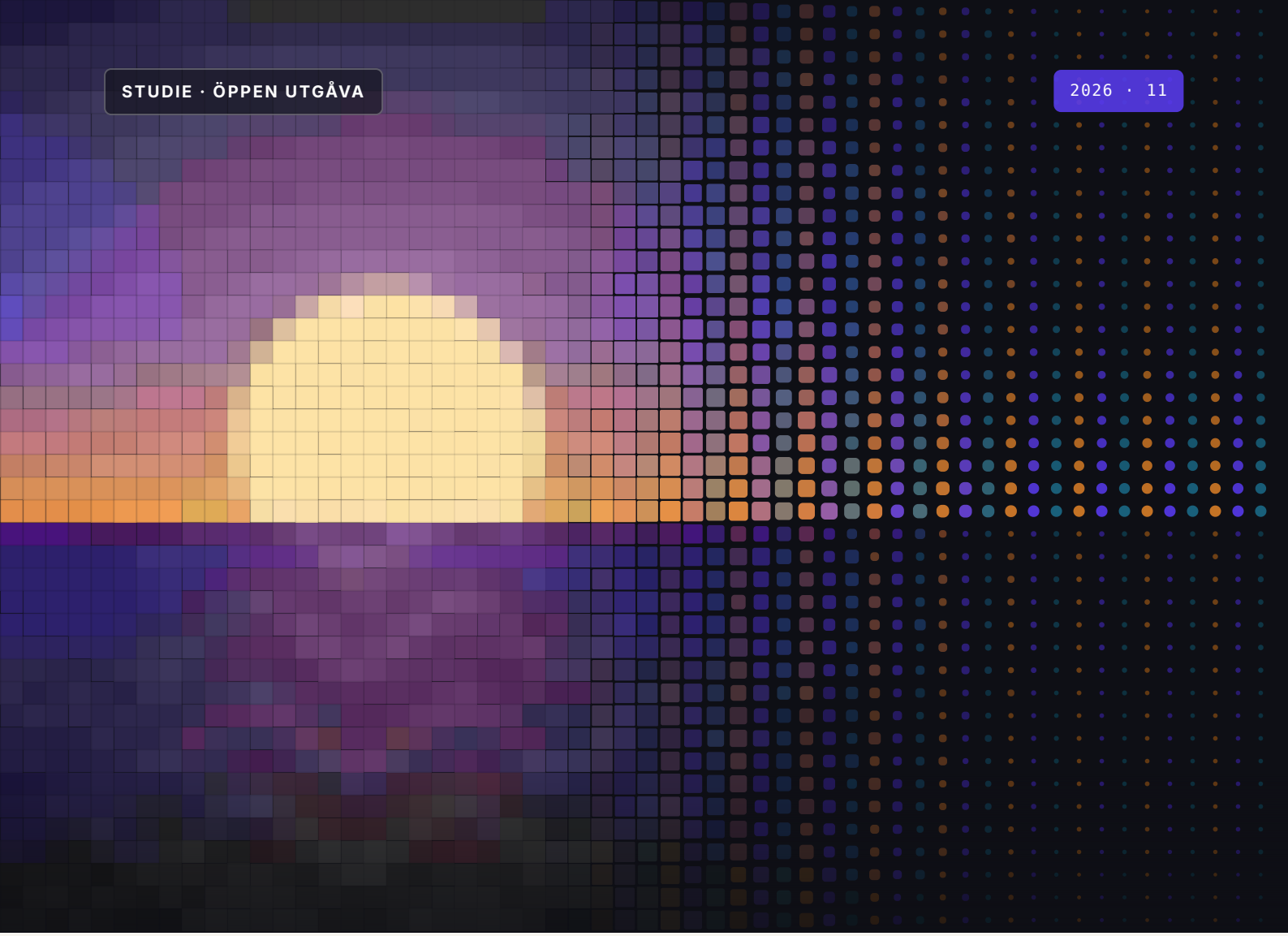


INLÄRD KOMPRIMERING · UPPSPELNING × KISEL

Neural kodning 2026

Den används runt codecen, inte i stället för den

Uppspelning · Fotografi · Samtal · Generativ AI · Spel · Kisel

SLUTSATSEN

Neurala nätverk arbetar redan i båda ändar av de codecar som bär världens video. Helt neurala codecar används bara i slutna system, och massmarknadens skärmar väntar på H.267.

Tan Shuai

Teknikexpert och grundare · AI-forskare och -utvecklare

tanshuai.com · i@tanshuai.com

Uppgifter per den 12 november 2026
Belopp i US-dollar

INNEHÅLL

Innehåll

◆ **Sammanfattning** — Redan etablerad, men runt codecen — 04

01 **Avgränsning och metod** — Vad ”neural kodning” betyder här, och hur jag gick till väga — 07

02 **Uppspelning och streaming** — Uppspelning och streaming: neuralt i båda ändar, standard i mitten — 08

03 **Fotografi och bilder** — Fotografi och bilder: neuralt i kameran, klassiskt på webben — 11

04 **Samtal och möten** — Samtal och möten: där helt neurala codecar faktiskt används — 12

05 **Generativ AI** — Generativ AI: den största utrullningen som ingen kallar en codec — 14

06 **Spel och grafik** — Spel och grafik: neural rekonstruktion är etablerad, neurala tillgångar är på väg in — 15

07 **Specialiserade länkar** — Specialiserade länkar: satelliter, drönare och robotar — 16

08 **Kisel** — Kiselperspektivet: energi, beräkning, minne och exakthet — 17

09 **Standarder och patent** — Standarder och patent: där neurala verktyg blir licensierbara — 21

