Has the peak attention been reached? Cross-media analysis of competition for attention

Wojciech Hardy (Faculty of Economic Sciences, University of Warsaw)

Weronika Motkowska (Faculty of Economic Sciences, University of Warsaw)

Preliminary version (May 2024) – please do not cite

Abstract

With the rapid digitization of the creative sectors, we have entered an era of abundance in terms of the supply of creative products. These new trends revived the interest in attention economics positing that an overabundance of information may lead to attention scarcity – consequently affecting consumer decisions. We test these assumptions in a market for attention, whereas different creative sectors compete for a finite supply of spotlight. In line with attention economics literature and debate, we hypothesise that attention has been increasingly scarce over the past decade (or the competition fiercer), but that the onset of COVID-19 might have relaxed this state to a point (due to an average increase in leisure time). We test this by constructing a novel country-level, monthly dataset on attention and time devoted to content in film, TV and games and analysing the relationship between these sectors. Consistent with the theory, we find an increasingly negative relationship between attention paid to different sectors that has somewhat relaxed during the pandemic.

Introduction

The XXI century has seen an explosion in the supply of entertainment content that went in pair with its accessibility. Digitisation greatly lowered the costs of production, distribution and, consequently, of access. However, the abundance of content has put a strain on a so far underappreciated consumer resource - attention. For the entertainment sector, this means that the competition for consumers escaped the boundaries of single subsectors and instead pitted all types of creative content against each other. This paradigm shift has been perhaps best expressed by the Netflix CEO in a 2019 letter to shareholders: *"We compete with (and lose to) Fortnite more than HBO".* Consequently, some industry experts claim that we have reached or are near reaching the point of peak attention – i.e. a moment when no attention can be given to particular content without taking it from another (Mulligan, 2019).

Yet, to the best of our knowledge, the notion of sectors competing for a limited resource of attention has not been empirically tested. Instead, most empirical work focuses on what determines the success and popularity of creative products within individual sectors – without consideration for a broader market competition for attention, which could constrain spending on any specific content.

To fill in this gap, we construct a novel dataset of attention measures for a set of creative sectors, covering more than 20 countries, with monthly data between 2015 and 2023. First, we collect sector-specific measures. For film and TV, we identify the most popular global and regional titles from all periods. We find their unique IDs from Wikidata and use that information to collect country-specific Google Trends data, which we standardize and rescale for a comparable set of indicators of attention. For video games, we collect monthly player counts for all PC games in the Steam catalogue from Steamcharts. We then merge this with information from Gamalytic to infer country shares among the players. In the final version of the paper, we will also include information on eSport viewership. Notably, we specifically focus on types of content that do not allow for multitasking (e.g. listening to music while cooking or driving) and instead require full attention. In the end, our data includes an attention measure that is: a) consistent over time; b) can proxy for the amount of time devoted to a particular type of content; c) is available at high frequency and for many countries.

The variation of the measures differs visibly between sectors. Our strategy thus is to base our analysis on the between-sector relationships of within-sector changes. In other words, we standardise the collected measures for each sector and country. In that way, our analysis picks up whether deviations in any sector relate to deviations in the other sectors. For some sectors, these measures will be comprehensive and derived from all meaningful titles, while in some sectors they will be based on subsets of the most popular titles that capture industry-level trends.

We conduct our analysis using econometric modelling and include seasonal and country-level fixed effects (in the panel approach). We analyse the intensity of the relationships across different subperiods and relate them to the key technological developments over the same period. We also analyse whether the pandemic's start had an impact on the observed patterns.

1. Literature review

The digital era brought about a growing supply of creative content, including that of high quality. Waldfogel (2017, 2018) documents the growing production of new titles across all creative industries, from film to music to books and games. He documents not only an abundance of choices but also that these choices do not overcrowd the markets. Instead, users consume an increasingly diverse selection of titles, with the overall quality of consumed content increasing (or at least not decreasing). With these trends, market players increasingly focus on facilitating product discovery, to allow sifting through the growing catalogues.

With the amount of content rapidly growing, curation and recommendation strategies have become important. The shift from physical stores to online marketplaces and platforms allowed for all of them, however niche, to enter distribution. This, however, also meant an increased challenge for selecting high-quality content to consume. Thus, consumers now place additional value on recommendation algorithms that help sift through the long catalogues (Sinha and Mandel, 2008; Aridor et al., 2022). Conversely, Brynjolfsson et al. (2023) highlight the value of time in the general valuation of free online goods.

Despite the support of recommendation systems and curators, finding time for all the released content has become increasingly difficult. More than half a century ago, Simon (1971) noted that "a wealth of information creates a poverty of attention and a need to allocate the attention effectively among the overabundance of information sources that might consume it" (pp. 40-41). This concept understandably received new traction in the internet era. Webster (2016) wrote about the "marketplace of attention", highlighting the growing importance of targeting and personalisation as ways of achieving success in the new media landscape.

Indeed, some media experts have already concluded that we have already reached "peak attention" and that no time is left for additional consumption (Mulligan, 2019). Purdy and Reznik (2019) additionally suggested that the ongoing competition for consumers' attention and engagement might and should (from a marketing point of view) extend to firms beyond the entertainment sectors. The authors argue that most firms should learn from the entertainment industries about the value of capturing and retaining consumer attention, in a way to form engagement and loyalty regarding their brand. This suggests that the competition for attention might get even fiercer.

