTok...), gaming or live streaming (Twitch).

21/01/2022

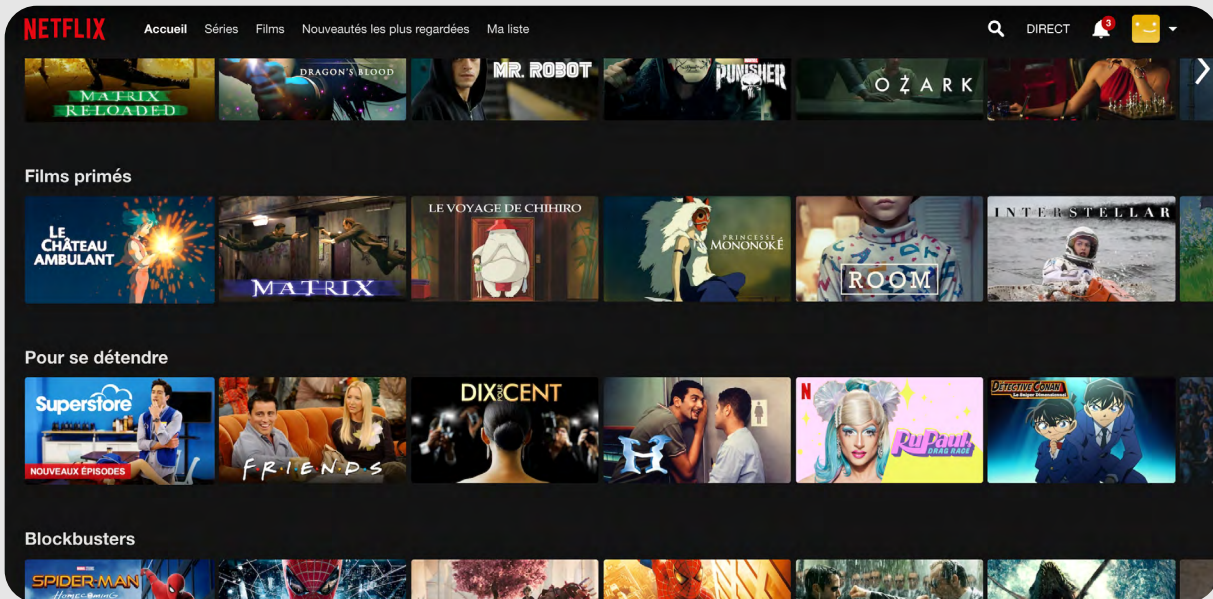
Algorithmic panic: European contents on the Netflix homepage

20 bots watched Netflix every day for 8 days, here is what their homepages tell us about the prominence of European works on the platform.



Partners





Screenshot of a Netflix homepage

Is Netflix doing enough to promote European contents to its European users? If a user never watches any European films or TV series, will European works simply disappear from the homepage? Doesn't algorithmic personalisation mean that filter bubbles will trap the user in his or her own tastes?

For the past couple of years, these questions have been repeatedly asked by a number of stakeholders in the European audiovisual industry, from regulators to producers, to government officials and analysts, bearing witness to a certain scepticism toward Netflix's recommender system. More than a decade since Pariser's blockbuster book¹ was published, it is still often taken for granted that all recommender systems create filter bubbles and, in doing so, constitute a threat to European contents' prominence online. **But are fears of disappearance from the homepage justified or are they unfounded?** Interestingly, no quantitative study has ever been carried out to answer that question and, despite a real lack of data pertaining to prominence on VOD services, the perceived threat that algorithms pose is enough to send professionals, regulators and decision-makers alike into what we might call an "algorithmic panic," prompting them to take steps to solve a "problem" that has not even been confirmed using scientific methods.

¹ Pariser, E. (2011) The Filter Bubble: What the Internet Is Hiding from You. London: Viking/Penguin Press.