10 **Vem som får betalt** — Vem som får betalt, och var ett företag kan satsa — 23

11 **Påståenden och belägg** — Påståenden du kommer att höra, och vad beläggen säger — 25

12 **Hållpunkter** — Att bevaka, 2027 till 2030 — 26

13 **Öppna frågor** — 27

◆ **Om författaren** — Tan Shuai — 28

◆ **Källor** — 29

FIGURER

Figurer

FIG. 01	Var neural kodning används 2026	06
FIG. 02	Tre vägar för neurala nätverk in i en codec	07
FIG. 03	Från färdig standard till massmarknadens skärmar går större delen av ett decennium	10
FIG. 04	I samtal slår en neural codec hårdvaru-HEVC med 65–76 %	13
FIG. 05	När ett företag styr hela stacken sprids neural rekonstruktion snabbt	15
FIG. 06	Energi per avkodad pixel: avkodare med fast funktion ligger en storleksordning före	18
FIG. 07	Ett neuralt verktyg kostar tiotusentals MAC per pixel; en hel neural avkodare kostar hundratusentals	19
FIG. 08	Uppdelningen i dagens produkter: hårdvara med fast funktion för bitströmmen, en neural motor för pixlarna	20
FIG. 09	Vem som får betalt för neural kodning 2026	23

Märkningar i texten

uppskattning — författarens beräkning eller slutsats utifrån citerade uppgifter

leverantörsuppgift — publicerat av företaget som säljer produkten, inte oberoende verifierat

analys — författarens tolkning av underlaget

Detta är en översättning av den engelska utgåvan, som gäller där versionerna skiljer sig åt.

Uppgifter per den 12 november 2026 · Belopp i US-dollar

SAMMANFATTNING

Redan etablerad, men runt codecen

SLUTSATSEN

Neural kodning är redan etablerad 2026, men den arbetar runt standardcodecarna, inte i stället för dem.

≈75 %
av den mobila datatrafiken är video, slutet av 2025 (Ericsson)

1,26 miljarder
smarttelefoner levererades 2025, och alla avkodar video i hårdvara med fast funktion

≈12×
energin per pixel hos ett uppmätt HEVC-avkodarchip med DRAM, för den enda publicerade chipkonstruktionen av en neural avkodare

≈22 TOPS
kontinuerlig beräkningskraft för en neural avkodare i DCVC-RT-klassen vid 1080p30 uppskattning

75,5 %
lägre bithastighet än hårdvaru-HEVC för MLVC vid 360p i samtal, på Microsofts egna testdata leverantörsuppgift

Slutet av 2029
H.267 version 1, den första standard som sannolikt för ut neurala verktyg till varje skärm

Var den används

Neurala nätverk körs nu i båda ändar av de codecar som bär nästan all video i världen. På vägen in formar de bilden: AI-baserade bildsignalprocessorer i telefonkameror och en neural nedskalare i Netflix kodningskedja. På vägen ut reparerar de den: YouTubes Super Resolution på tv-apparater, superupplösning hos Meta och tv-processorer som marknadsförs med hundratals neurala nätverk. Däremellan är bitströmmen fortfarande AV1, HEVC, AVC eller VVC, avkodad av hårdvara med fast funktion. Ingen streamingtjänst sänder ett neuralt videoformat till tv-apparater eller telefoner; där en neural videoavkodare körs på en enhet i dag, körs den inuti ett enskilt företags egen app.

Slutna system

Helt neurala codecar, där både kodaren och avkodaren är inlärda, används bara i slutna system, där ett företag styr båda ändarna och kan falla tillbaka på en standardcodec. Microsoft har kört en AI-baserad ljudcodec i alla tvåpartssamtal i Teams och Skype sedan 2021, och i juli 2026 släppte företaget MLVC som öppen källkod, en neural videocodec som nu rullas ut i peer-to-peer-samtal i Teams: vid 360p behöver den omkring 75 % lägre bithastighet än hårdvaru-HEVC för samma bedömda kvalitet. Varje modell för tal- och videogenerering körs på en inlärd codec, från taltokeniserare på 12,5 Hz till video-VAE:er med kompressionsförhållandet 1:192. Neural texturkomprimering har kommit med i ett AAA-spel. Det här är verkliga driftsättningar, och de har ett villkor gemensamt: ingen annan behöver avkoda strömmen.

Tidsplan

Uppspelning av neurala bitströmmar på massmarknaden väntar på en standard och en kiselcykel. Den första standarden som sannolikt kommer att föra ut neurala verktyg, VVC:s efterföljare som jag kallar H.267, utfärdade sin begäran om förslag i juli 2026, utvärderar svaren i januari 2027 och siktar på en första version i slutet av 2029. AV1 behövde nästan åtta år från färdig

specifikation till 30 % av tittandet på Netflix. Enligt det mönstret kommer de första H.267-chipen omkring 2031 och massmarknaden i mitten av 2030-talet uppskattning. Under tiden är den codec som når skärmarna härnäst AV2, färdig i maj 2026 utan någon neural avkodare, och tre IP-leverantörer presenterade kunder för AV2-avkodare inom tre veckor.

Kisel

Skälet finns i kislet. Den enda publicerade chipkonstruktionen av en neural videoavkodare behöver 14,6 nJ per pixel, omkring 12 gånger så mycket som ett uppmätt HEVC-avkodarchip räknat med sitt DRAM och omkring 47 gånger så mycket som enbart dess kärna. Microsofts realtidscodec för forskning, DCVC-RT, behöver enligt min räkning av dess egen tabell per modul omkring 175 tusen MAC-operationer per avkodad pixel, eller ungefär 22 TOPS kontinuerlig beräkningskraft för 1080p30 uppskattning: omkring hälften av de 40 TOPS som definierar en AI-dator, för en enda videoström. En neural avkodare måste dessutom ge bitidentiska resultat på varje enhet, annars faller entropiavkodningen sönder. Därför byter MLVC bort komprimering för att avkoda identiskt på olika NPU:er, och körs i 540p, inte 4K.

