OPEN PLATFORM

ANDROID TV: THE ERA OF OPEN PLATFORMS

HOW THE SYSTEM CAPTURES THE TELEVISION MARKET AND HOW IT AFFECTS THE BUSINESS OF LOCAL AND BIG IPTV/OTT OPERATORS.

REVIEWS AND FORECASTS

DIALOGS

TECHNOLOGIES

ANDROID TV:
THE ERA OF OPEN PLATFORMS

How the system captures the television market
and how it affects the business of local
and big IPTV/OTT operators.

Author: Hennadii Mitrov

Can a local IPTV/OTT operator build an internationally recognised service? How can they compete with giants who invest millions of dollars in development? Android TV can help.

Android TV is an open platform with built-in Google services.
This is an operating system for premium set-top boxes and smart TVs.
The solution is being improved but will remain free.

ANDROID TV HISTORY

In 2010, Google Inc., in conjunction with Intel, Sony, and Logitech, released Google TV, an interactive platform for smart TV. This solution ran Android OS with an integrated Chrome browser. In 2014, its Android 5.0 Lollipop-operating successor, Android TV, appeared. Nexus Player, created by Google in partnership with Asus, became the first device operating on the new platform. In May 2015, the Nvidia Shield set-top box was released.
With the advent of Android TV 6.0, more and more manufacturers began to produce STBs and

smart TVs based on the new OS. Android TV 6.0 offered out-of-the-box voice control, built-in PVR and PiP (picture in picture) functions, as well as an elaborate Leanback library that allowed operators to change the user interface.

Channel API and the TIF framework allowed developers to display content from apps directly on the Android TV start screen. Users were able to search for content in all of their apps at once.

ANDROID TV: STATISTICS AND PROSPECTS

More than half a million Android TV devices are sold every month. According to the S&P Global Market Intelligence forecast, more than 40.1 million devices will be sold in 2021, 27 million of which will be smart TVs and 11.9 million of which will be set-top boxes.

In 2022, over 100 million devices will run Android TV, including most of the new set-top boxes.

Source: Rethink

Prospects for Android TV involve smart TVs as well. According to Strategy Analytics, in 2018, 157 million smart TVs were sold, and every tenth unit ran Android TV. The platform's main competitor is Tizen OS from Samsung: it occupies 20% of the market (with 32 million devices sold in 2018). If we consider the sales of AOSP-operating smart TVs, we can see that the market leader is Android.

DIFFERENCES BETWEEN THE ANDROID TV, AOSP, AND ANDROID TV OPERATOR TIER

Android Open Source Project (AOSP) is an Android version that lacks access to Google services. Manufacturers of smart TVs and set-top boxes can change and modernise the solution as they desire, so it is large companies that choose it more often. Their resources allow them to develop apps and implement new features themselves.

Android TV is a comprehensive solution for smart TVs and premium set-top boxes. The system offers out-of-the-box voice control, personal recommendations, PVR, a store with more than 5000 games and apps, access to streaming services, and much more. To release a device for Android TV, manufacturers undergo mandatory Google certification.

SUBSCRIBE

The Android TV Operator Tier is a solution for operators that allows them to brand the user interface, change the boot screen, and manage user accounts, security, analytics, and billing. It helps IPTV/OTT operators save on OS development and updates, while offering access to many features and regular updates.

Devices for Android TV are available in retail stores. Android TV set-top boxes and TVs are not dependent on the operator and can operate on any network.

Google reported at the Connected TV World Summit in London that by March 2019, Android TV would be used by 140 IPTV/OTT operators. This number has increased by more than seventeen times in the past three years.

Operator Tier is a fast and cost-effective way to enter the market, but Google does not offer turnkey solutions.

The number of operators who choose Android TV

Android TV Partners by Year
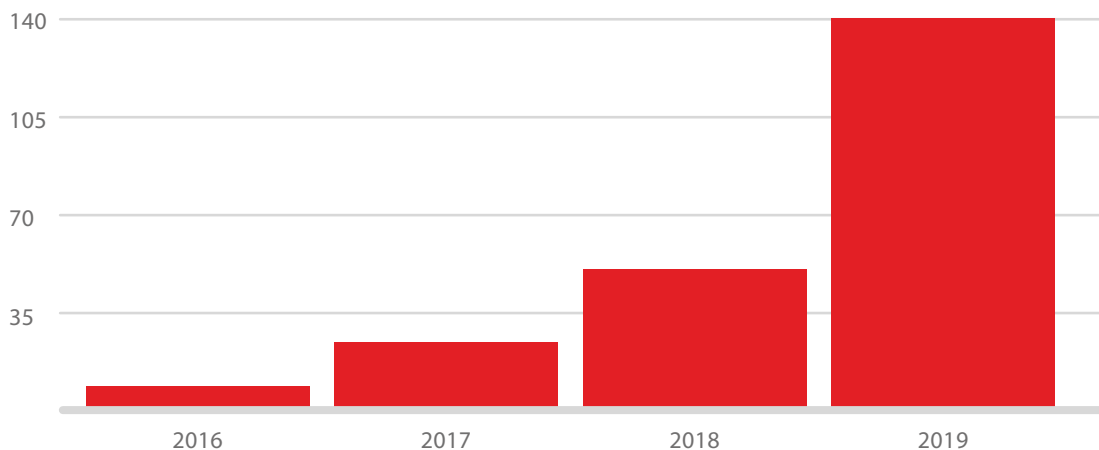
Source: 9to5Google

SUBSCRIBE

"

'Is Android TV sufficient to launch a Pay TV service? Absolutely not, Android TV only offers the middleware sitting on the box, but the entire back end (Service Delivery Platform, Head End) has to be developed and no standard offering exists today from Google'.

Brian Jentz, Senior Director of Product Management, Technicolor

ADVANTAGES OF ANDROID TV

Relevant solution

✔ More than 5000 games and apps are available to users

✔ Regular security updates and new features

More possibilities

✔ Home security, IoT, VR, and much more

Savings

✔ Lower OS support costs

✔ Developing apps for the Android TV ecosystem is cheaper than developing them for proprietary operating systems.

Premium UI/UX

✔ Google has invested years into the development of a convenient UI/UX, in order to make it accessible to each operator.

SUBSCRIBE

IS GOOGLE IN CONTROL?

Operators fully control their own set-top boxes and smart TVs. However, Google sets some requirements, for example, mandatory updates every ninety days.

The corporation collects only data to display as customised ads. IPTV/OTT service statistics and viewers' personal data are secure. Users can install apps that collect more data, but neither Google nor the operator can influence this.

One of Android TV's security measures is Widevine DRM support. The system encrypts the content so that only an authorised user can play it. The solution supports more than a billion devices: smart TVs, set-top boxes, smartphones, game consoles, and tablets. To minimise the risk of piracy and cyberattacks, some operators integrate third-party security systems into Android TV.

THE BATTLE FOR HDMI 1

Android TV does not restrict access to Google Play. For this reason, the Android system remains free — Google makes money on an ecosystem of apps. Some companies choose AOSP, where the app store is under their complete control.

Operators who have chosen Android TV have accepted the viewer's freedom. By allowing the installation of competing apps, the service provider retains control over the HDMI 1 port and the main place in the search results. To subscribe to Netflix, the viewer doesn't need a third-party device capable of supplanting the operator's device in the future.

SUBSCRIBE

THE ERA OF OPEN PLATFORMS

IPTV/OTT operators may abandon proprietary solutions in favour of open platforms in the coming decade. Companies will rely on Android TV or similar solutions, such as RDC from Comcast or Frog from the French company WyPlay. At the same time, set-top boxes with proprietary operating systems will dominate in undeveloped markets, where reliable broadband connections are not available to viewers.

With Android TV, local IPTV/OTT operators build premium-level services and compete with companies that invest millions of dollars into research and development. Larger operators also appreciate the system's advantages are because in an international battle for viewers' attention, they need platforms with great potential. One of them is Android TV.

High-performance set-top boxes with 4K and Dolby Digital Plus™

MAG520 w3
MAG520

Linux set-top box

Learn more

MAG500 A

Android TV set-top box

Learn more

Enjoy the cinema experience from the comfort of your home

Powered by an Amlogic chipset, the Series 5 MAG set-top boxes are high-performance devices for IPTV/OTT services. The media players support the HEVC codec that enables smooth 4K playback even with limited bandwidth. And Dolby Digital Plus fully immerses viewers in films, TV shows, and concerts.

SMART
MUSIC PLAYER APP

The Whooshi music player app has everything you love about music players and a little magic.

Create your personal audio profile in just a few minutes. Pass the hearing test, and Whooshi will do the rest, using sophisticated algorithms.

Plug sound effects in and get a new musical experience.

PORTABLE
SOUND AMP

It will make your favorite headphones wireless and will amplify the sound.

The Whooshi amplifier fits the hand perfectly, which makes track-switching, volume- and play-controlling very simple.

It is easy and firmly fastened to clothes, so go for a run, ride your bike, and train with Whooshi.

WHOOSHI

8K:
AHEAD OF TIME

We're talking about the future of the 8K standard, and when technology will change the IPTV/OTT sphere.

Author: Marina Sluka

4K is just on the verge of becoming a market standard, but its successor, 8K, has already arrived and it's running into problems. For one thing, provider networks aren't ready for new bit rates. At present, the equipment is expensive, and there is not enough content. But can 8K change the IPTV/OTT sphere?


The terms 8K, Full Ultra HD, and 4320p all refer to 7690×4320 pixel resolution The standard has four times as many pixels as 4K and 16 times as many as Full HD. The technology is officially promoted by Samsung, Hisense, and TCL, all of which are members of the 8K Association.

> " The Association develops an ecosystem and popularises 8K products.
>
> Dr. Weidong Liu,
> leading Hisense researcher

THE NEW STANDARD: FEATURES
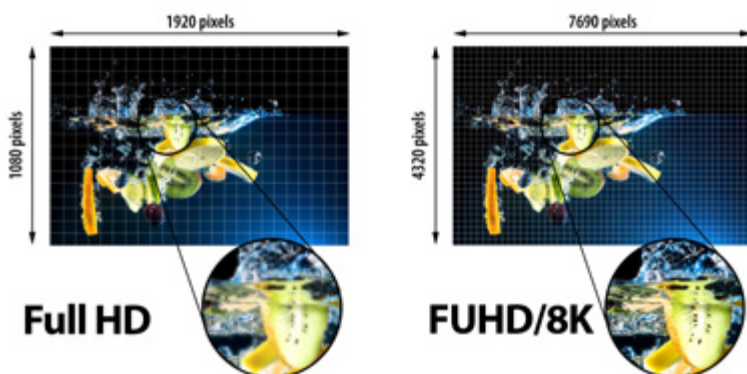
Even from a few centimetres away, pixels are indistinguishable on the 8K display with a diagonal of 70 inches. From this distance, the viewer will see only a fuzzy, granular image on a comparable Full HD display.
But the difference between 8K and 4K isn't critical: the differences aren't noticeable if you are viewing them from the same distance.