Methodology

As part of our ongoing effort to understand the mechanisms that underpin content prominence online, we have developed a robust methodology² that allows us to study the effects of viewing behaviour on the placement of content on the Netflix homepage. We employ it here to survey the placement of European works with regard to past viewing behaviour. To put it briefly, our method consists of 20 Raspberry Pi's (tiny single-board computers), each provided with a different newly created Netflix account, watching 3 hours of content and collecting the position of every title on the homepage each day. The bots' actions are pre-programmed and there is no human intervention during the whole process.

For the purposes of this experiment, 10 bots watched only European contents and the 10 remaining bots watched non-European contents, films and TV series alike. Before the start of the experiment, each bot was given a list of titles to watch, selected at random from the whole Netflix catalogue with the only criteria being their European³ nature. Our goal in setting up this experiment was to remove as many biases as possible. The fact that the creation of a new profile, the parsing of the homepage and the viewing of contents are all happening simultaneously for all bots ensures that our results are not skewed by time-based discrepancies. What's more, the random selection of titles eliminates selection bias, although it does result in extremely atypical viewing patterns from our bots. Finally, all bots are housed in the same "cluster" (see photo below), meaning that they share the same IP address, thus eliminating any location bias.

The experiment lasted for 8 days, which we felt was an appropriate length of time from which to gain valuable insights. As mentioned above, each bot watched 3 hours of content each day, from 6pm to 9pm, and from October 12th to October 19th 2021.

² For a detailed explanation of the methodology, please refer to the document "A method for measuring content prominence on Netflix", available online on pcen.fr or as a pdf upon request.

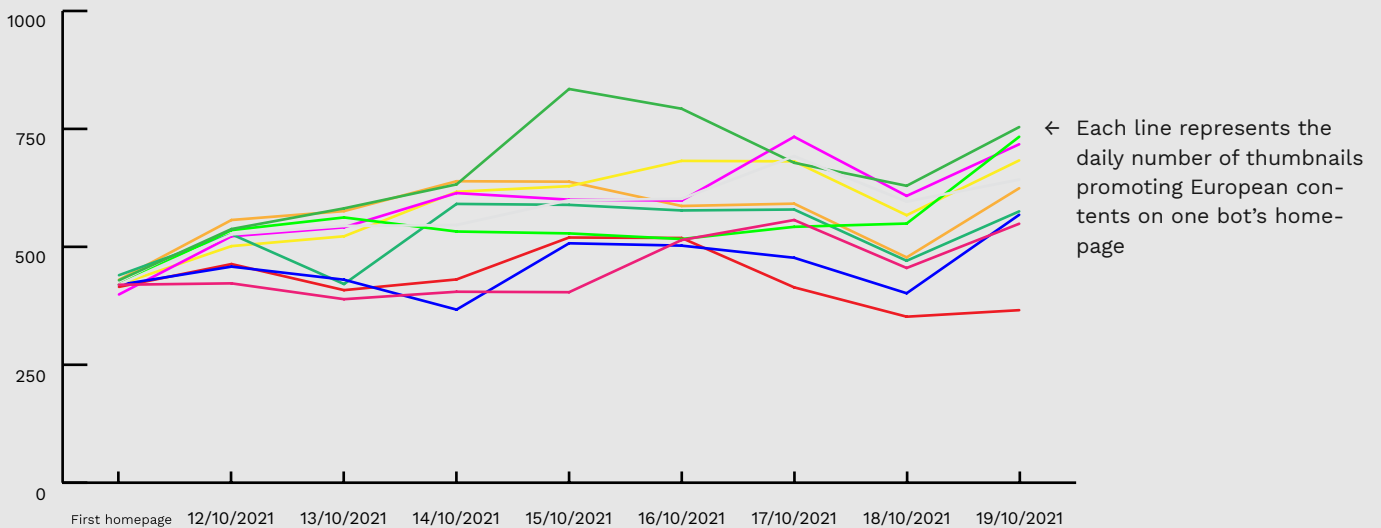
³ As per the signatories of the Television Without Borders agreement and determined by metadata collected from IMDb