Affärer

För ett företag följer pengarna standarderna och plattformarna, inte fristående neurala codecar. Plattformarna behåller besparingarna från neural för- och efterbearbetning. IP-leverantörer licensierar avkodare med fast funktion för varje ny standard, plus NPU:er för bildbehandling. Patentpooler tar betalt per enhet: 0,50 dollar per telefon och 1,00–1,50 dollar per 4K-tv för VVC i USA och andra marknader i Region 1. De tre mest kända startupbolagen inom neurala codecar, WaveOne, iSIZE och Deep Render, köptes av Apple, Sony och InterDigital. De öppna möjligheterna finns i neural för- och efterbearbetning, i produkter för slutna system, i kisel som kombinerar en codec med fast funktion med en NPU, och i utvärderings- och standardiseringsarbetet däremellan.

Fråga	Bedömning	Nyckeltal (källor i kapitlen)
① **Var används neural kodning i stor skala i dag?**	Runt codecen: kamera-ISP:er, nedskalning på kodarsidan, superupplösning vid uppspelning. Helt neuralt bara i slutna system: samtal, speltillgångar, tokeniserare för generativ AI	Video ≈75 % av mobildatan; 1,26 miljarder telefoner och ≈200 miljoner tv-apparater levererade 2025, alla med avkodning i hårdvara med fast funktion; DLSS 4 i över 250 spel och appar
② **Kommer neurala codecar att ersätta AV1, HEVC och VVC i uppspelning? När?**	Inte före en standard och en enhetscykel. H.267 är den första realistiska bäraren, och den kommer sannolikt att blanda klassiska och neurala verktyg	H.267: begäran om förslag juli 2026, utvärdering januari 2027, version 1 slutet av 2029; massmarknad mitten av 2030-talet uppskattning
③ **Vad kostar det i kisel?**	En storleksordning mer energi per pixel än avkodning med fast funktion i dag, plus minnestrafik och ett krav på bitexakthet	14,6 nJ/pixel (neural) mot 1,19 (HEVC, med DRAM); ≈175 kMAC per avkodad pixel ≈ 22 TOPS vid 1080p30 uppskattning
④ **Vad tillåter standarder och patent?**	Neurala efterfilter kan redan signaleras i VVC-strömmar; JPEG AI är en ISO-standard; H.267 avgör vilka neurala verktyg som blir licensierbara	ECM (klassisk) ≈28 % bättre än VVC; neurala loopfilter 6–14 % vid 10–22+ kMAC/pixel; ByteDance 645 patentfamiljer i arbetet bortom VVC
⑤ **Vem får betalt, och var kan ett företag satsa?**	Plattformar, IP-leverantörer, patentinnehavare. Rena startupbolag inom neurala codecar köptes upp. Satsa på för- och efterbearbetning, slutna system, kisel med codec plus NPU, utvärdering	VVC-poolen 0,50 dollar per telefon; avkodar-IP för AV2 licensierad inom veckor efter specifikationen; tre förvärv, villkoren inte offentliga
⑥ **Vilka påståenden bör man vara skeptisk till?**	Vinster som anges mot gamla codecar, vid låg upplösning, utan avkodningskostnad eller från demonstrationer i stället för levererade produkter	MLVC: 88 % lägre än H.264 men 65–76 % lägre än hårdvaru-HEVC; texturer: 85 % VRAM-besparing i en demo, omkring 30 % i det utgivna spelet

FIG. 01

Var neural kodning används 2026

Neurala nätverk körs redan i båda ändar av de etablerade codecarna. Helt neurala codecar används bara där ett och samma företag styr både kodaren och avkodaren.

	Runt en standardcodec förbehandling, förbättring, uppskalning	Helt neural codec inlärd kodare och avkodare	Inuti en standard neurala verktyg i en standardbitström eller ett API
01 **Uppspelning och streaming**	I STOR SKALA Netflix nedskalare; superupplösning hos YouTube och Meta; AI-uppskalare i tv-apparater	INTE I BRUK Ingen streamingtjänst sänder ett neuralt videoformat; tv-apparater och telefoner avkodar standardcodecar i hårdvara	STANDARDISERAS H.267 version 1 väntas i slutet av 2029; neurala verktyg finns bland kandidaterna
02 **Fotografi och bildbehandling**	I STOR SKALA AI-ISP:er i flaggskeppstelefoner; en helt neural ISP visad som IP på CES 2026	DEMO, PILOT JPEG AI är en ISO-standard: demonstrationer i telefoner, ingen produkt	INTE I BRUK Formatet som vinner mark är klassiskt: JPEG XL förvalt påslaget i Chrome 155
03 **Samtal och möten**	I STOR SKALA Superupplösning på mottagarsidan och neural brusdämpning	I STOR SKALA Neuralt ljud i tvåpartssamtal i Teams sedan 2021; MLVC-video rullas ut i Teams	I BRUK, BEGRÄNSAT Opus 1.5 lägger till valfria, bakåtkompatibla ML-verktyg
04 **Generativ AI**	—	I STOR SKALA Varje video- och röstmodell kör en inlärd tokeniserare, t.ex. en video-VAE med 1:192	—
05 **Spel och grafik**	I STOR SKALA Neural uppskalning och bildrutegenerering: enbart DLSS 4 i över 250 spel och appar	I BRUK, BEGRÄNSAT Neurala texturer i ett utgivet AAA-spel; Nvidias SDK är i betaversion	STANDARDISERAS DirectX får matrisoperationer för shaders (förhandsversion)
06 **Specialiserade länkar: satellit, robotar**	DEMO, PILOT Återställning av flöden med låg bithastighet på serversidan	DEMO, PILOT Förlusttåliga neurala codecar i forskning och piloter	—

Källa: Dokument från företag, standardiseringsorgan och utvecklare som citeras i rapporten; läget den 12 november 2026. Raderna är ordnade efter hur många människor varje område når.