Beyond the abundance of content, attention strain could also result from internal shifts in the creative sectors. Within the film and TV market, there has been a pronounced shift in the distribution of high-profile titles across different VOD services. For example, just between 2016 and 2019, the top 10 most popular (most sought-after) TV shows among IMDb users have shifted from mainly traditional TV shows to almost exclusively VOD or premium TV shows. Moreover, while the most attention-grabbing VOD shows came from 2-3 such services in 2015 and 2016, this number increased to 9 by November 2019.[1] Indeed, as shown by Nason (2019), an increasing share of US households subscribes to an increasing number of different video streaming services.

These trends contribute to what has been dubbed "subscription fatigue". Indeed, several reports show a growing frustration with subscription services, in terms of their number, their combined price, the way the content is increasingly spread, the overwhelming choice, increasing searching costs and difficult management of several subscriptions (Gilsenan, 2019; Westcott et al., 2019; Inman, 2019).

[1] The numbers are based on own analysis using data reported by IMDb and archived by Internet Archive Wayback Machine.

These findings go in line with the effects of attention scarcity or saturation, whereas the viewers find it increasingly difficult to manage their consumption behaviour.

Over the past two decades, other kinds of content have been taking increasing shares of users' time. The video game market has been rapidly growing, overshadowing other creative sectors. By 2022, its revenues have been estimated at $203B (Wijman, 2022), almost four times the combined revenues of the box office and music revenues (Loria, 2022; IFPI, 2022). In the same period, new types of digital content have been emerging, putting additional strain on the attention resources, including eSports and live-streaming channels, as well as social media platforms such as TikTok. With a growing and older video game player base and entirely new types of content, the general supply of time and attention is likely becoming an increasingly important factor influencing success.

One major change potentially relaxing these strains was the start of the COVID-19 pandemic. Aksoy et al. (2023) note the COVID-19 pandemic of the 2020s greatly altered daily schedules, permanently shifting some of the work home and contributing to increased leisure time. During that time the overall internet traffic greatly increased not only due to remote work (Feldmann et al., 2020). Live performances, festivals and theatres suffered great losses, while film production has been largely halted (Nhamo et al., 2020). Moreover, the reduced commuting and mobility during the pandemic reduced the demand for music streaming (Sim et al., 2020), while the demand for VOD streaming services soared to a point where streaming quality reductions had to be enforced across the EU. The gaming industry also noted unprecedented increases (that currently, after a large slowdown, contribute to massive layoffs across firms).

2. Hypotheses

Our aim in this study is to put the notion of the changing competition for attention to the test. In particular, owing to the changes induced by digitisation, we hypothesise a negative relationship between the attention ascribed to different creative sectors. In particular, when studying time trends, we suspect that deviations in the demand for a particular type of goods (e.g. due to blockbuster titles being released) can be related to lower interest in other types of goods, as the attention and time budgets decrease.

H1: Attention devoted to a creative sector is negatively related to the attention devoted to other creative sectors.

Studying prior trends in the creative sectors suggests that the competition for attention has only been growing larger, with the arrival of new types of content, the popularization of some others and the growing dispersion of high-profile content across services. In this line, we suspect that the relationship between attention paid to different creative sectors has been increasingly negative over the years.

H2: The relationship between attention devoted to the creative sector has been increasingly negative over the past decade.

Finally, as the COVID-19 pandemic reduced the production output and increased the time of leisure, we suspect that it has relaxed the attention strain induced by the competition of different creative sectors.

H3: The COVID-19 pandemic intensity reduced the negativity of the relationship between the attention given to different creative sectors.

3. Data and methods

To verify the hypotheses we focus on three types of creative goods: long-format audiovisual products (i.e. movies and TV series), PC video games and live-streaming. Three factors dictate our choice. First, we focus on types of content that require full attention for consumption, e.g. unlike music listening, which can be combined with other activities. Second, all three sectors are prominent and have been growing in recent years or changing in a way that could contribute to attention strain. Third, we can approximate attention or demand in each of these sectors with high (monthly) frequency and for numerous countries.

3.1. Film and TV data

To measure the level of attention devoted to films and TV series we follow the steps:

1) We identify a large set of highly popular titles covering the years 2015-2023.
2) We collect IDs related to these titles that are recognisable by Google as individual topics.
3) Using the IDs, we collect the Google Trends information for these titles for a set of countries, covering all months between 2015 and 2023.
4) We standardise the Google Trends information to achieve comparable values.

The title sample for relevant content is constructed by collecting the Top 100 titles in separate charts for movies and series according to their popularity on IMDb determined through the interest of website users. IMDb has been the most popular film and TV-devoted website, compiling a broad range of information on titles, actors, production studios, etc. Its Moviemeter and TVmeter top 100 lists provide a ranking of the titles currently most sought by IMDb user base. We use the Internet Archive's Wayback Machine[2] to access historic rankings and collect the titles of at least monthly frequency, starting from 2015 (the list of historic dates for the rankings, available through Wayback Machine, is provided in Appendix Table 1). In total, we collect 5157 titles in this way (Appendix Table 2). This sample contains a considerable amount of internationally popular and relevant movies and series.[3]

To measure the joint attention given to the titles in our sample across countries and dates, we use Google Trends data. The Google Trends tool provides high-frequency, country-specific information on the relative search volumes for specified keywords or topics. For film and TV it thus constitutes a reliable measure of attention. It has also been proven to provide valuable information correlated with the current level of activity in various sectors such as retail sales, automotive sales, home sales or travel (Varian & Choi 2009).