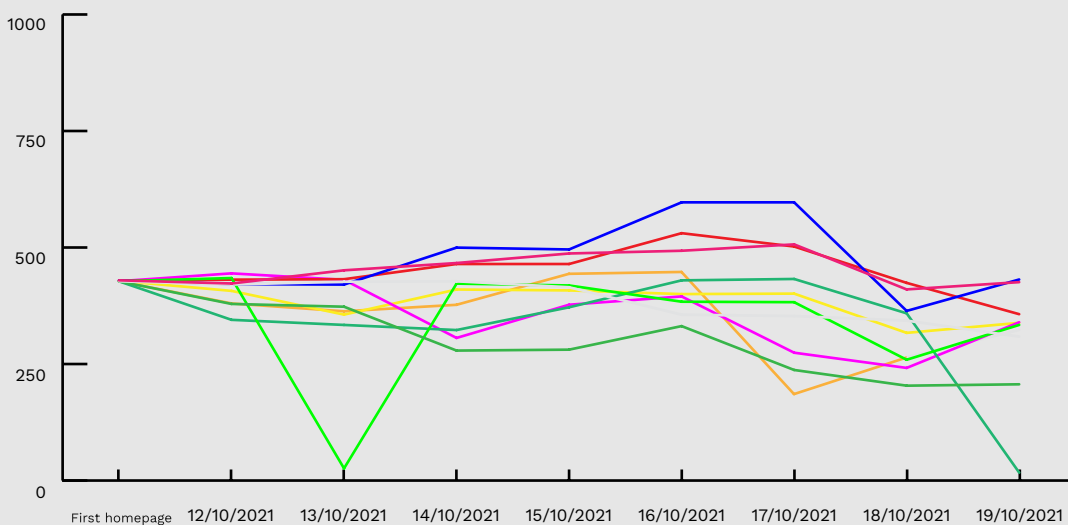
Our Raspberry Pi cluster

Results

Number of thumbnails promoting European contents on the homepages of bots that watched only European contents



Number of thumbnails promoting European contents on the homepages of bots that watched only non-European contents