AVGRÄNSNING OCH METOD

01 Vad ”neural kodning” betyder här, och hur jag gick till väga

Tre vägar för neurala nätverk in i en codec

”Neural kodning” täcker tre olika saker, och de flesta meningsskiljaktigheter om ämnet kommer av att de blandas ihop.

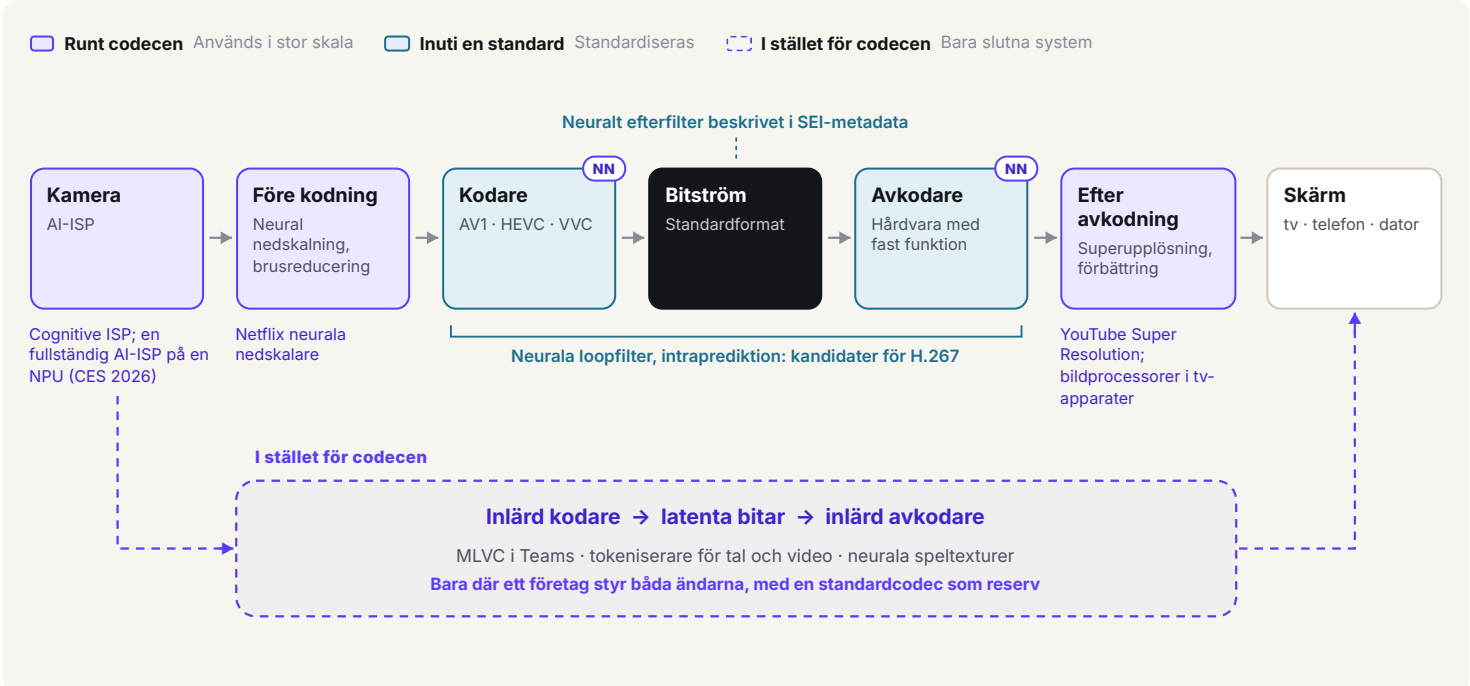
01 **Runt codecen.** Ett nätverk bearbetar pixlar före kodningen eller efter avkodningen, och bitströmmen förblir standard. Exempel: en AI-baserad bildsignalprocessor (ISP) i en kamera, Netflix neurala nedskalare före kodaren, superupplösning i en tv. Vilken enhet som helst kan spela upp strömmen; nätverket ändrar bara vad som går in eller vad som kommer ut.

02 **Inuti en standard.** Ett nätverk blir ett verktyg i en standardcodec, till exempel ett neuralt loopfilter eller en intraprediktor som varje avkodare måste köra, så det måste specificeras exakt, vanligen i heltalsaritmetik. En lättare variant är ett neuralt efterfilter som bitströmmen beskriver i metadata och som avkodare kan välja att tillämpa.

03 **I stället för codecen.** Både kodaren och avkodaren är inlärda, från ände till ände. Bitströmmen är en uppsättning kvantiserade latenta värden som bara det matchande nätverket kan avkoda. Exempel: Microsofts MLVC i Teams, taltokeniserare i röstmodeller, neurala texturer i ett spel.

FIG. 02 **Tre vägar för neurala nätverk in i en codec**

Det mesta som används 2026 arbetar runt en standardcodec. Neurala verktyg inuti en standard håller på att specificeras. Helt neurala codecar ersätter codecen bara där ett och samma företag styr båda ändarna.