[2] https://web.archive.org/

[3] In the final paper, we plan to additionally collect the most popular regional productions to cover the attention and interest in titles that did not reach international popularity. We used the IMDb advanced search tool, which can be used to return titles released in specific periods and produced in specific countries and sort them according to their popularity in a given country. We thus identified the most popular local releases from countries in the sample, collecting at most 10 top series titles and 10 top movie titles from every month for any country. The collection of attention expressions for these titles is still ongoing.

Google Trends can provide search popularity for keywords and phrases, but also for particular topics that can be identified through unique IDs as defined by Freebase or Google Knowledge Graph (GKG) IDs. We take advantage of this by collecting the Google-recognised IDs of our movie and TV series titles from Wikidata and Google API. We do so by utilising the IMDb IDs from our initial title sample and retrieving relevant IDs recognisable by Google Search (Freebase and GKG). We then collect search interest data for all the titles and countries in our sample. This approach gives us two advantages over merely using titles as keywords for Google Trends. First, by relying solely on topic IDs, we can be sure that the search interest is not conflated with search terms for other topics with similar wording (e.g. a search for "witcher" could otherwise refer to a TV show, but also to books, video games and other media about the character). Second, we can easily identify the same titles across different countries, even if the titles were translated locally (e.g. the "Witcher" show is named "Wiedźmin" in Polish and "Zaklínač" in Czech, so the search term "Witcher" would only provide reliable results in some countries, while collecting popularity of the Witcher show object by its ID will provide consistent results for all countries).

Google Trends offers valuable insights into search trends, but it comes with certain limitations. Firstly, the data provided by Google Trends does not present actual search volumes but relative scales. Essentially, Google Trends divides the volume of a specific query by the total number of searches within a specified time frame and geographic region. This normalized volume is then set with 100 as the peak of searches, providing a comparative measure rather than exact figures. Secondly, there's a constraint on the number of queries that can be analyzed in a single request, with a maximum limit of five queries. This means that when using Google Trends, researchers can only directly compare the search volumes of up to five titles simultaneously.

We develop a framework that overcomes these limitations and provides us with comparable sets of measures. First, we collect the information for each title with the addition of a 'reference' title for the second time series. We then rescale the numbers for each title using the reference as units.[4] In the end, for all titles and countries, we end up with numbers that consistently reflect the same unit of measure through the reference to the benchmark title and country.

Finally, for each country, we standardize the values for a mean of 0 and a standard deviation of 1. Our Google Trends data units are not interpretable in any conventional way. Standardising them allows us to interpret the numbers in terms of changes in magnitudes comparable with the standard deviations of the attention devoted to films and TV products.

3.2. Video games

To measure the level of attention devoted to video games we follow the steps:

1) Identify all video games in the Steam catalogue.
2) Collect information on monthly player counts from Steamcharts.
3) Collect information on country share among players from Gamalytic.
4) Estimate the number of players from each country.
5) We standardise the players' information to achieve comparable values.

[4] In detail, for each title and country we set the Google Trends for the Breaking Bad series in the USA as the reference. We then rescale show's popularity by defining it in reference to the popularity of Breaking Bad in January 2015. Breaking Bad was chosen due to its status of the highest-rated television series in the IMDb rankings, coupled with its sustained and consistently high search volume over the nine-year duration of our analysis. An example comparison of Breaking Bad and other series can be viewed in Appendix X.

For video games, we focus on PC titles as it is the only platform for which comprehensive statistics are publicly available. PC video games account only for a part of the total video game market, but the trends in gameplay are likely parallel across different consoles, with major titles often released across different systems. Steam remains the largest digital store with PC video games, holding a majority share of the market. Steam also provides access to API services that can be used to retrieve current information about video games. Steamcharts uses that information to compile up-to-date data on the average number of players playing particular titles. These data are provided for most of the Steam titles, including historic monthly counts.

Notably, while some estimates of PC video game sales are also available, the player counts are preferable for this study. Video games differ widely in the time required for their completion. Moreover, many video games focus not on completable plots but on repetitive multiplayer experiences. The ongoing player counts – even for older titles – are thus a better measure of actual attention paid to the video games at any point in time. They can also reflect players returning to previously played games, etc.

Using the Steam IDs we then collect information on player countries from Gamalytic. Gamalytic compiles information and estimated numbers based on data from Steam. Their approach for estimating country shares likely builds on the country information provided in user profiles who leave reviews for games or ones with particular games listed in their "owned" lists.[5] As the information is generally thus accessible for only a subset of actual players it includes some missing information. In particular, it is quite reliable for titles with large player counts, but smaller titles might miss information on players from countries with smaller populations. Still, the number of video games with data in our sample is approximately 50,000. Moreover, the titles with large player counts drive the changes we are interested in. Finally, we are interested in within-country changes in gameplay, so the numbers identified for particularly important titles are enough to provide consistent patterns.

As in the case of Film/TV titles, we standardize the player counts at country levels to arrive at measures with means of 0 and standard deviations of 1.

3.3. COVID trends

We consider two separate measures for COVID. First is a dummy variable indicating the period between March 2020 and December 2022. Notably, the choice of the period is subjective, with the severity of the pandemic varying largely over the studied period. Many of the effects we consider stem from work-from-home or pandemic restrictions which can vary between countries and depend on the current severity of COVID-19. Thus, as a second measure, we introduce Google Trends for the topic of the COVID-19 disease in particular countries. Google search volumes for COVID-19 have been previously shown to accurately predict the severity of COVID-19 infections across time and can be thus used as a more consistent measure than tested cases or government restrictions (see e.g. Chu et al., 2023; Minakawa et al., 2022; Pan et al., 2020; Sato et al., 2021)

[5] The exact way of estimating these shares will be described in more detail in future versions of the paper.