SUBSCRIBE

Most viewers watch TV at an average distance of 2.63 m from the display.

This distance hasn't changed since 2004, although the size of the display has increased significantly.

Source:
BBC Research & Development

2.63 m
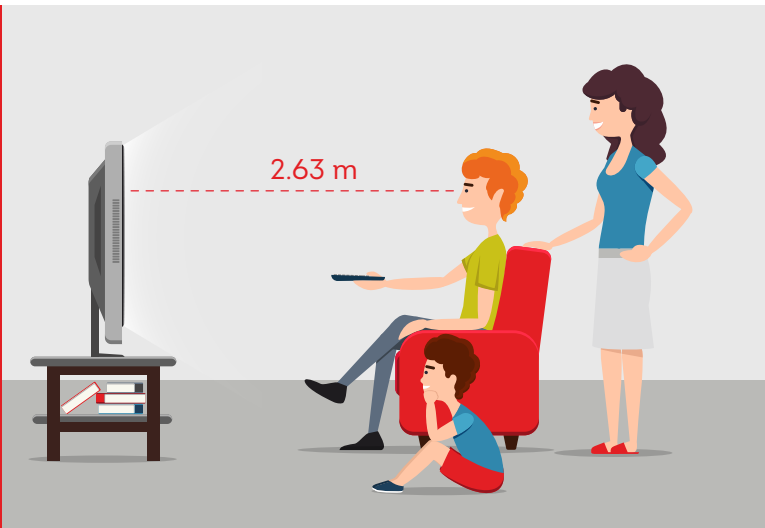
It is impossible to sense the difference between UHD and Full UHD at this distance, even on a display with a 55" or 65" diagonal. The differences can be seen only on devices with diagonals of more than 75". In 2022, the price of an 8K TV with a 75" diagonal is likely to drop to $2000; a similar 4K device will cost half as much.

49"
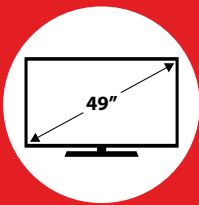
Demand for big displays will only increase: according to Statista, in 2009, only 7% of households had a TV with a diagonal larger than 49"; in 2019, every third house had a TV this size.

4K has quickly gained popularity: the prices of the new-standard TVs have declined rapidly, and their sales have steadily increased since these devices hit the retail market seven years ago. IHS predicts that by 2022, six out of ten households in North America and 40% of homes in Western Europe will have a 4K TV.

Worldwide sales of 4K TVs, mln (OVUM)
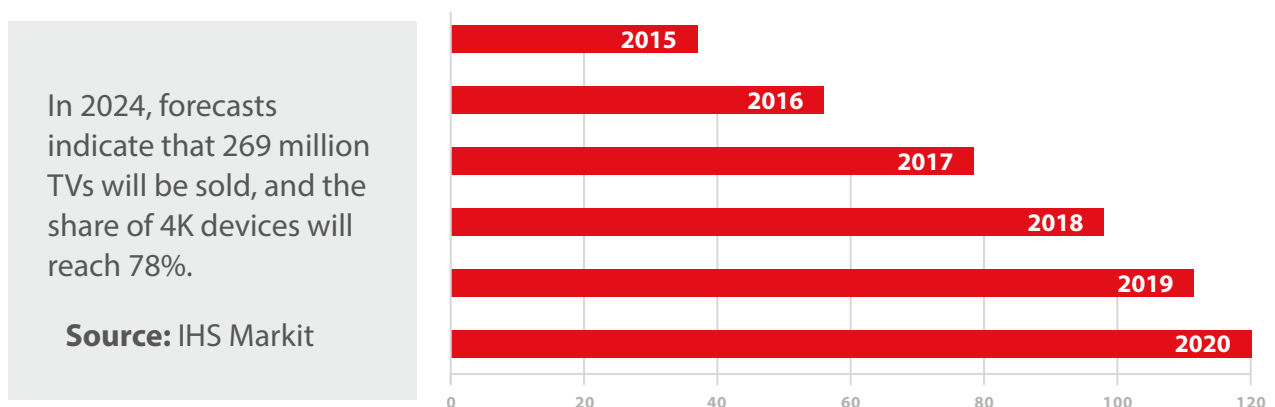
In 2024, forecasts indicate that 269 million TVs will be sold, and the share of 4K devices will reach 78%.

Source: IHS Markit

| | 2015 |
| 2016 |
| 2017 |
| 2018 |
| 2019 |
| 2020 |

0   20   40   60   80   100   120

8K TVs appeared later. Sharp demonstrated its prototype at CES in Las Vegas in 2012. LG introduced an 8K model at IFA in Berlin in 2014. At that time, Samsung presented its 8K TV. In addition to these, retail models are also being offered by Sony, Hisense, and TCL in 2019.

MAIN ISSUES ASSOCIATED WITH 8K

The introduction of the new standard is hampered by the same problems that plagued 4K five or six years ago and Full HD prior to the introduction of 4K. Neither the IPTV/OTT operator networks nor the physical media were ready for the new bit rates. Currently, there is almost no compatible content, and 8K devices are inaccessible to most users.

Infrastructure is not ready

To watch content in 8K, the viewer needs not only a suitable TV but also a dedicated satellite or high-speed broadband connection. Delivery of content at this resolution requires completely new technologies and infrastructure.
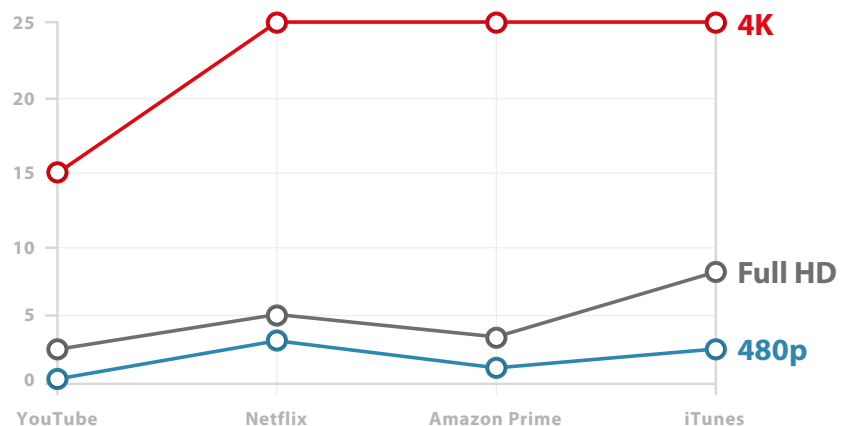Streaming videos in 4K requires a connection with a speed of 15–25 Mbit/s. With 8K, you need six times more bandwidth 80 to 150 Mbit/s. Most families have more than one television. If the devices work simultaneously, stable streaming requires a 1 Gbit/s connection. Internet providers cannot offer this.
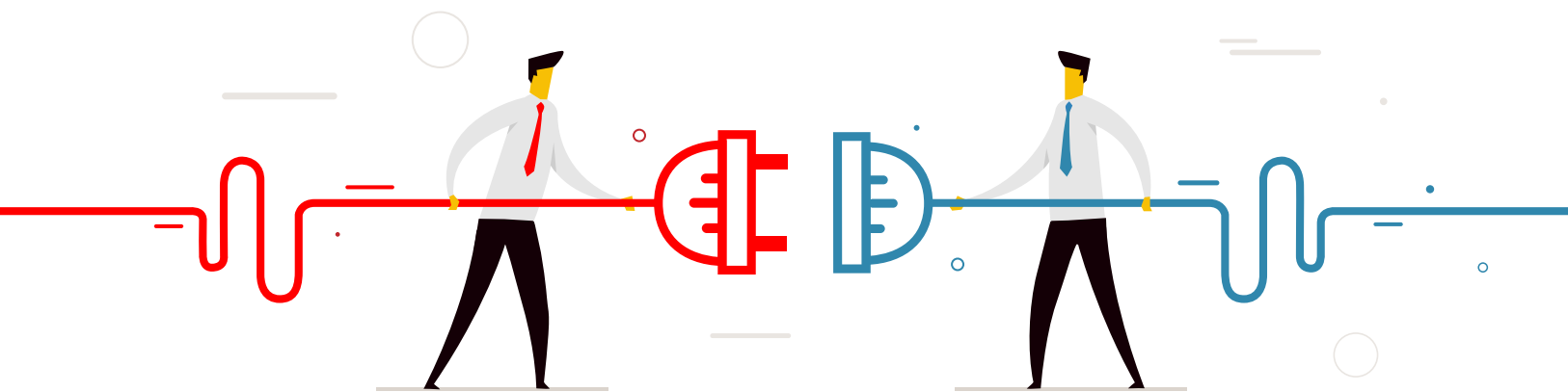
Streaming services: recommended connection speed (cable.co.uk)

The average speed of mobile networks in 2019 is 26.1 Mbit/s, for fixed-line networks, it is 57.9 Mbit/s.

Source:
Speedtest Global Index

	YouTube	Netflix	Amazon Prime	iTunes	
4K	15	25	25	25	
Full HD	2.5	5	3.5	8	
480p	0.5	3	1	2.5	

SUBSCRIBE

The 4K standard has also faced problems associated with infrastructure and media. It was impossible to compress UHD videos to a bit rate that allowed them to be transmitted through existing networks, especially mobile networks. To fix this, the Joint Collaborative Team on Video Coding (JCT-VC) developed an HEVC codec (h.265), which was standardised in 2013. The codec became the successor to h.264 and allowed videos to be compressed twice as much without losing image quality.

There are no suitable physical media. Even a compressed film in 8K takes up a few hundred gigabytes, and maximum-capacity Blu-ray discs contain only 100 GB.

HEVC has accelerated the introduction of 4K, and the impetus for 8K could be the format that evolves from it — the FVC (Future Video Codec). Its developers promise the new codec will be able to compress video 50% more efficiently. By October 2019, the test version of H.266 will be released, and by June 2021, the first hardware codecs will appear.

Manufacturers of HDMI cables are ready for the introduction of 8K. In November 2017, the HDMI 2.1 standard was released, which supports 8K transmission at 60 fps. More importantly, the technology allows 4K HDR transmission at 120 fps.