Källa: Författarens sammanställning av källorna i rapporten; läget den 12 november 2026.

Den första typen kräver ingens tillstånd. Den andra kräver en standard. Den tredje kräver kontroll över båda ändarna. Det är i den ordningen de når människor, och det förklarar det mesta av det som följer.

Ordnat efter den verkliga världen, inte efter forskningsagendan

Forskningsartiklar studerar mest den tredje typen, eftersom det är där komprimeringsrekorden finns. Den här rapporten ordnar områdena efter hur många människor de når, eftersom det avgör vart pengar och ingenjörsarbete går: uppspelning och streaming först, sedan fotografi, samtal, generativ AI, spel, och de specialiserade länkarna (satelliter, drönare, robotar) sist. Video utgör omkring 75 % av all mobil datatrafik, enligt Ericssons räkning i slutet av 2025. Under 2025 levererade branschen omkring 1,26 miljarder smarttelefoner (IDC) och ungefär 196–209 miljoner tv-apparater (uppskattningar från TrendForce och Omdia). Var och en av dem avkodar video i hårdvara med fast funktion.

Hur jag arbetade

Det här är en skrivbordsstudie, gjord på samma sätt som jag skulle granska ett chip eller en produktplan. Jag läste standarddokument och mötesrapporter (JVET, MPEG, JPEG, AOMedia, IETF), företags teknikbloggar och versionsinformation, chipartiklar, patentpoolers prislistor, texter om exportkontroll och pressen. Varje siffra anges med datum. Där jag själv härleder en siffra, till exempel beräkningsbehov per pixel ur en artikels tabeller, säger jag det och märker den **uppskattning**. Påståenden som jag inte kunde kontrollera mot en primärkälla är märkta **leverantörsuppgift**. Belopp anges i US-dollar. Läget gäller per den 12 november 2026.

Jag jämför lika med lika där jag kan, och säger till där jag inte kan. Komprimeringsvinster betyder lite om de inte åtföljs av fyra saker: jämförelsebasen (vinst jämfört med vad), måttet (PSNR, VMAF eller en panel av tittare), innehållet och upplösningen, och vad det kostar att avkoda. Det mesta av oenigheten om neurala codecar försvinner när de fyra har angetts.

KAPITEL 02

UPPSPELNING OCH STREAMING

02 Uppspelning och streaming: neuralt i båda ändar, standard i mitten

Det tittarna tar emot är fortfarande en standardbitström

De största videotjänsterna har avslutat ett codecskifte och påbörjar nästa, och inget av dem involverar en neural avkodare. I december 2025 rapporterade Netflix att AV1 stod för omkring 30 % av tittandet, tjänstens näst mest använda codec efter AVC. AV1 nådde tv-apparaterna först när hårdvaruavkodarna gjorde det: Netflix lade till AV1 i sin enhetscertifiering 2019, streamade det till tv-apparater från slutet av 2021 och uppger att 88 % av de enheter med stor skärm som lämnades in för certifiering mellan 2021 och 2025 hade stöd för det, nästan alla sedan 2023. YouTube uppgav att mer än 90 % av katalogen, viktat efter visningstid, fanns tillgänglig i AV1 per den 19 mars 2026.

AV2, AV1:s efterföljare från Alliance for Open Media, nådde version 1.0.0 den 28 maj 2026, med omkring 30 % bättre komprimering än AV1 enligt AOMedias objektiva mätningar. YouTube planerar AV2-strömmar på webben ”redan i början av 2027”. Kislet följde direkt: Chips&Media färdigställde en AV2-avkodare i hårdvara i maj och licensierade den i juni till ett icke namngivet stort nordamerikanskt teknikföretag; Allegro DVT hade levererat sin

AV2-avkodar-IP till tre kunder den 16 juni; VeriSilicon hade släppt en video-IP med stöd för AV2 till flera kunder den 10 juni. AV2 är en klassisk codec: vinsterna kommer från bättre partitionering, prediktion, transformer och loopfilter (ett av dem ett ”datadrivet” styrt detaljfilter), inte från ett nätverk från ände till ände.

VVC, den nuvarande standarden från ITU och ISO, fick sitt första hem i tv-sändningar i Brasilien, vars system DTV+ (TV 3.0) togs i kommersiell drift i juni 2026, lagom till fotbolls-VM. I juli 2026 fanns ingen tv till salu i Brasilien med inbyggd DTV+-mottagare; de första väntas mellan slutet av 2026 och början av 2027.

Vid kodaren: Netflix neurala nedskalare

Det renaste exemplet på neural kodning i uppspelning har körts sedan 2022 och ändrar ingenting av det som en tittares enhet måste göra. Netflix kodar varje titel i flera upplösningar, och nedskalningssteget som tar fram 1080p, 720p och 540p ur en 4K-master använde traditionellt ett fast filter som Lanczos. Netflix ersatte det med ett litet neuralt nätverk, ett förbehandlingsblock följt av ett block för storleksändring, tränat för att förbättra den kvalitet som syns efter en vanlig uppskalare i enheten. Netflix beskriver det som en lösning som kan sättas in direkt: ingen annan ändring på kodningssidan eller på klientsidan, så att ”miljontals enheter som stöder Netflix-streaming automatiskt får nytta av det”. I Netflix preferenstester valde omkring 77 % av tittarna den neurala nedskalarens resultat leverantörsuppgift. För att hålla kostnaden rimlig i Netflix skala räckte det med ”bara några få lager i det neurala nätverket”.

Det är mönstret att komma ihåg. Nätverket körs en gång per titel i datacentret, där beräkningskraft är billig och kan schemaläggas, och resultatet är en vanlig AV1- eller HEVC-ström.