3.4. Final sample

Our final sample for this paper version covers 27 countries[6] with full data on the attention for film/TV and PC games, spanning months from January 2015 until December 2023. The final derived scores for each country can be viewed in the Appendix Graph 5.

4. Results

To verify the hypotheses we estimate a series of models using the panel OLS regression with country-level fixed effects. Standardised player counts constitute the explained variable, whereas film/TV attention is our main variable of interest. We start with the base model including an additional control for a global linear yearly trend and monthly seasonal effects in the model (1). We subsequently add country-specific year trends in model (2). We then test for the effects of the COVID pandemic with a COVID dummy indicator (models 3 and 4), and with COVID country-specific trends (models 5 and 6), as well as their interactions with the film/TV measure. Table 1 summarises the results of all models.

Table 1. Panel OLS regressions of standardized player counts on the standardised level of attention towards film/TV content

Player counts on:	(1)	(2)	(3)	(4)	(5)	(6)
Film/TV attention	-0.04*** (0.01)	-0.04*** (0.01)	-0.09*** (0.02)	-0.09*** (0.02)	-0.04*** (0.01)	-0.04*** (0.01)
Film/TV attention # COVID dummy^{X}			0.16*** (0.02)	0.16*** (0.04)		
Film/TV attention # COVID trend^{XX}					0.05*** (0.01)	0.05*** (0.01)
Year	0.30*** (0.00)		0.28*** (0.00)		0.29*** (0.01)	
COVID dummy^{X}			0.10** (0.04)			
COVID trend^{XX}					0.00 (0.02)	0.00 (0.02)
Month FE	Yes	Yes	Yes	Yes	Yes	Yes
Country FE	Yes	Yes	Yes	Yes	Yes	Yes
Country-specific year trends	No	Yes	No	Yes	No	Yes
Observations	2,916					
Countries	27					

Notes: Robust standard errors in parentheses. P-value: *** < 0.01; ** < 0.05; * < 0.1;
^{X} dummy for 2020.3-2022.12. ^{XX} standardised Google Trends for the topic of COVID-19 disease.

The relationship between the analyzed two sectors is negative and significant in all models with a coefficient next to film and television attention equal to -0.04 in models (1), (2), (5) and (6) and an

[6] Argentina, Australia, Canada, Belgium, Brazil, Czechia, Finland, France, Germany, Hungary, India, Italy, Japan, Mexico, Netherlands, New Zealand, Norway, Poland, Portugal, Slovakia, South Korea, Spain, Sweden, Switzerland, Turkey, United Kingdom, United States of America.

even stronger negative correlation in models (3) and (4). Taken as is, the coefficients would suggest that a one standard deviation shift in attention towards film and TV titles relates to a 0.04 SD shift in player counts (or time devoted to video games) in the opposite direction.

Furthermore, we find that the relationship has become more positive during the pandemic. In models (3) and (4), the joint relationship becomes positive, whereas in models (5) and (6) which include more detailed measures of COVID-19 trends, the relationship with Film/TV attention during harder COVID times essentially drops closer to 0. To verify the robustness of our estimations, models (2), (4) and (6) include country-specific linear time trends. However, this additional control does not affect our coefficients of interest.

To better understand the relationship between the demand in the two sectors we additionally estimate models with interactions for each year in the sample. A whole sample model estimations are reported in the first column of Table 2. The relationship changes through the studied period, becoming more negative between 2015 and 2019 (with an outlying positive relationship in 2018), then experiences three years of positive or non-significant relationship in 2020-2022 and returns to negative in 2023.

As a final check, we additionally split our sample into subsets of countries with different levels of attention devoted to film / TV titles. To do so, we use the non-standardised film/TV attention measures, which remain comparable between countries. This split allows us to focus on countries for which the identified film/TV titles consume the largest shares of available consumer attention. Notably, we lose much of the statistical power as we reduce the number of analysed countries to 6-7 per regression. Still, for the fourth quartile, we see the strongest reflection of the down-slope trend between 2015-2019, its 2020-2022 rebound and continued decrease in 2023.

Table 2. Panel OLS regressions based on the level of attention towards film/TV content

Player counts on:	Whole sample	Quartile 1	Quartile 2	Quartile 3	Quartile 4
Film/TV attention					
# 2015	0.13*** (0.03)	0.11 (0.06)	0.18*** (0.04)	0.19*** (0.03)	0.04 (0.07)
# 2016	0.05* (0.03)	0.08 (0.08)	-0.01 (0.08)	0.11*** (0.02)	-0.02 (0.03)
# 2017	-0.01 (0.03)	0.07 (0.06)	-0.05 (0.03)	0.06 (0.05)	-0.09* (0.04)
# 2018	0.15** (0.06)	0.23 (0.16)	0.22 (0.14)	0.11** (0.03)	-0.05 (0.05)
# 2019	-0.20*** (0.02)	-0.18** (0.04)	-0.20*** (0.02)	-0.23*** (0.03)	-0.21** (0.06)
# 2020	0.04** (0.02)	0.08 (0.04)	0.04 (0.04)	0.08** (0.03)	-0.01 (0.05)
# 2021	0.05 (0.04)	-0.13 (0.11)	0.13** (0.04)	0.15** (0.05)	0.03 (0.02)
# 2022	0.03 (0.05)	0.10 (0.06)	0.03 (0.13)	-0.12 (0.10)	0.06 (0.08)
# 2023	-0.10*** (0.03)	-0.14 (0.08)	-0.17* (0.09)	-0.02 (0.05)	-0.08 (0.05)

Countries	All	CH, FI, JP, KR, NL, SK	BE, BR, DE, IN, PT, SE, TR	AR, CZ, ES, FR, IT, NO, PL	AU, CA, GB, HU, MX, NZ, US

Notes: Robust standard errors in parentheses. P-value: *** < 0.01; ** < 0.05; * < 0.1;
All regressions include monthly fixed effects, country fixed effects and country-specific year trends.