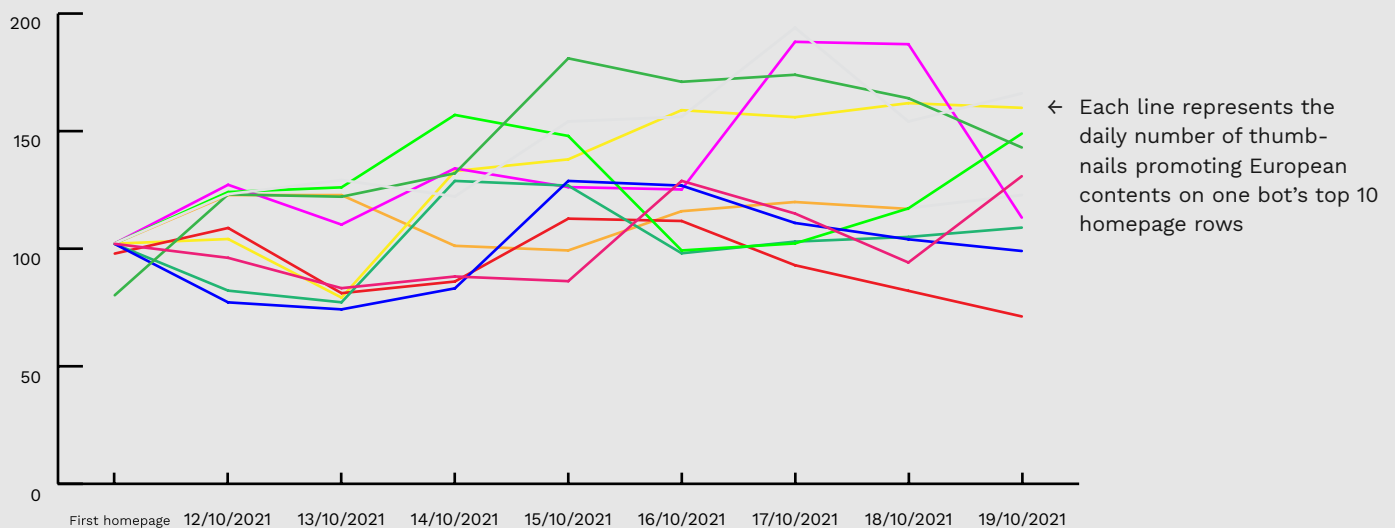


The graphs above show the daily number of thumbnails advertising European titles on each bot's homepage. Note that the two sudden drops on the second graph correspond to significantly smaller homepages (3 and 2 rows respectively), a rare occurrence due not to a bug in our code but to as-of-yet unexplained behaviour displayed by the recommender system.⁴

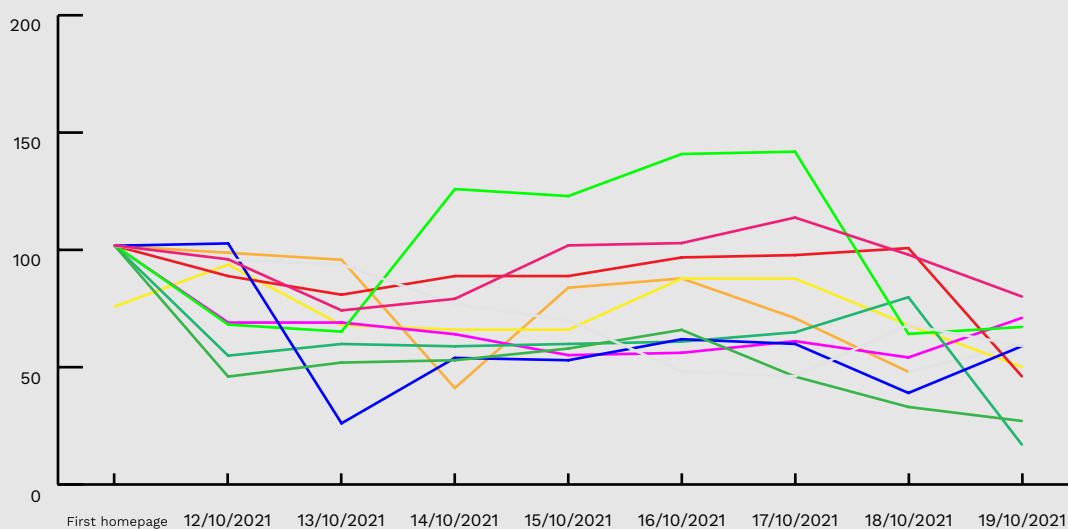
⁴ One can surmise that the randomness in the bot's selection of titles confounds the algorithm somewhat, and renders it incapable of displaying a full page of recommendations, although this is only an unverified hypothesis.

A cursory glance at both graphs reveals a slight upward trend for bots that watched only European contents, and a corresponding downward trend for those bots that watched only non-European contents. Given the extreme nature of both sets of bots' viewing habits, the scale and speed of the evolution towards more or fewer European contents on the homepage seem moderate. The algorithm does respond, in most cases, to a user's attitude toward European contents: **watching only European works on Netflix tends to result in an increase in the number of European thumbnails on the homepage and vice-versa.**

Number of thumbnails promoting European contents on the top 10 rows of the homepage (billboard included), for bots that watched only European contents