Vid skärmen: superupplösning i appen och i tv:n

I andra änden reparerar nätverk det som codecen och källan förlorade. Den 29 oktober 2025 presenterade YouTube Super Resolution för sin tv-app: videor som laddats upp under 1080p, från 240p till 720p, skalas upp till HD med maskininlärning, med 4K som uttalat mål. De uppskalade versionerna är märkta, tittarna kan växla tillbaka till originalet och kreatörer kan välja bort funktionen. Meta beskrev i oktober 2025 hur företaget kör videosuperupplösning både på sina servrar vid uppladdning, på vanliga x86-processorer i samarbete med Intel, och i klienten vid uppspelning, och fann att de två kompletterar varandra. Tv-tillverkarna gör samma sak i kisel: Samsungs flaggskeppsprocessor för 8K, NQ8 AI Gen3, marknadsförs med 768 neurala nätverk för uppskalning.

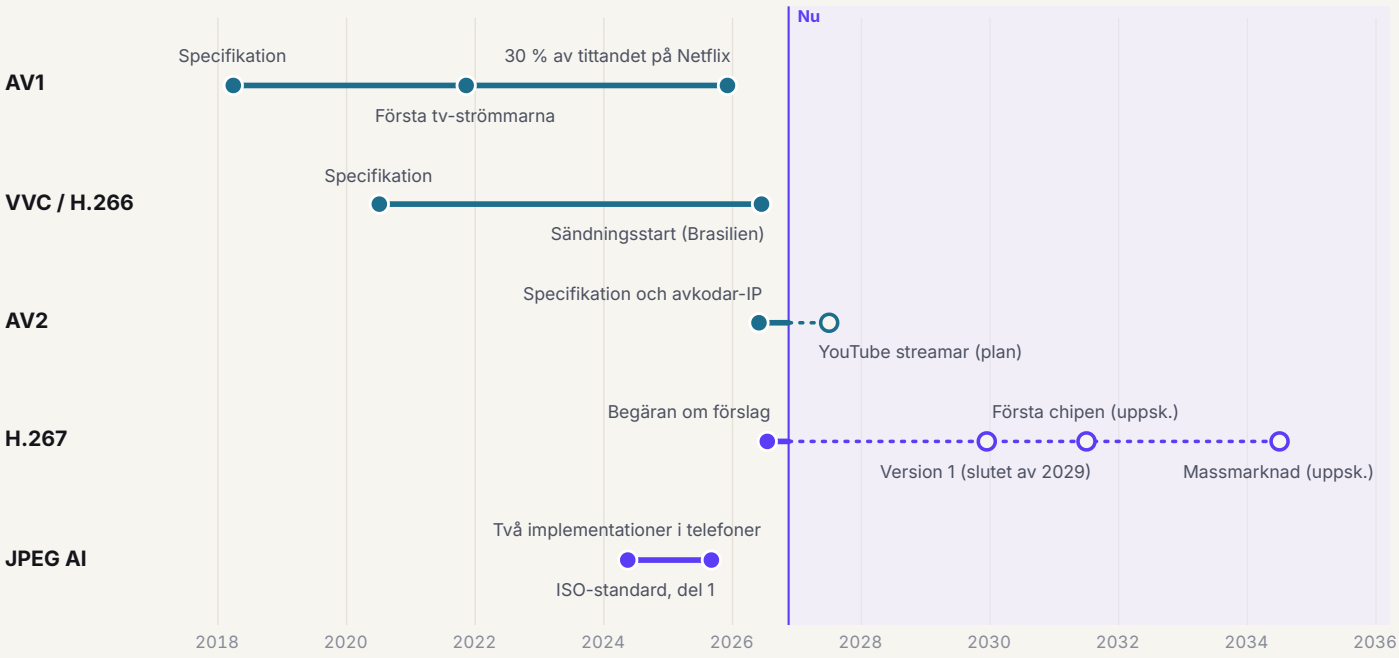
Ekonomin skiljer sig åt beroende på var nätverket körs. Bearbetning på serversidan kostar plattformen beräkningskraft för varje video den hanterar. Bearbetning i enheten kostar plattformen ingenting; den använder en NPU som köparen redan har betalat för. Det är därför de största vinsterna 2026 kommer från båda.

Varför ingen tv avkodar en neural bitström

Ett uppspelningsformat måste kunna avkodas på varje enhet som någonsin ska spela upp det, och de enheterna håller i åratal. Det gör att vägen från färdig standard till massmarknadens skärmar tar lång tid även för klassiska codecar. AV1:s specifikation blev klar i mars 2018; mjukvaruavkodaren dav1d kom inom några månader; tv-apparater streamade det från slutet av 2021; det nådde omkring 30 % av tittandet på Netflix i december 2025, nästan åtta år efter specifikationen. VVC blev klar i juli 2020 och nådde sin första sändningstjänst i juni 2026.

FIG. 03 **Från färdig standard till massmarknadens skärmar går större delen av ett decennium**

Uppspelning kräver en hårdvaruavkodare i enheten. AV1 behövde nästan åtta år från färdig specifikation till 30 % av tittandet på Netflix; den första standarden som sannolikt för ut neurala verktyg väntas i slutet av 2029.



Källa: Netflix Tech Blog (dec. 2025); AOMedia; Allegro DVT; TV Technology; Fraunhofer IIS om Brasiliens DTV+; MPEG-155; JPEG-kommittén; ISO. Ofyllda markeringar är mål eller författarens uppskattningar.