Conclusions

We study the market for attention in which different creative sectors compete for the same limited resource of attention. We hypothesise that if the recent decades have seen growing attention scarcity, the relationship between the studied sectors should become increasingly negative. On the other hand, as the COVID-19 pandemic relaxed some of the constraints on time budgets, we also expect the relationship between the studied sectors to become more positive in that period.

To verify these hypotheses we study country-level monthly changes in attention and time devoted to PC video games and TV/film titles. In line with our intuition, we find evidence for a direct competition between the two sectors. As the attention and time devoted to one of them grows, the other sector suffers. These relationships mostly come from non-seasonal, unexpected deviations from the country-specific trends, suggesting that sudden spikes in attention (e.g. due to high-profile premieres) might negatively affect the performance of titles in other sectors.

At the same time, the identified relationships become non-significant or even slightly positive during the COVID-19 pandemic, suggesting that indeed it might have relaxed some of the competition for attention. Notably, this change largely disappears by 2023, suggesting a return to the prior trends.

To the best of our knowledge, this is the first analysis of this type. From the perspective of economics, it highlights the need to consider the competition between seemingly unrelated markets, as well as that money is no longer the only resource that matters for consumer decisions. For marketing, this provides a clear view of the need to account for trends and release schedules across different sectors, in line with what captures the most attention at any given point in time. It also bolsters the need for more research on how such relationships could be monetized, e.g. through brand extensions. Indeed, recent years have seen an increasing number of film/TV products based on video games[7], whose high profile likely biases our results upward.

We plan several extensions to the current analysis in future versions of the paper. First, we plan to include a broader set of countries to increase the statistical power of our analyses. Second, we plan to cover a broader range of film and TV titles to account for local productions that can capture much attention but only in subregions (and thus not appear in the IMDb Top 100 charts). Third, we plan to introduce at least one more sector in the form of live streaming, to make our analysis of the market of attention more comprehensive.

[7] To name a few, the TV shows: Halo (2022), Last of Us (2023), Witcher (2019-ongoing; based on books but with popular game series); films: Sonic the Hedgehog 1,2,3 (2020, 2022, 2024), Super Mario Bros. Movie (2023).

References

Aksoy, C. G., Barrero, J. M., Bloom, N., Davis, S. J., Dolls, M., & Zarate, P. (2023). Time Savings When Working from Home (Working Paper No. 30866). National Bureau of Economic Research.

Aridor, G., Goncalves, D., Kluver, D., Kong, R., & Konstan, J. (2022). The Economics of Recommender Systems: Evidence from a Field Experiment on MovieLens (arXiv:2211.14219). arXiv.

Brynjolfsson, E., Kim, S. T., & Oh, J. H. (2023). The Attention Economy: Measuring the Value of Free Goods on the Internet. Information Systems Research.

Chu, A. M., Chong, A. C. Y., Lai, N. H. T., Tiwari, A., & So, M. K. P. (2023). Enhancing the Predictive Power of Google Trends Data Through Network Analysis: Infodemiology Study of COVID-19. JMIR Public Health Surveill, 9, e42446. doi: 10.2196/42446. PMID: 37676701; PMCID: PMC10488898.

Minakawa, S., Saito, N., Itoga, M., & Kayaba, H. (2022). Google Trends search volumes indicate habituation against COVID-19 in Japan. 弘前医学, 72(1-4), 1-5.

Mulligan, M. (2019). The Attention Economy has Peaked. Now what? Music Industry Blog.

Netflix. (2019). Letter to shareholders.

Pan, Z., Nguyen, H. L., Abu-gellban, H., & Zhang, Y. (2020). Google Trends Analysis of COVID-19 Pandemic. In 2020 IEEE International Conference on Big Data (Big Data) (pp. 3438-3446). Atlanta, GA, USA. doi: 10.1109/BigData50022.2020.9377852.

Sato, K., Mano, T., Iwata, A., et al. (2021). Need of care in interpreting Google Trends-based COVID-19 infodemiological study results: potential risk of false-positivity. BMC Medical Research Methodology, 21, 147. https://doi.org/10.1186/s12874-021-01338-2.

Simon, H. A. (1971). Designing organizations for an information-rich world. In M. Greenberger (Ed.), Computers, Communication, and the Public Interest (pp. 37-72). Baltimore, MD: The Johns Hopkins Press.

Sinha, R. K., & Mandel, N. (2008). Preventing digital music piracy: the carrot or the stick? Journal of Marketing, 72(1), 1-15.