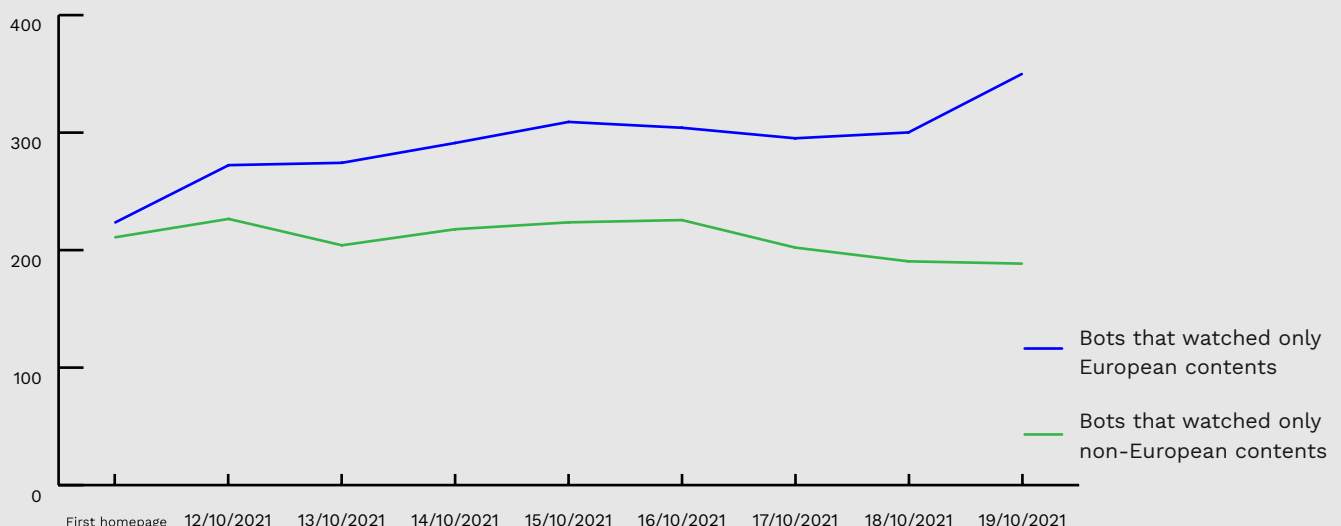
Number of thumbnails promoting European contents on the top 10 rows of the homepage (billboard included), for bots that watched only non-European contents



When it comes to prominence on the homepage itself, similar patterns emerge. For the two graphs above, we chose to look at only the top 10 rows on the homepage.⁵ As with the homepage as a whole, bots that have watched only European contents tend to see an increase in the number of thumbnails promoting European works within the top 10 rows. With a few exceptions, the inverse is also true. Again, for those bots that have not watched any European contents, we can clearly see that they have not disappeared from the top of the homepage. It is important to also bear in mind that the top 10 rows include non-recommended rows such as “Continue watching for...” and the country’s “Top 10 in...”, which feature titles that have not been selected by the recommender system.

⁵ An arbitrary choice, which may need some fine-tuning in the future. For reference, on average, the top 10 rows contain between 250 and 300 thumbnails.

Number of thumbnails promoting European contents on the top 10 rows of the homepage (billboard included)



Shifting our focus from the number of thumbnails to the number of different titles on the homepage⁶ reveals that the upward trend caused by watching European content does not coincide with a corresponding downward trend when the inverse is true. For bots that have watched only European works, the average number of European titles on the homepage gradually increases, reaching 350 on the last day of the experiment. This number remains more or less steady however for our “non-European” bots.

⁶ Bear in mind that each title may—and often does—appear multiple times on the homepage, in different thumbnails.

Number of thumbnails per country of production, all homepages combined, descending order

| Sessions → | All | Sessions → | European contents | Sessions → | Non-European contents |
|-------------|---------------|-------------|-------------------|-------------|-----------------------|
| Country | Nb thumbnails | Country | Nb thumbnails | Country | Nb thumbnails |
| USA | 124 063 | USA | 59 997 | USA | 64 066 |
| France | 32 229 | France | 19 190 | France | 13 039 |
| UK | 25 330 | UK | 13 559 | UK | 11 771 |
| Japan | 17 490 | Japan | 6 750 | Japan | 10 740 |
| Canada | 13 237 | Canada | 6 317 | Canada | 6 920 |
| Spain | 8 338 | Spain | 4 595 | South Korea | 4 107 |
| Mexico | 7 749 | Germany | 4 476 | Mexico | 4 068 |
| Germany | 7 719 | Mexico | 3 681 | Spain | 3 743 |
| South Korea | 7 549 | South Korea | 3 442 | Germany | 3 243 |
| Australia | 4 937 | Italy | 3 317 | India | 2 651 |