En neural bitström tillför två problem som klassiska codecar löste för länge sedan: energi per pixel och bitexakt avkodning. Båda behandlas i kapitel 8. Tillsammans innebär de att ingen tillverkare av tv-apparater eller telefoner kommer att bygga in en neural videoavkodare i varje enhet innan en standard talar om exakt vad som ska byggas.

Vad som skulle ändra det: H.267

Standarden som skulle kunna föra ut neurala verktyg till varje skärm är VVC:s efterföljare, som utvecklas gemensamt av ITU-T och ISO/IEC i Joint Video Experts Team (JVET). Namnet är inte slutgiltigt; jag kallar den H.267, som stora delar av branschen gör.

I oktober 2025 utvärderade JVET svaren på en begäran om belägg: fem svar från 16 företag och forskningslabb, med 20 testade codec-konfigurationer. Kommittén fann belägg för bättre komprimering än VVC och bedömde både förbättrad konventionell teknik och teknik baserad på neurala nätverk som ”intressanta”. Neurala metoder visade tydliga perceptuella vinster, särskilt på 8K HDR-innehåll, och mindre vinster på spelinnehåll. InterDigital, en av de svarande, lämnade in en konstruktion baserad på den konventionella Enhanced Compression Model (ECM) och en som lade AI-verktyg ovanpå VVC.

Den formella begäran om förslag utfärdades vid mötet i Genève i juli 2026. Huvudbidragen skulle lämnas in i slutet av oktober 2026, ett kompletterande paket ska lämnas i mitten av december, och svaren utvärderas vid JVET-mötet i Brisbane i januari 2027. JVET räknar med en första testmodell under andra halvåret 2027 och en första version av standarden senast i slutet av 2029, och kallar arbetet ”förmodligen en större insats än någonsin tidigare för att dra i gång en ny standard”.

Det klassiska spåret sätter ribban högt. ECM, byggd med konventionell signalbehandling, sparar ungefär 28 % jämfört med VVC i random access-konfiguration. JVET:s utforskning av neurala nätverk redovisar loopfilter som sparar omkring 6 % vid 10 tusen MAC-operationer per pixel, 8 % vid 22 tusen och 14 % vid en driftpunkt med hög

komplexitet. Min tolkning är att H.267 med störst sannolikhet blir en klassisk kärna med några neurala verktyg där de sparade bitarna motiverar avkodarkostnaden, och att utvärderingen i januari 2027 kommer att visa hur många analys.

Det finns redan en brygga. VVC- och HEVC-strömmar kan bära kompletterande förbättringsinformation (SEI) som beskriver ett neuralt efterfilter för avkodarsidan att tillämpa, och de nya utgåvorna av VVC och av dess SEI-standard (VSEI), färdiga i oktober 2025, utökade de meddelandena. En ström kan ange vilket nätverk som ska köras efter avkodningen utan att tvinga varje avkodare att köra det.

Vad det här betyder för företag inom uppspelning

Sälj det som fungerar runt codecen nu: förbehandling på kodarsidan för tjänster, och block och finjustering för efterbearbetning i chip för tv-apparater och digitalboxar. Bygg inte en uppspelningsaffär på en neural bitström som måste avkodas på tredjepartsenheter före 2030-talet. Om du tillverkar kisel är intäkterna på kort sikt AV2- och VVC-avkodare; den neurala delen är en NPU-väg för förbättring vid sidan av dem (kapitel 8).

KAPITEL 03

FOTOGRAFI OCH BILDER

03 Fotografi och bilder: neuralt i kameran, klassiskt på webben

Kamerakedjan är redan neural

Varje foto från en nyare flaggskeppstelefon har passerat neurala nätverk innan det komprimeras. Qualcomms Snapdragon 8 Gen 2, i telefoner från 2023, introducerade en ”Cognitive ISP” som kör semantisk segmentering i realtid vid foto- och videoinspelning och justerar ansikten, hår, himmel och bakgrund var för sig; Samsungs Galaxy S23 Ultra var den första telefonen som använde den i selfiekameran. I januari 2026 visade Chips&Media och Visionary.ai på CES vad de kallar den första fullständiga AI-baserade ISP:n: en bildkedja från ände till ände som körs helt på en NPU, Chips&Medias WAVE-N, i 16-bitars flyttal, utan en ISP med fast funktion, och som kan uppdateras trådlöst. ”Befintliga kedjor skruvar fast neurala block på en ISP med fast funktion. Vi ersätter den konventionella ISP:n helt”, sade Visionary.ai:s vd. Den erbjöds tillverkare för utvärdering efter mässan.

Sedan sparas resultatet som JPEG eller HEIF, klassiska format som vilken enhet som helst kan öppna. Mönstret är detsamma som i uppspelning: neuralt inom ett företags egen kedja, standard vid gränsen.

JPEG AI: en färdig neural standard utan produkt ännu

JPEG AI är den första standarden för inlärd bildkodning från ISO/IEC och ITU-T (ISO/IEC 6048, ITU-T T.840). Del 1, kärnkodningssystemet, publicerades 2025. JPEG-kommittén redovisar komprimeringsvinster på 12,5 % till 27,9 % jämfört med VVC:s intrakodning, över ett spann av avkodarkomplexitet från 7 till 216 tusen MAC-operationer per pixel. Konstruktionen har två kodare och tre avkodare, tränade gemensamt, så att en mottagare kan implementera bara en avkodare och ändå läsa varje ström. I maj 2024 rapporterade kommittén två implementationer i smarttelefoner. Parallellt publicerade IEEE i december 2024 IEEE 1857.11-2024, en standard för bildkodning baserad på neurala nätverk och utformad för NPU:er.

I november 2026 hittade jag ingen telefonkamera, fotoapp eller webbläsare som skriver eller läser JPEG AI som förval.