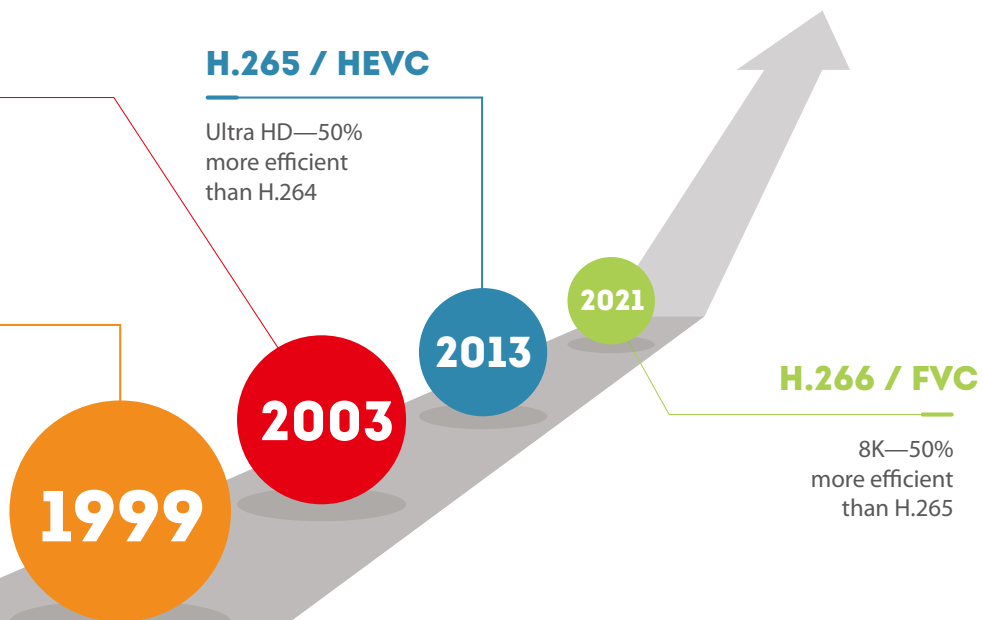HDMI™
HIGH-DEFINITION MULTIMEDIA INTERFACE
2.1

H.264 / AVC

Internet, computers, portable and HD TVs

H.265 / HEVC

Ultra HD—50% more efficient than H.264

MPEG-4

For low bit rate videos

2021

2013

H.266 / FVC

8K—50% more efficient than H.265
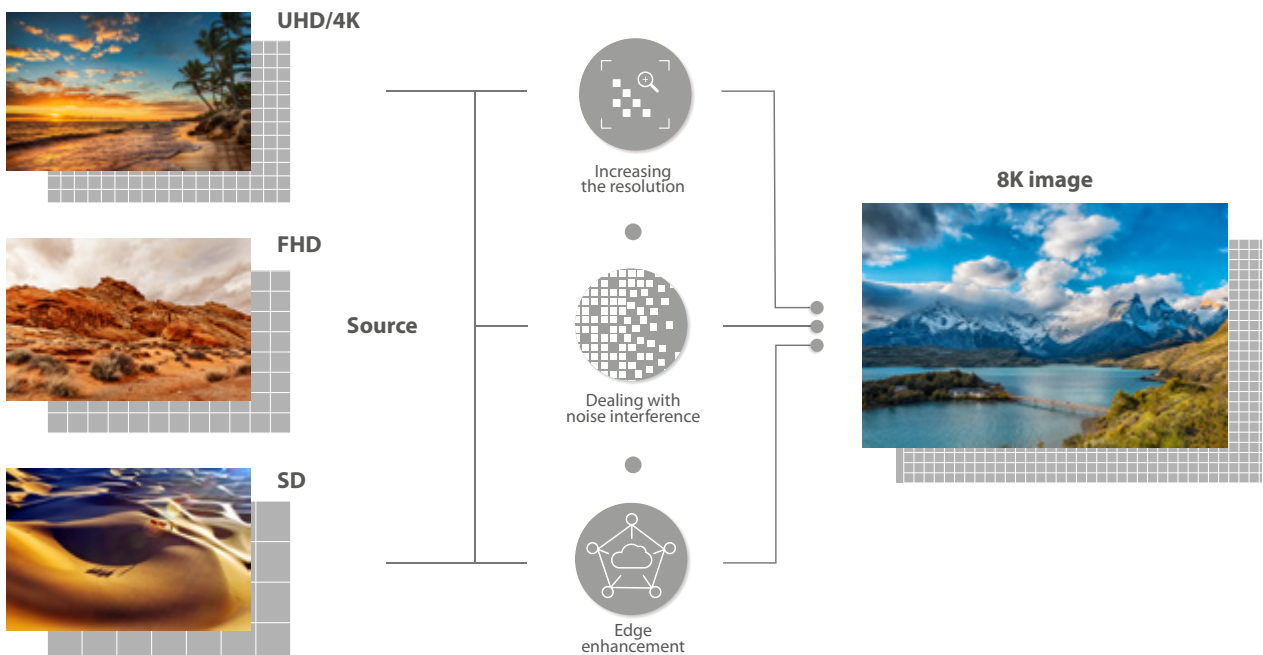
2003

1999

16

LACK OF CONTENT

Films and TV shows are currently made in 8K. For example, *The Hobbit* trilogy, *House of Cards,* and *Guardians of the Galaxy Vol. 2,* were shot in this new resolution. Japan and Italy have already launched FUHD channels covering major sporting events like the Olympic Games. But there is not enough content. However, there is no lack of content in the 4K standard.
The main streaming players like Netflix and Amazon Prime Video have a stake in 4K. For example, the Netflix database has 630 films and TV shows available in 4K. Linear television is keeping up as well: there are 142 UHD channels with regular broadcasting in the world. Forty-six of them are broadcast in Europe. Rakuten TV is the first streaming service that promises to introduce 8K, but this won't happen before 2020.

4K: the number of films and TV shows released on Blu-ray

More than 200 titles are released on Blu-ray every year.

Source: HD Report
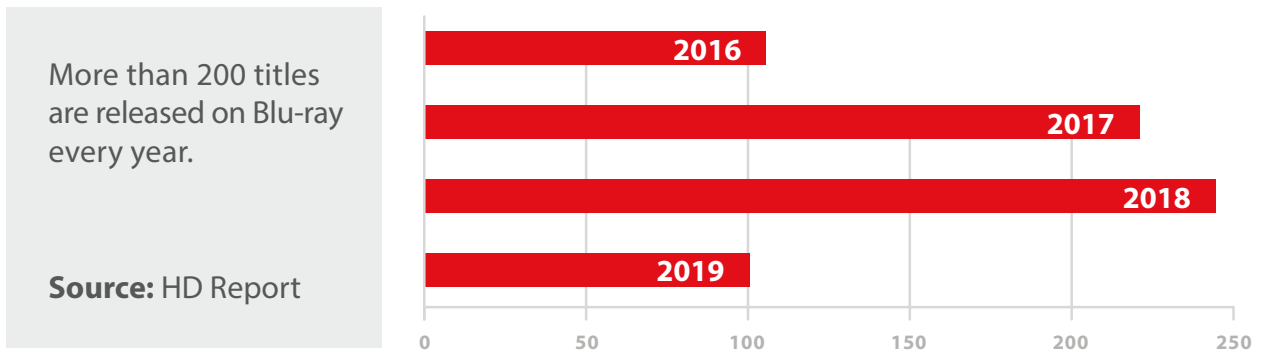
Thankfully, the problem has a solution: 8K TVs increase resolution programmatically. For example, the neural network of Samsung devices has been trained by sampling of millions of images of different resolutions, which allows it to generate missing pixels. The system automatically 'fills in the gaps' in the image and improves the quality up to Full UHD.

UHD/4K

FHD

Source

SD

Increasing the resolution

Dealing with noise interference

Edge enhancement

8K image

SUBSCRIBE

OTHER APPLICATIONS OF 8K

This technology is already being applied in medicine. Endoscopes with high-resolution cameras can be used to examine patients and these devices can even be operated them
in a minimally invasive way.
High resolution is also used in virtual reality (VR), where it is important to be able to zoom in on images with no loss of quality. Even at 16× magnification, VR images maintain Full HD quality.

**8K has great potential, but in the coming years, Ultra HD will remain the standard.
The amount of content is growing; networks are ready for new bit rates, and modern 4K TVs are available to viewers.
The prices for Full HD and 4K set-top boxes are almost equivalent now, so IPTV/OTT operators are launching projects based on more modern devices.**

SUBSCRIBE

MAG520   MAG520 w3

4K Linux set-top box

Linux 4.9

4K and HEVC support

Dual-band Wi-Fi 2.4 and 5 GHz
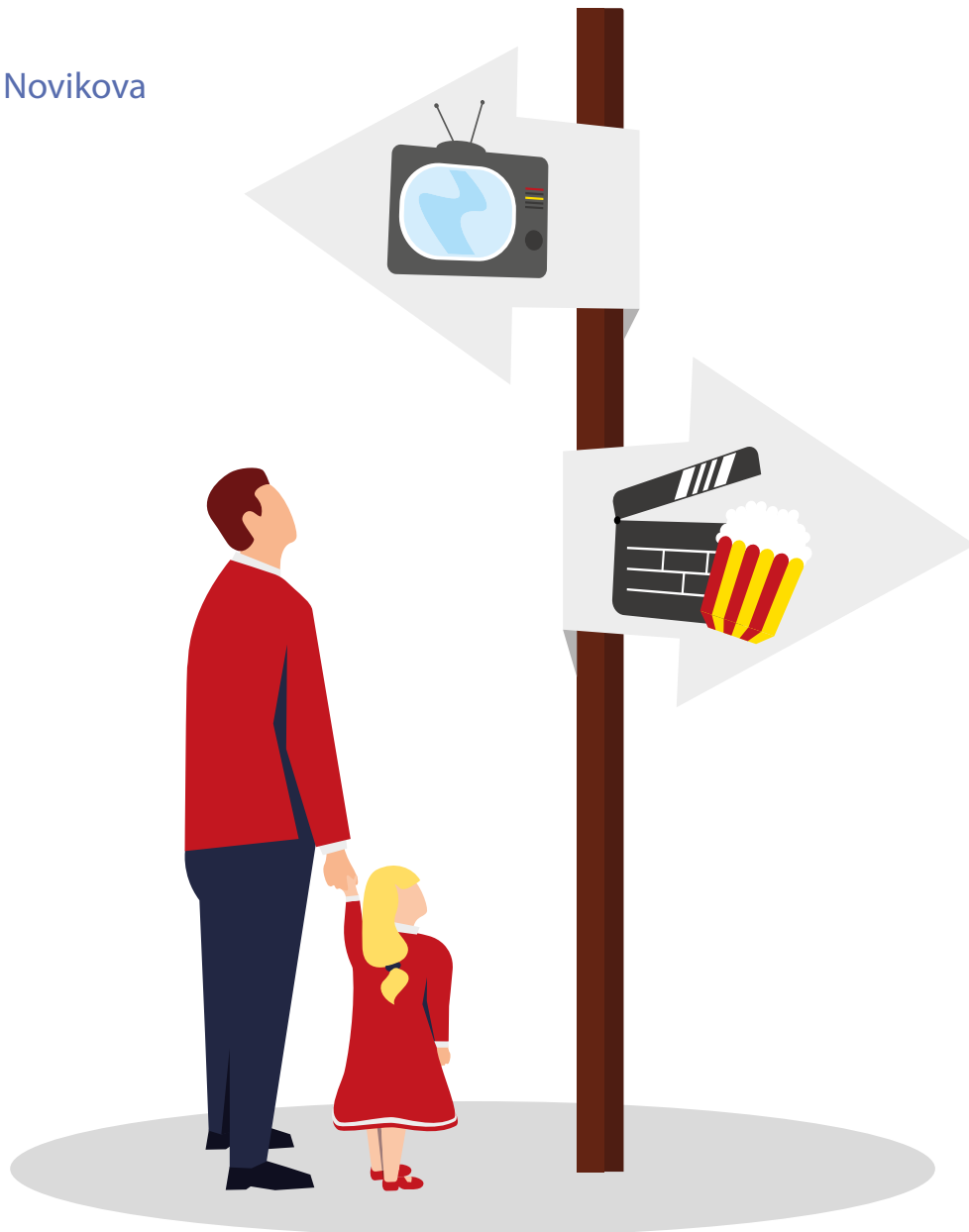(only for MAG520w3)