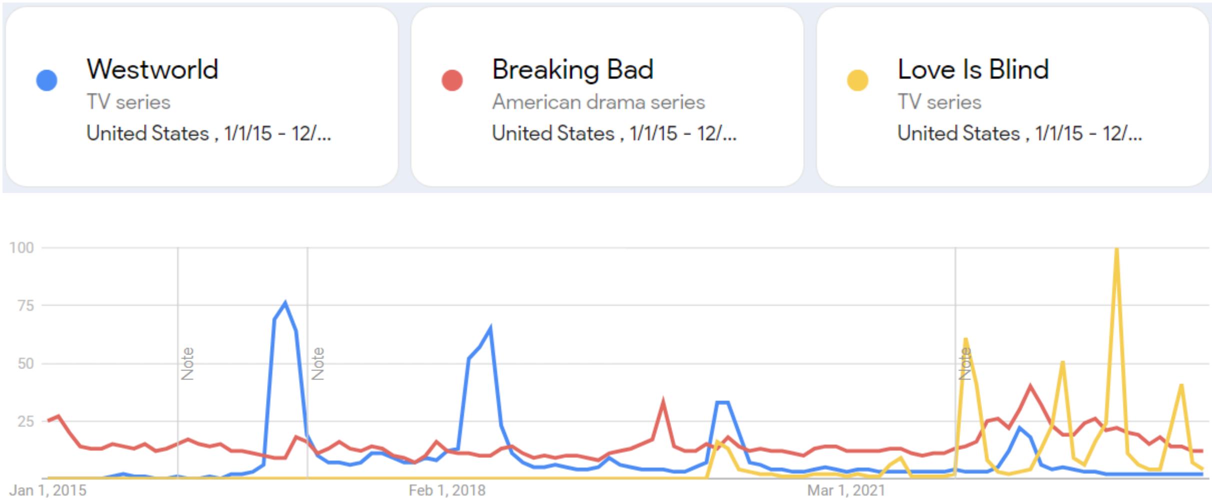
Waldfogel, J. (2017). How Digitization Has Created a Golden Age of Music, Movies, Books, and Television. Journal of Economic Perspectives, 31(3), 195-214.

Waldfogel, J. (2018). Digital Renaissance: What Data and Economics Tell Us about the Future of Popular Culture. Princeton University Press.

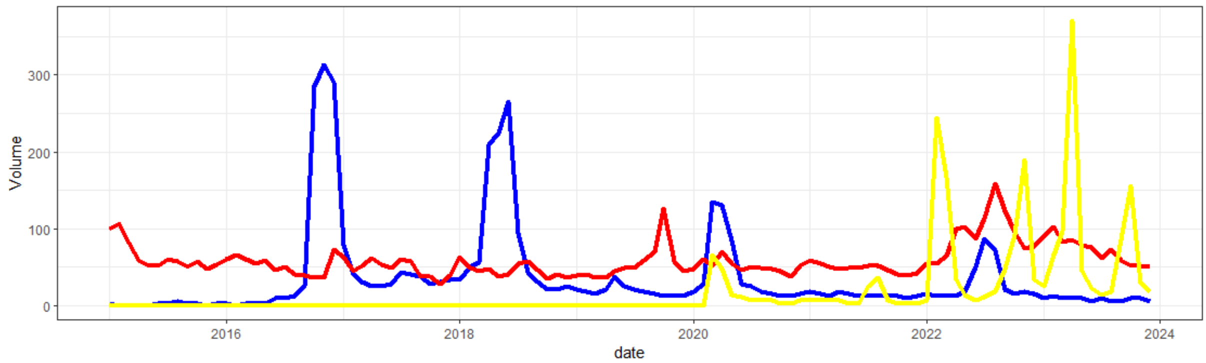
Webster, J. (2016). The Marketplace of Attention: How Audiences Take Shape in a Digital Age. London: The MIT Press.

Varian, H., & Choi, H. (2009). Predicting the Present with Google Trends. Economic Record, 88. doi: 10.2139/ssrn.1659302.

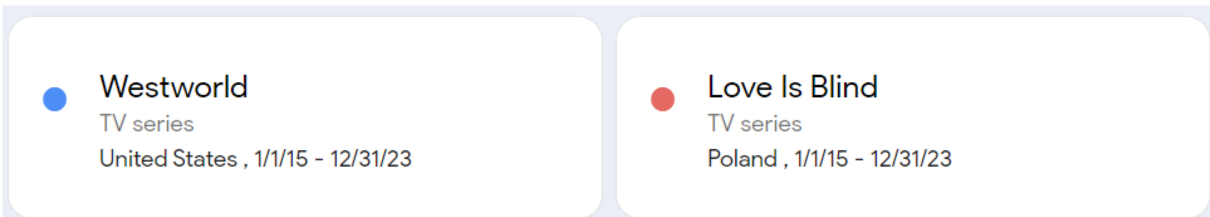
Appendix

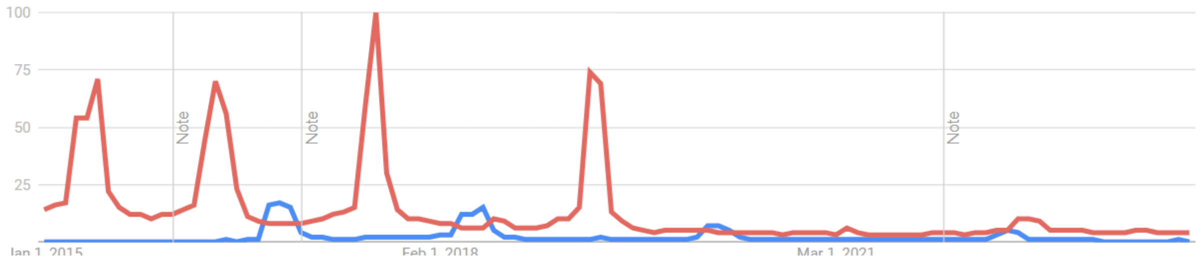


Graph 1. Interest over time output for titles Westworld, Breaking Bad and Love is Blind in the United States from 1st of January 2015 till 31st of December 2023. Available online at https://trends.google.com/trends/explore?date=2015-01-01%202023-12-31,2015-01-01%202023-12-31,2015-01-01%202023-12-31&geo=US,US,US&q=%2Fm%2F011c6zcc,%2Fm%2F03d34x8,%2Fg%2F11fnb5385q&hl=en