As an aside, and moving away from a strictly European viewpoint, we looked at the countries whose productions were featured the most on the homepage⁷ and, rather surprisingly, the top 5 are the same for both categories of bots. **France, where the experiment took place, has the second highest overall number of thumbnails promoting French productions,**⁸ albeit far behind the US, the leader by a factor of around four. Our methodology does not allow us to determine whether this is the result of high demand for French content or an editorial choice on Netflix's part. Japanese productions occupy fourth place, probably due in no small part to the popularity of Japanese anime on the platform. Mexico is a rather unexpected entry at seventh place.

⁷ Again, using IMDb as our source.

⁸ No distinction was made here between co-productions with other countries and single-country productions.

Over the course of our 8-day experiment, **92% of all European titles available on Netflix have appeared at least once on one of the bots' homepages.** What this means is that the vast majority of European works on Netflix have been "pushed," in varying degrees, to the homepage at least once. Some have been featured multiple times in the top rows, others only once in the bottom rows, but only very few have never appeared on the homepage at all. The scale and speed at which the homepage evolves in response to a user's viewing behaviour is such that, overall, there does not seem to be many "dormant" titles buried deep within the catalogue and never surfacing to users' screens.

Conclusion

Our findings indicate that European works do not suddenly disappear from the homepage, even when a user only watches non-European titles on Netflix. It would seem then that the fear of “invisibility by way of the algorithm”, expressed by many European professionals and regulators, is unfounded. Objectively, there appears to be no cause for concern, and so we might begin to examine the reasons for such a widespread distrust of Netflix’s recommender system. Why then were these fears relayed by the media and some professionals?

We put forward two hypotheses, each warranting further discussion:

- **A lack of confidence in the attractiveness of European works.** If all contents were to compete for attention on an equal footing, what would make European contents more attractive than others? After decades of US hegemony at the box office and in TV ratings, European professionals tend to assume that audiences will always favour foreign over local content. When applied to recommender systems which, we are told, personalise recommendations based on past viewing behaviour, this line of thinking can only come to a rather defeatist conclusion: if nothing is done, our unattractive European works will simply never be recommended. Hopefully, the recent proliferation of European success stories on streaming services will start to chip away at that idea, and European stakeholders will begin to view recommender systems in another light.
- **The fear of a loss of cultural sovereignty.** European lawmakers like to ensure a minimum degree of exposure to European works on every media format, and they do so for different, more or less explicit reasons: to promote so-called cultural diversity, to protect otherwise endangered sectors of industry and to safeguard against perceived American cultural imperialism. All of these endeavours require a certain level of control, typically exerted through the imposition of financing obligations and broadcast quotas. However, the necessary relinquishing of control that recommender systems entail makes regulation pertaining to content prominence inadequate and difficult to enforce. Viewed in this light, the fear of seeing European works disappear from the homepage of a service like Netflix stems not from an actual, witnessed scenario but from what *could* in theory happen when control over prominence is transferred to an unknown, impersonal

entity: the algorithm. Left to its own devices, it is presumed that the algorithm will act against the European Union's best interests, leaving no room on the homepage for local content. Moreover, the fact that the algorithm was built in the US, Europe's long-time "cultural aggressor," only serves to reinforce the prejudice.

Whatever the reason for Europeans' scepticism toward cultural recommender systems, there is a pressing need for actionable data from which to confirm or alleviate those concerns. Our methodology aims to do just that, and has shown itself to be reliable enough to provide quantitative insights into content prominence in the context of an algorithmically generated homepage. We will continue to put it to use in future experiments, where we will explore other variables such as release dates, languages and genres.

Thank you



arte



NETFLIX



Contact