4x ARM Cortex-A53

1 GB RAM

Dolby Digital Plus™ sound

Learn more

CINEMA VS. TELEVISION

Insights into the eternal confrontation between Cinema and Television and which will ultimately conquer the hearts of viewers.

Author: Anna Novikova

Throughout their history, Cinema and Television have fought for viewer attention. Always trying to outperform each other, the industries attracted fans with new technologies, genres, and opportunities. Who will win the war and how will the confrontation affect the development of the entire entertainment industry? We'll be delving into that question today.

Let's go back to the beginning

The period from the end of the 19th to the mid-20th century, we can safely call the golden age of the film industry. Black and white films with legends like Charlie Chaplin, Greta Garbo, and Rudolph Valentino — gathered thousands of enthusiastic fans. At this time, Cinema's only real rival was Theatre. It was significantly inferior to the new technology in both capabilities and accessibility. Having no other competitors, the studios competed with each other, astonishing audiences with more and more sophisticated scenarios and special effects.

Then, everything changed with the mass production of the Television set. Families who lived in the suburbs, away from city entertainment, in particular, welcomed this new "technological miracle" with open arms. Family evening in front of the TV — became a pastime.
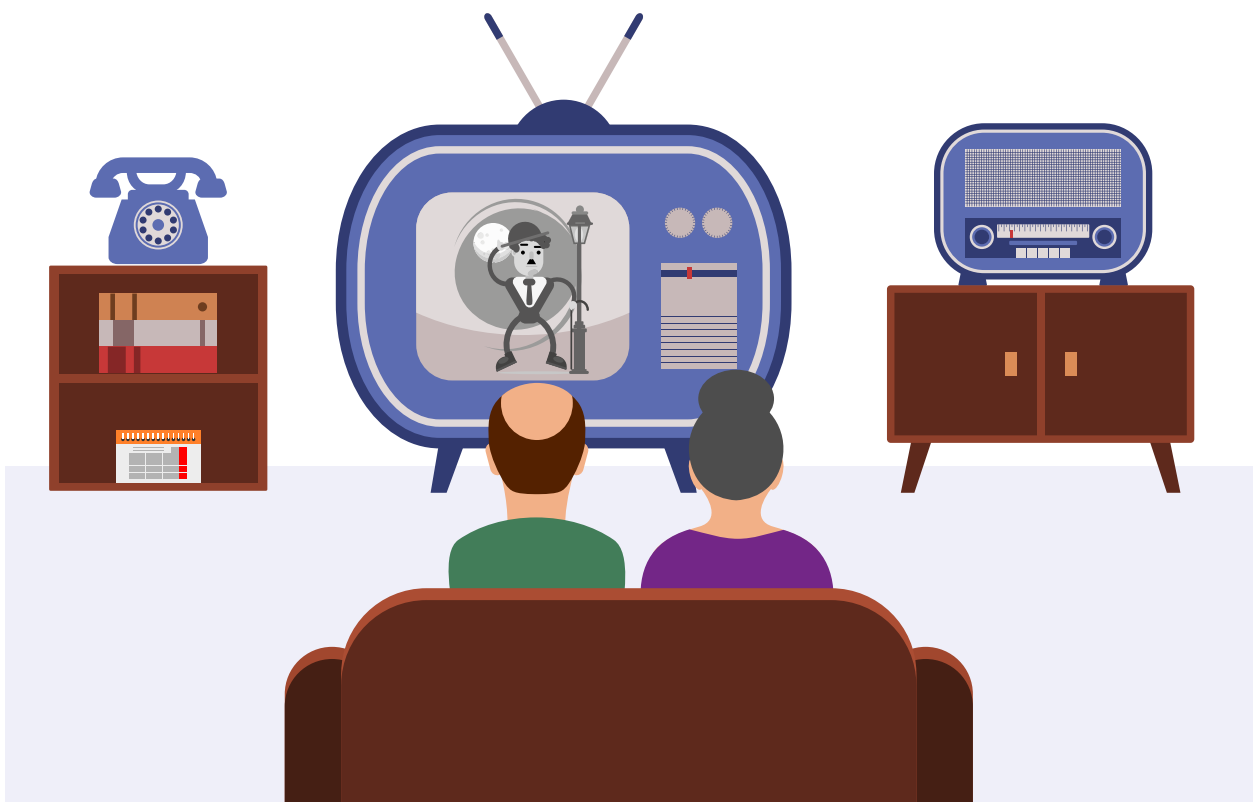
Many townspeople also prefered to sit in front of their own personal screen rather than spend time and money going to the cinema to see something that they might not even like. With their own television, they could easily just turn off the program and do something more pleasant. But the main thing was — the television was free!

The fall in film revenues made Hollywood nervous. To rectify the situation, they had to offer viewers something new.

From the 40s to the 60s, cinemas tried different tricks, for example, "CinemaScope", a new widescreen format, became a real hit. The large curved screen with stereo attracted long lines of people, and the technology became a phenomenon of Western culture, about which even songs were composed.
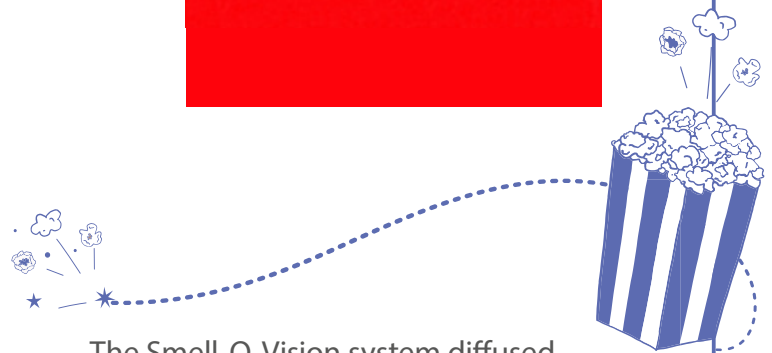
The Smell-O-Vision system diffused odours in cinemas, in sync to screen events. It was first used in 1960, for the film *Scent of Mystery*: each of 30 unique aromas was accompanied by a different soundtrack.

Between the years 1952-1954, the first experiments with a three-dimensional cinema picture were undertaken. But, due to the high cost and imperfection of cinema's 3D-technology, the 3D film disappeared for several decades.
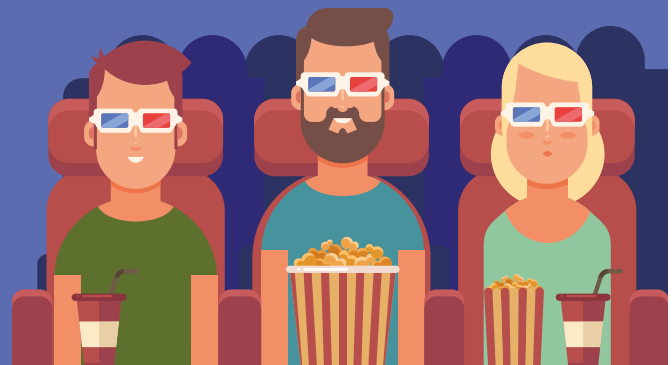
A successful solution to the competition was the division of audience niches. Hollywood began to focus on teenagers who wanted to get out of the house or go on a date. Especially popular were drive-in theaters. The content of the films also changed: teenage cinema had more rebellious characters, rock 'n' roll in the soundtrack, and the script could include dialogue and vocabulary, that was not suitable for TV.

Later, they compromised: TV-channels bought old films to broadcast, and the film industry began to use the advertising opportunities to attract people to premieres. In 1975, after a large marketing campaign, the film Jaws broke all records and laid the foundation for the summer blockbuster genre.

This collaboration served to create a TV/Movie quasi-audience. And now even streaming services are also involved in the process.
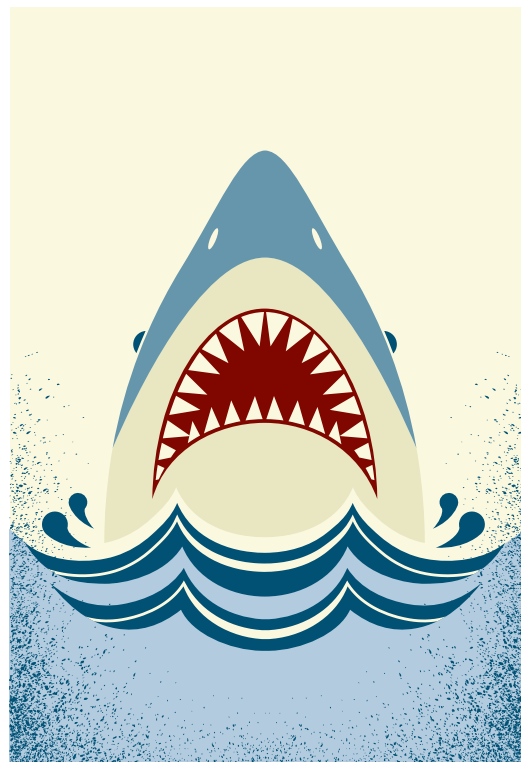
WHO IS THE MODERN FILM AND TV VIEWER?

In order to more accurately answer the question: "What do viewers prefer: cinema or TV", — we have to consider who the viewers are. TV and cinema segment their audiences in different ways, so let's consider each one in more detail.