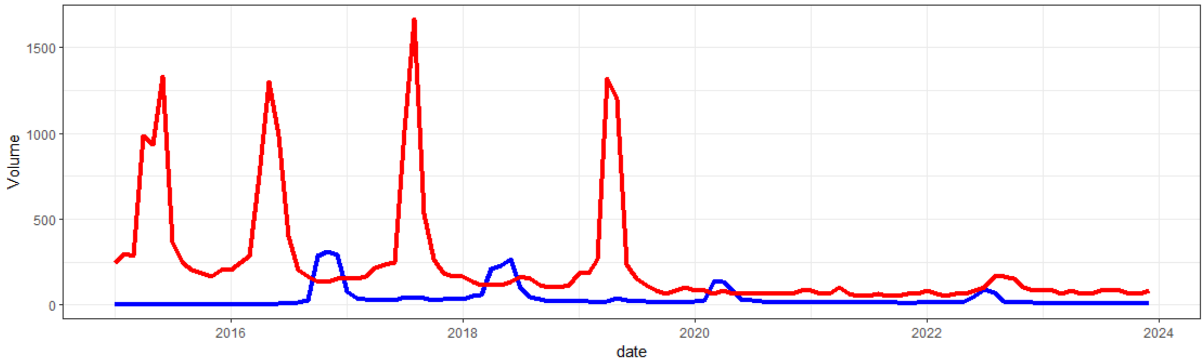


Graph 2. Relative popularity index for titles Westworld, Breaking Bad and Love is Blind in the United States from 1st of January 2015 till 31st of December 2023 as calculated by our approach.





Graph 3. Interest over time output for titles Westworld in the United States and Game of Thrones in Poland from 1st of January 2015 till 31st of December 2023. Available online at https://trends.google.com/trends/explore?date=2015-01-01%202023-12-31,2015-01-01%202023-12-31&geo=US,PL&q=%2Fm%2F011c6zcc,%2Fm%2F0524b41&hl=en



Graph 4. Relative popularity index for titles Westworld in the United States and Game of Thrones in Poland from 1st of January 2015 till 31st of December 2023 as calculated by our approach.

Graph 5. Standardized trends by country over 2015 till 2023.

Graph 6. Google Trends Tool schema (part 1).

Python code which scrapes IMDb TV and movie popularity charts URLs from the Wayback Machine for multiple years.

```
years_urls = []
for year in range(2014, 2025):
    url_tv = "https://web.archive.org/web/" + str(year) + "0101000000*/https://www.imdb.com/chart/tvmeter"
    url_mov = "https://web.archive.org/web/" + str(year) + "0101000000*/https://www.imdb.com/chart/moviemeter"
    years_urls = years_urls + [url_tv] + [url_mov]

urls = []
for url in years_urls:
    d = webdriver.Chrome()
    d.get(url)
    sleep(2)
    elements = d.find_elements("xpath", "//div[@class='calendar-day ']//a")
    href_attributes = [element.get_attribute("href") for element in elements]
    urls = urls + href_attributes
    d.quit()

wayback_links = pd.DataFrame(urls, columns=["url"])
```

Python code which retrieves IMDb titles and their URLs from a given list of Wayback Machine URLs.
The IMDb IDs are then extracted from the URLs.

```
imdb_titles = pd.DataFrame()
for url in wayback_links.loc[:, "url"]:
    try:
        Web = req.get(url)
        S = BeautifulSoup(Web.text, 'lxml').find_all("td", {"class": "titleColumn"})
        A = [element.find("a") for element in S]
        titles          = [element.get_text()            for element in A]
        titles_imdb_urls = [element.get("href") for element in A]
        imdb_titles_new = pd.DataFrame({"title": titles, "imdb_url": titles_imdb_urls})
        imdb_titles = pd.concat([imdb_titles, imdb_titles_new], axis = 0)
        sleep(3)
    except:
        pass

imdb_titles["IMDb ID"] = imdb_titles["imdb_url"].str.split("/", expand = True)[7]
imdb_ids = imdb_titles.loc[:,"IMDb ID"].unique()
```

Python code which retrieves the Google ID (named in here "mid" as its in the middle of google trends link) by searching the Google Autocomplete API and finding the relevant topic for a given IMDb title.

```
def request_type_link(id):
    url = f"https://trends.google.com/trends/api/autocomplete/{id}?hl=en-US&tz=-60&key={api_key}"
    response = httpx.get(url=url)
    topics_data = json.loads(response.text.replace(")]}',", ""))
    data = topics_data["default"]["topics"][0]
    result = pd.DataFrame.from_dict([data])
    result["mid"] = result["mid"].str.replace('/','%2F')
    result["IMDb ID"] = id
    return result

result_df = []
for id in imdb_ids:
    try:
        result_df.append(request_type_link(id))
    except:
        pass
df = np.array(result_df).reshape(-1,4)
google_ids = pd.DataFrame(df, columns = ["mid", "title", "type", "IMDb_ID"])
```