Film audience

The film industry uses several methods of viewer classification, but more often than not classifies viewers by age in a similar way to the film certification categories (U, PG, 12A, 15, 18).

Age group

| Children (5-11 years old) | Families | Teenagers / Students / Young Couples | Adults |

UK AIM's committee study offers several new segmentations. The lifestage classification focuses on what each group expects from a visit to the cinema. For example, **teens prefer to eat popcorn** while watching comedies and horrors, while **people over 40 drink good wine** during a thought-provoking film.

Attitude classification aims to highlight cinema fans who often go to the cinema and encourage others to go.

Relation categories

| Enthusiasts / Frequent Spectators | Those who go to the cinema due to boredom | Reluctant | Non-attenders |

In addition to desire, the decisive factor is the ability to go to the cinema. For example, teenagers from the suburbs would like to go to the cinema every Friday, but without a car, this is a problem. Big families often refuse to go because tickets, travel expenses and food can be too expensive. Then there are older people who rarely go to late sessions.

SUBSCRIBE

TV AUDIENCE

Ericsson ConsumerLab has proposed one of the most popular ways to classify television audience. Depending on television habits and preferences, they divide users into six groups:

Couch Traditionalists — watch exclusively on stationary TV

Average TV Joes — watch TV average amount of time

Screen Shifters — freely change the viewing device

Mobility Centrics — prefer to watch on mobile devices

**TV Zero** — watch TV minimally

Computer Centrics — prefer to watch on PC

The fastest-growing groups are Mobility Centrics and Screen Shifters. This is a favourable trend for owners of modern services with multiscreen functionality.

SUBSCRIBE

In the struggle between TV and cinema, this feature is also one of the most powerful advantages. The ability to view content on any device is not available for cinema and attracts more and more subscribers to the IPTV/OTT services every year.

The Streaming Video Alliance divided the audience into two age groups: millennials (18-34 years old) and people over 35 years old. The preferences of these categories are significantly different, that's very important during the launching and promoting of a new service.

35+

18-34
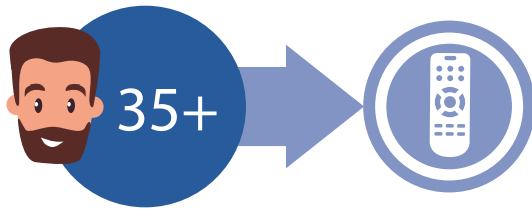
Older people prefer traditional television. They have several favourite series and channels, and it's hard for them to give up their viewing habits. It is more likely that this group will pay closer attention to high-quality IPTV, with a large selection of channels, good signal and simple control.
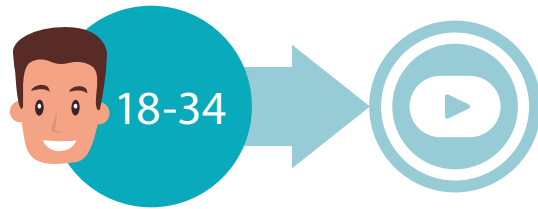
Millennials' favourite content is short videos (e.g. YouTube) and TV series. They are the main reason for empty seats in cinemas. Their reasoning is that a high-quality series delves more deeply into the plot and heroes of a story and is able to hold the viewer's attention for a long time contrary to a film. A film's duration, the amount of time that can be devoted to these aspects is very limited, that is why this group's interest in cinema is declining.

What's it all about?

ONLINE CINEMA
★ ★ ★ ★ ★
watch your favorite movie

TICKET
ADMIT ONE
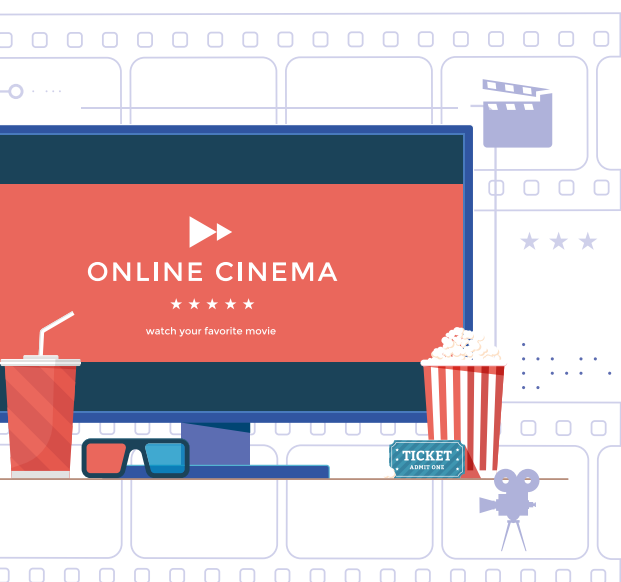
At the beginning of this decade, we may have been able to say with confidence that cinema audiences consisted mainly of youths, while TV focused on the older generation, but now the situation is not so clear. Yes, traditional TV is mostly watched by the older generation and this trend will continue. This is due to the fact that older people prefer the comfort of their own home to an exhausting trip to the cinema. The current youth, over time, will join the ranks of homebodies.
In addition to the increasing of an older audience, we can expect an influx of young users. They will be attracted by new features of modern services: free choice and growth of content quality, multiscreen, access to niche and unique content.

> The retention of TV audiences will also be affected by the development of genres that are impossible to show in cinema. For example, series and sports.
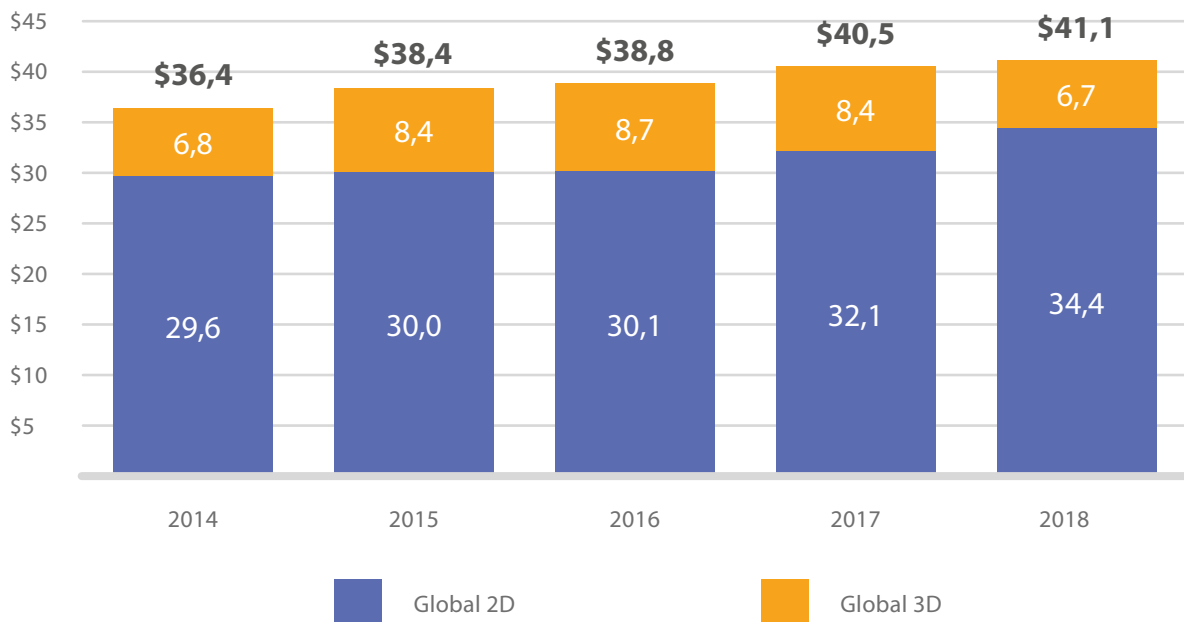
At the same time, we should not expect that cinema ranks will become empty. Studios have learned a lot from TV series creators. One of the most striking examples of studios learning from TV series is the creation of shared film universes (e.g. *Marvel* or *Star Wars*). Popular franchises captivate the audience and pay off.
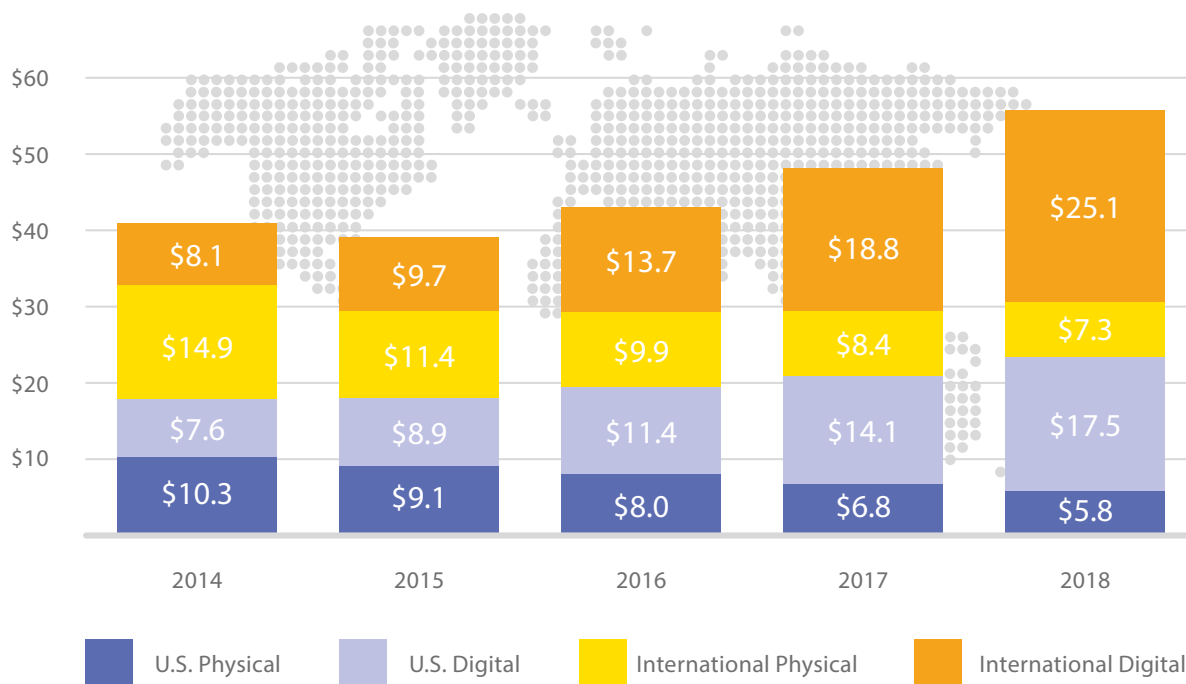
WHO IS THE REVENUE LEADER: CINEMA OR TV?

According to the THEME Report, over the past few years—both have a steady growth. Over the past 5 years, cinemas' revenue grew by 13% and reached $41.1 billion in 2018.

Global 3D Box Office — All Films (US$ Billions)

Year	Global 2D	Global 3D	Total
2014	29,6	6,8	$36,4
2015	30,0	8,4	$38,4
2016	30,1	8,7	$38,8
2017	32,1	8,4	$40,5
2018	34,4	6,7	$41,1

Legend: Global 2D, Global 3D

Source: THEME Report 2018

However, revenue growth is generated not by cinemas, but by streaming services. In 2018, online platforms generated 57.7% of total revenue, and total home viewing revenue was $55.7 billion (+16% compared to 2017). The growth came mainly due to online subscribers. The physical media part (e.g. Blu-ray discs) is decreasing.
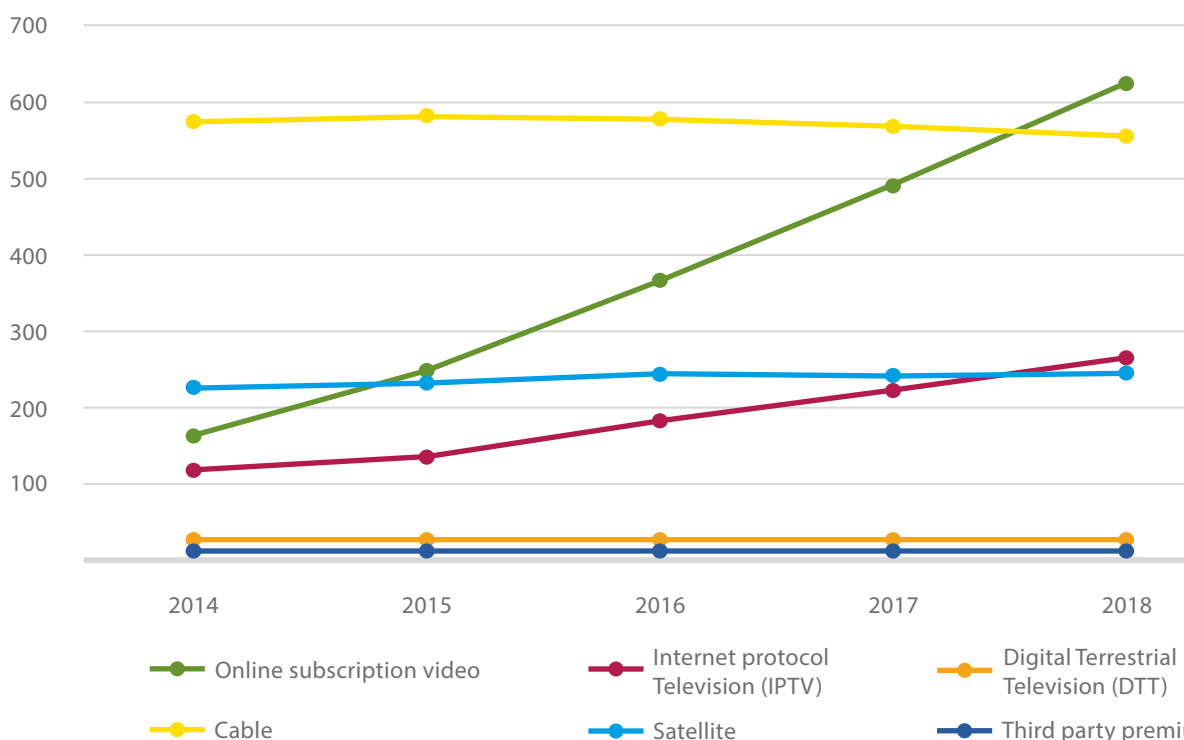
SUBSCRIBE

Global Home Entertainment Consumer Spending (US$ Billions)

	2014	2015	2016	2017	2018
International Digital	$8.1	$9.7	$13.7	$18.8	$25.1
International Physical	$14.9	$11.4	$9.9	$8.4	$7.3
U.S. Digital	$7.6	$8.9	$11.4	$14.1	$17.5
U.S. Physical	$10.3	$9.1	$8.0	$6.8	$5.8

Source: IHS Markit

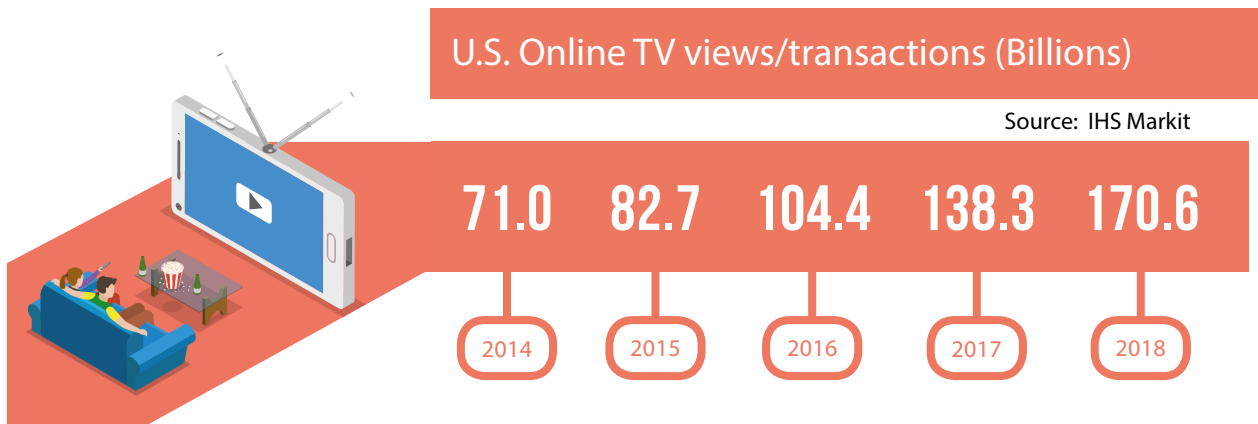■ U.S. Physical ■ U.S. Digital ■ International Physical ■ International Digital

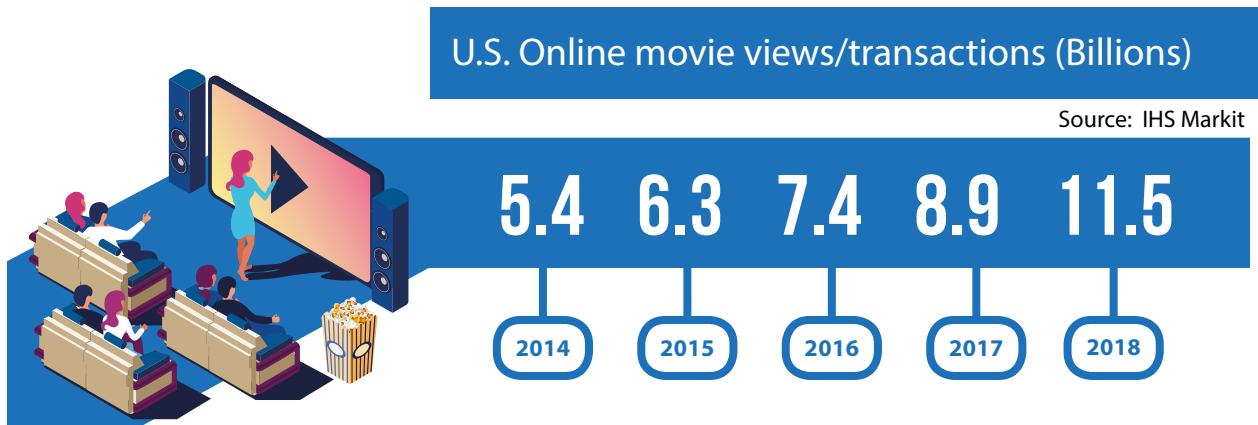The largest increase occurred due to IPTV and online subscriptions (e.g. Netflix, Amazon Prime). The study indicated that most families use TV and online video simultaneously.

Global Pay TV & Online Video Subscriptions (Millions)

Source: IHS Markit

— Online subscription video — Internet protocol Television (IPTV) — Digital Terrestrial Television (DTT)
— Cable — Satellite — Third party premium

29

SUBSCRIBE

TV channel watching is still the most popular leisure: users watch TV more than 15 times more than films. However, cinema is growing a little faster.

U.S. Online movie views/transactions (Billions)

Source: IHS Markit

5.4	6.3	7.4	8.9	11.5
2014	2015	2016	2017	2018

U.S. Online TV views/transactions (Billions)

Source: IHS Markit

71.0	82.7	104.4	138.3	170.6
2014	2015	2016	2017	2018

WHO IS THE WINNER?

The long-standing confrontation has grown into mutually beneficial cooperation. Income from films is no longer limited to their broadcast in cinemas, as now many fans of home viewing have joined the ranks of film viewers. Online services, in turn, are increasingly releasing films for the big screen.

Competitors actively learn from each other's successes and mistakes, which allows them to provide viewers with the most appropriate content and services. All this positively affects the entertainment industry quality. So in conclusion: in the long confrontation between Cinema and TV, the winner is the viewer.
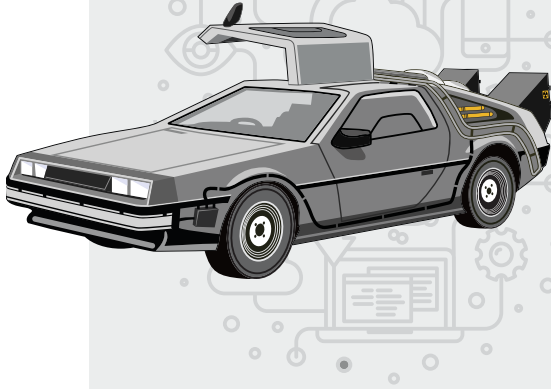
In recent years, the income gap between cinema and TV services has widened, but we cannot say that television is a clear winner.

SUBSCRIBE

YOUR WISH
IS MY COMMAND

BROADVISION experts explain how speech recognition works and how voice assistants can improve your experience.

Author: Anna Novikova

30 years ago

a TV that responded to voice commands in the movie *Back to the Future Part II* was an idea as fantastic as a flying car, and now almost everyone has access to devices with voice interfaces. In this article, we'll explain how speech recognition works and show you how voice assistants operate, using Google Assistant as an example.

EARLY EXPERIMENTS IN **VOICE RECOGNITION** »»

Surprisingly, the first home product with a voice interface (VI) was a toy. In 1987, the Julie doll, which could be trained to recognise children's speech, was released. Julie reacted to stimuli like turning off the light and read the books that came with it aloud.

A few years later, the first voice-recognition dictation software appeared, in addition to devices for the visually impaired and people who couldn't use computer keyboards due to physical limitations.