A slower way to gather the Google IDs, which creates the same results as Autocomplete API, is to query the Wikidata API.

```
def get_entity_id(imdb_id):
    query = """
    SELECT ?item ?itemLabel ?freebase_id ?google_kg_id WHERE {
      ?item wdt:P345 "%s".
      OPTIONAL { ?item wdt:P646 ?freebase_id }
      OPTIONAL { ?item wdt:P2671 ?google_kg_id }
      SERVICE wikibase:label { bd:serviceParam wikibase:language "
[AUTO_LANGUAGE],en". }
    }
    """ % imdb_id
    wikidata_endpoint = "https://query.wikidata.org/sparql"
    try:
        response = requests.get(wikidata_endpoint, params={'query':
query, 'format': 'json'})
        data = response.json()
        entities = data['results']['bindings']

    except:
        entities = None

    if entities:
        entity = entities[0]
        freebase_id = entity.get('freebase_id', {}).get('value') if
'freebase_id' in entity else None
        google_kg_id = entity.get('google_kg_id', {}).get('value') if
'google_kg_id' in entity else None
        return imdb_id, freebase_id, google_kg_id
    else:
        return None, None, None


imdb_ids_list, freebase_ids_list, google_kg_ids_list = zip(*
(get_entity_id(imdb_id) for imdb_id in missing_topic_ids))

google_ids = pd.DataFrame({'IMDb_ID': imdb_ids_list, 'Freebase_ID':
freebase_ids_list, 'Google_KG_ID': google_kg_ids_list})
google_ids = google_ids.assign(mid = lambda x:
np.where(x["Freebase_ID"].isna(), x["Google_KG_ID"],
x["Freebase_ID"]))
```

Graph 7. Google Trends Tool schema (part 2).

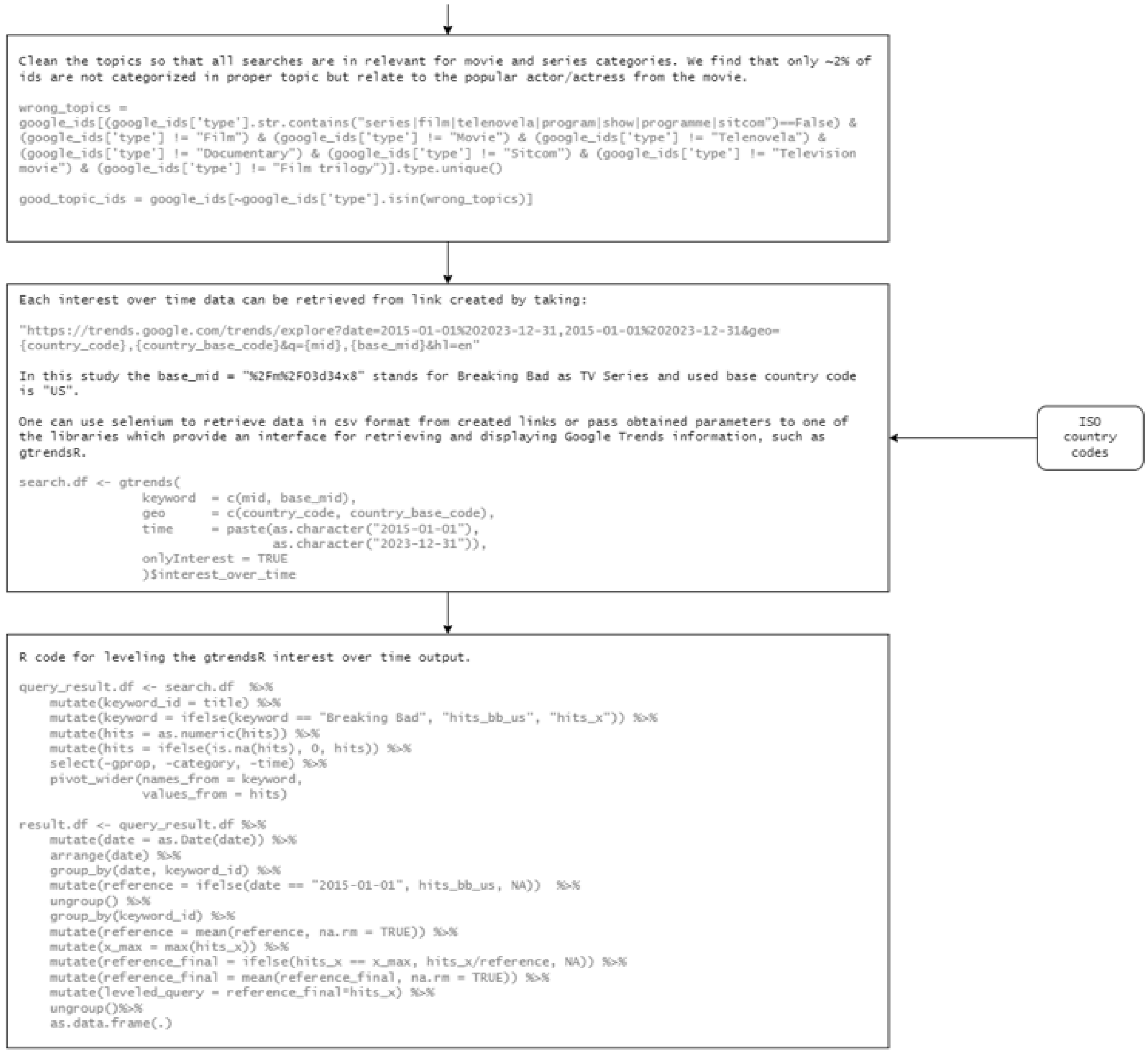


Table 1. Dates of top 100 popular lists available through the Wayback Machine (included in the final paper).

Table 2. TV series and movie titles (included in the final paper).