Later in the 1990s, other voice interfaces were introduced to automate business processes. For example, the VAL portal from BellSouth processed telephone inquiries and told customers about the company's services. Unfortunately, these early solutions were inaccurate and required lengthy training.

The technology steadily improved over time and today, many of the 'smart' devices available to consumers come equipped with voice interfaces. Manufacturers of technology products have incorporated voice recognition into their devices to offer their customers greater ease of use and hands-free operation. Speaking is easier than typing when you're in the midst of daily activities — while you're driving or commuting, or when you're in front of your TV.

Present-day speech recognition systems help us find information, transcribe text, and schedule appointments. They are used in interactive self-service systems, for example, technical support services.

Julie Talking Doll — the first voice-controlled consumer device.

In 1990, a license for DragonDictate, the **first 'voice typewriter',** cost $9000.

Susan Bennett — is an actress known as the American voice of Siri.

1987

1990

2011

HOW DOES **SPEECH RECOGNITION** WORK?

Devices perceive speech differently than humans. Instead of isolated words, a device hears a continuous signal, in which sounds flow smoothly into each other. A device will detect a single phrase pronounced with different intonations or voiced by different speakers as different signals. Because of the high degree of variability in human speech, voice recognition accuracy has not yet reached 100%.

The main task of voice-recognition algorithms is to interpret what has been said regardless of the pronunciation peculiarities of the speaker or the presence of background noise and other interference.

ELEMENTS OF A SPEECH-RECOGNITION SYSTEM
Speech recognition systems consist of four components:
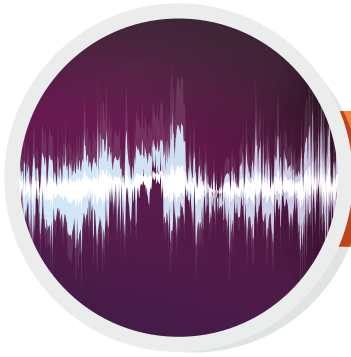
a sound-cleaning module
(to remove background noise)

an acoustic model
(to discern the speech sounds)

a language pattern
(to predict the most likely word sequences)

a decoder
(to combine the data output from the acoustic model with the language pattern to provide the final result)

1 Sound cleaning

The first task of a speech-recognition system is to evaluate the quality of the sound input and separate the desired signal from sound interference, or noise. Depending on the nature of the unwanted sound, different approaches can be used to filter speech from background noise.

NOISE SUPPRESSION

There are several ways to suppress noise in a speech-recognition system. One of them is to introduce artificial noise, including recordings of common man-made noises (for example, the sound of a car engine, wind, or rain), to the system to 'teach' the acoustic model to distinguish sound from background noise. However, when the system later encounters unfamiliar noise, the model will likely send an 'error' message.

Other noise suppression approaches rely on hardware solutions. Some smartphones are equipped with two microphones: the first microphone, located on the front of the device, catches speech with interference while a second microphone on the back picks up ambient background noise. In theory, all you have to do to get a clear signal is simply subtract the second sound signal from the first.

EXTRANEOUS VOICES

It is more difficult to single out a particular voice when several people are speaking at the same time than to distinguish speech from non-speech sounds. To isolate an individual speaker from other voices, the acoustic model adapts to the user's voice and remembers the peculiarities of his/her pronunciation.

SUBSCRIBE
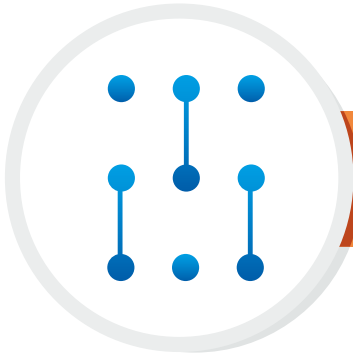
2 Dividing a speech signal into words

The next task for a voice-recognition system is to single out individual words in a continuous sonic flow and determine their meaning.

At its most basic level, speech can be seen as alternating sounds and silences. The silences can be understood as word 'separators'.

To analyze a speech signal in terms of words and word separators, an audio recording is first split into frames, i.e., small sections that are approximately 10 ms in length. These frames are not strictly consecutive: the end of one section is superimposed on the beginning of another.

Frame shift   Frame size
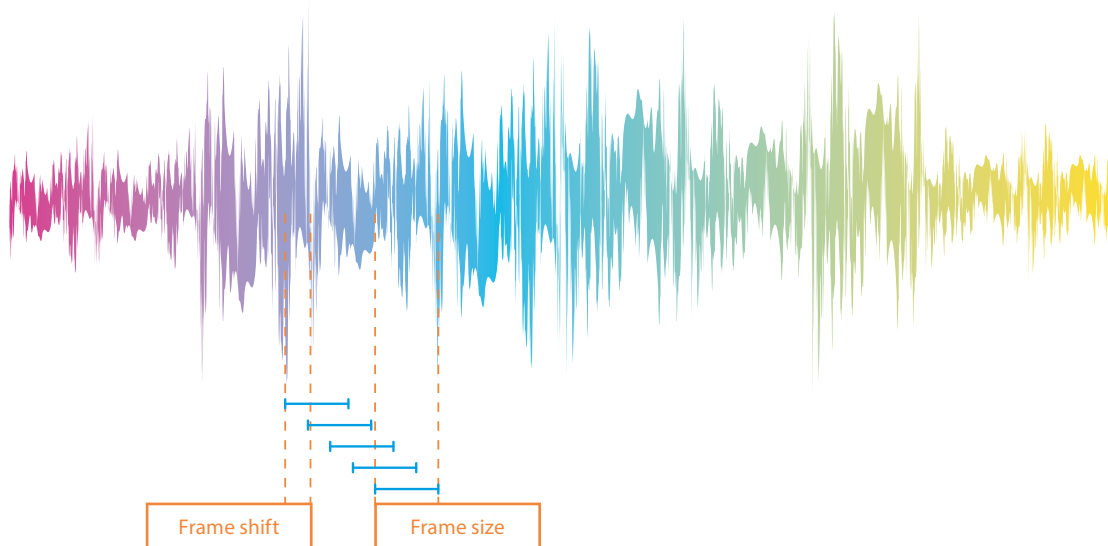
To determine which of the frames contain human voices, the system sets a limit. Values above the limit are considered to be words, while values below are understood as silence. There are several options for setting the limit value:

π Setting it as a constant (this constant can be used when sound is generated in the same way and under the same conditions).

{...} Defining a number of values that correspond to silence (if silence occupies a significant part of the recording).

Performing an entropy analysis (this requires determining how strongly the signal 'oscillates' within a given frame. The amplitude of oscillations for silent parts of a recording is usually lower).

SUBSCRIBE

Of the three, entropy analysis is considered the most reliable, although it has its flaws. For example, entropy may decrease when vowels are drawled or increase with a slight noise. To resolve the problem, two concepts a 'minimum distance between words' and a 'minimum word length', are introduced. The algorithm merges fragments that are too short and cuts off noise.

3 Interpretation of words

Most often, neural networks, combined with an apparatus containing hidden Markov models, are used to interpret words.

A-Z

Aa

0.8 E
0.15 O
0.05 A

Artificial brain

T — h — ?

Most probably, 'e'
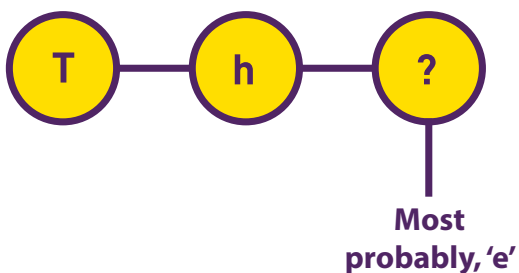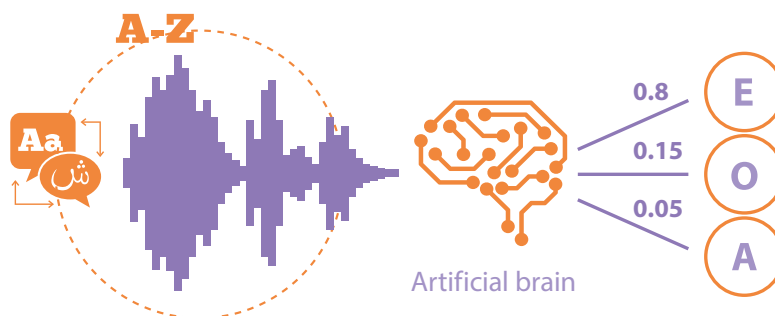
HIDDEN MARKOV MODELS

While researching early 20th century literary texts, mathematician Andrey Markov initially assumed the probability of the occurrence of a letter depended on the letter that preceded it. It turned out that this value remained constant in different parts of the same text.

Probability indicators are unique for each author. This makes it possible to use Markov models to spot plagiarism.
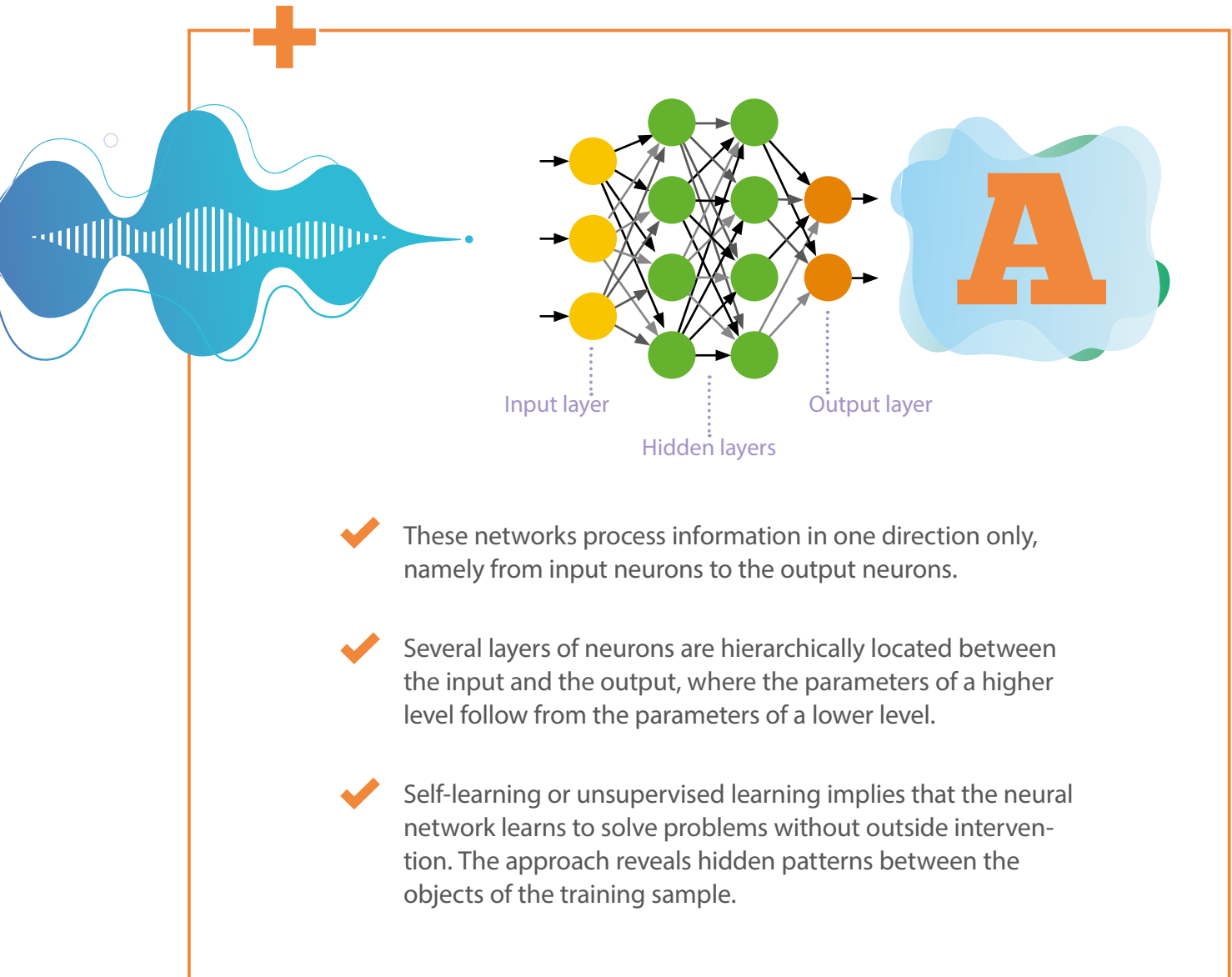
SUBSCRIBE

In Markov models, printed text consists of sequences of characters while speech is treated as a sequence of phonemes. While all of the symbols in written text are known, voice recordings contain the manifestation of phonemes and not the phonemes themselves (for example, there are several ways to pronounce the 'R' sound).

The device does not know which phoneme has been pronounced; it perceives only the parameters of the sound wave at a certain moment in time. In addition to estimating the probability of the occurrence of a particular phoneme, the system has to associate phonemes with appropriate signal variants.

FEEDFORWARD NEURAL NETWORKS

Until recently, self-learning neural networks with numerous layers were most often used in speech recognition.

Input layer   Output layer

Hidden layers

✓ These networks process information in one direction only, namely from input neurons to the output neurons.

✓ Several layers of neurons are hierarchically located between the input and the output, where the parameters of a higher level follow from the parameters of a lower level.

✓ Self-learning or unsupervised learning implies that the neural network learns to solve problems without outside intervention. The approach reveals hidden patterns between the objects of the training sample.

SUBSCRIBE

The result, representing a set of probabilities of the occurrence of a particular phoneme, is compared with the prediction of the Markov model. A pronounced sound is determined quite accurately.

RECURRENT NEURAL NETWORKS

Speech-recognition systems are gradually moving away from the use of simplified hidden Markov models. More and more, acoustic models are being built on recurrent neural networks, where internal memory and backpropagation are used for more efficient recognition.

Neurons not only receive information from the previous layer but also send the results of their own processing to themselves. This makes it possible to consider the order of data occurrence.

Input layer   Hidden layers   Output layer
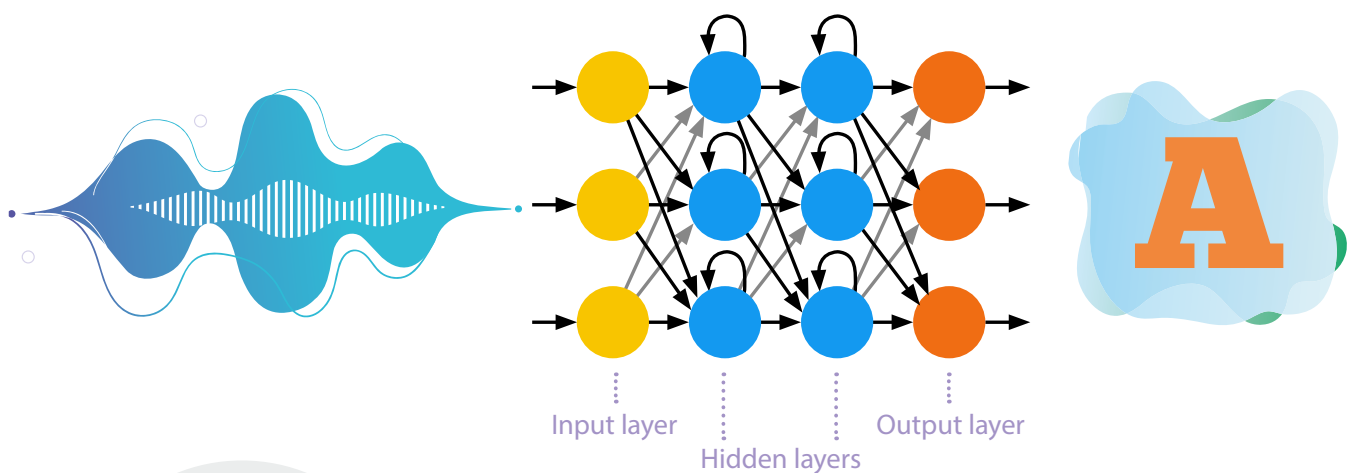
4 Phrasing

The principle of distinguishing phrases and sentences is very similar to word decoding.

Previously, N-gram type models were used for this task, where the probability of occurrence of a word depending on N previous words (usu. N = 3) was determined based on the analysis of large text blocks.

Deep learning and development of recurrent neural networks significantly improved the

I — have — been / bean

linguistic model and allowed it to take into account the context of what was said. The restriction on the use of only N previous words also disappeared.

Linguistic models were now able to guess the words missed or not recognised for a number of reasons. This turned out to be especially important for languages with random word order like Russian, where not only the previous words were important but the entire phrase.

This is how most speech-recognition systems work. But it is not enough to understand what was said. To be useful, the system must also be able to respond to incoming commands: it needs to answer questions, open apps, and manage other functions. Voice assistants are responsible for these tasks.

SPEECH RECOGNITION IN **MAG500A**

Included with MAG500A is a remote with Google Assistant voice control support. Voice control makes your interaction with the media player more convenient and user-friendly.

Here are the main features of the voice assistant:

✓ Search for films and information

✓ Answer questions (e.g., about the weather, traffic, or film cast)

✓ Playback control (app launch, volume adjustment)

✓ Interaction with peripherals

WHAT IS GOOGLE ASSISTANT?

Google Assistant is a virtual voice assistant first introduced at the Google I/O 2016 conference in California. Like Apple's Siri, Amazon's Alexa, and Microsoft's Cortana, the application provides contextual information at the user's request and is able to perform certain actions (such as entering search queries, setting reminders, opening applications, and controlling playback).

Google Assistant uses computer-assisted learning and Natural Language Processing (NLP) technology. The system is able to single out sounds, words, and ideas in speech.

The Assistant operates a billion devices and supports over thirty languages, but the Android TV version only speaks nine languages so far: English, French, German, Indonesian, Italian, Japanese, Korean, Portuguese, and Spanish.

HOW DOES GOOGLE ASSISTANT WORK?

First, the application records the speech it detects. Interpreting speech requires extensive computational power, so Google Assistant sends requests to Google data centres. When the sound data reaches one of these centres, the solid signal is divided into sounds. Google Assistant's algorithm searches a database of speech sounds and determines which words are the best match for the recorded combination of sounds.

The system then singles out the 'main' words from the user's statement and decides how to respond. For instance, if Google Assistant notices words like 'weather' and 'today', it will respond with today's weather forecast.

SUBSCRIBE

Google servers send the information back to the device, and the Google Assistant app performs the desired action or responds with a voice.

Google is changing the way Google Assistant works so that speech is recognised and commands are processed directly on the user's device. Using the capabilities of recurrent neural networks, the company has developed a new model for speech recognition and understanding. The size of the database of acoustic models has undergone a hundredfold reduction, so the artificial intelligence of Assistant can already work locally. The application processes speech in real-time and with almost zero delay, even without Internet access.

The new-generation Google Assistant responds to a request almost ten times faster. In 2019, Google's new Pixel smartphone models will support Google Assistant. In the future, the app will be available on other devices as well.

Today, Android TV's voice interface is available not only to companies with million dollar budgets, but also to local IPTV/OTT operators. This is a great opportunity for operators to attract new audiences, create content searches and accessing services easier and more convenient for users so they can stand out against their competitors.

Google and Android TV are trademarks of Google LLC.

SUBSCRIBE

MAG500A

4K-capable Android TV set-top box

by Infomir

Android 9.0
Android TV Platform

4K, HDR, HEVC
High-resolution content playback

Google Assistant
Voice Search

Dolby Digital Plus™ sound
8-channel surround audio

Chromecast built-in
Cast content to the big screen straight from your phone

Wi-Fi and Ethernet
Choose a connection type that best suits you

Learn more

EDITORIAL:

BROADVISION magazine is a quarterly online periodical devoted to the IPTV/OTT industry and modern technology produced by Infomir's marketing specialists.

We hope you'll enjoy reading our magazine as much as we enjoy working on it. Sign up for our newsletter and share this issue on social networking sites. It's the best way to show us that you like what we're doing.
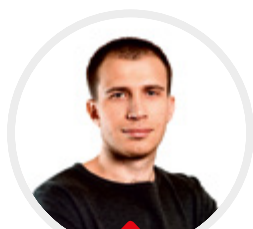
And feel free to let us know what you would like to see on the pages of *BROADVISION*.

E-mail your ideas and suggestions to us at **broadvision@infomir.com**.

Let's grow together!
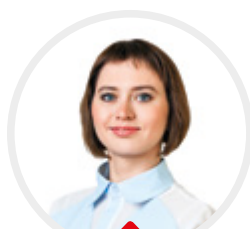
PRODUCTION TEAM

Project Manager: **Editor:**   **Design:**

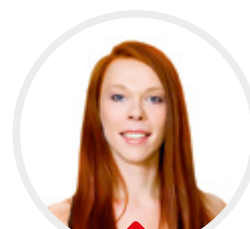Igor Oklander   Hennadii Mitrov   Kateryna Drozdovska   Kristina Usova

Contributors:

Marina Sluka   Anna Novikova

We would like to thank **Infomir Group** marketing team members.

As always, it's thanks to your collaborative efforts that we're able to get this issue out to our readers.

BROADVISION by im infomir

www.infomir.eu