Vad webben valde i stället: JPEG XL

Det bildformat som faktiskt vann mark 2026 är klassiskt. Den 6 oktober 2026 slog Google på avkodning av JPEG XL som förval i Chrome 155, med en ny minnessäker avkodare skriven i Rust, fyra år efter att företaget tog bort det experimentella stödet. Safari har avkodat JPEG XL sedan 2023, och Mozilla har rört sig åt samma håll. Googles argument är bättre komprimering än JPEG, förlustfri omkomprimering av befintliga JPEG-filer, HDR och progressiv avkodning.

Lärdomen gäller generellt. Ett format som måste gå att öppna överallt vinner med en snabb, deterministisk avkodare som körs på vilken CPU som helst och kan göras minnessäker. En neural bildavkodare behöver en NPU eller GPU och exakt aritmetik på alla enheter. Neural kodning vinner inom en enhets egen kedja och i system där samma företag kodar och avkodar, till exempel en molntjänst för foton som aldrig lämnar ut den komprimerade filen till någon annan.

Vad det här betyder för företag inom bildbehandling

Pengarna finns i kamerakedjan: IP för AI-ISP, NPU:er dimensionerade för bildbehandling och det finjusteringsarbete som kameratillverkare köper. För lagring och distribution är en inlärd bildcodec i dag ett verktyg för slutna system, inte ett format att satsa på.

KAPITEL 04

SAMTAL OCH MÖTEN

04 Samtal och möten: där helt neurala codecar faktiskt används

Ljudet kom först

Microsoft presenterade Satin, en AI-driven talcodec, i februari 2021 och sade att den redan användes för alla tvåpartssamtal i Teams och Skype. Satin levererar superbredbandigt tal, upp till 16 kHz ljudbandbredd, vid 6 kbps, där Skypes tidigare codec Silk var smalbandig vid samma bithastighet. Den bithastighet som sparas läggs på redundans, så att samtalen klarar paketförluster bättre.

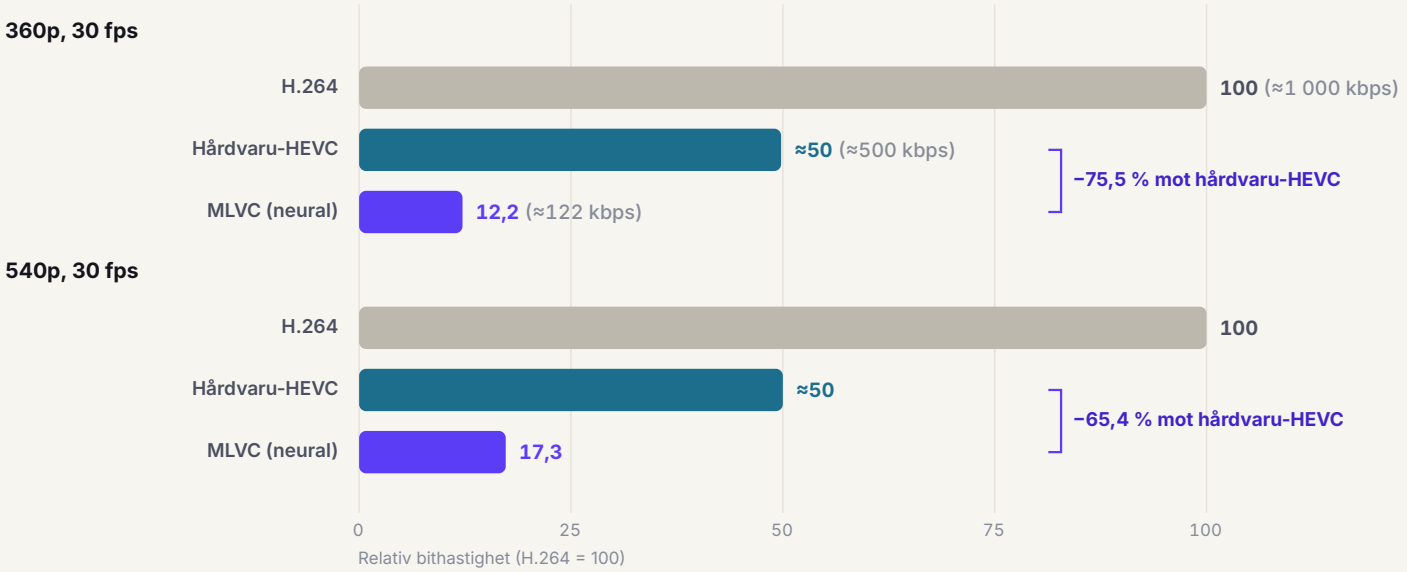
MLVC: en neural videocodec i samtal i skarp drift

Den 24 juli 2026 släppte Microsoft MLVC, sin videocodec baserad på maskininlärning, som öppen källkod under MIT-licensen, med modellkod, tränade vikter, träningsskript och de konverteringsverktyg som riktar sig mot olika NPU:er. MLVC är produktversionen av den DCVC-familj som Microsoft Research har publicerat sedan 2021. Microsoft uppger att MLVC rullas ut i Teams för peer-to-peer-samtal, med telemetri i realtid, A/B-tester och de konventionella codecarna som reserv.

MLVC körs i 540p och 360p med 30 bilder per sekund och är konstruerad för att använda högst halva NPU:n på Apples Neural Engine och på NPU:er från Qualcomm och Intel, med arbetet uppdelat mellan CPU och NPU. I subjektiva tester enligt ITU-T P.910 på Microsofts eget Video Conferencing Dataset behövde den 87,8 % lägre bithastighet än H.264 och 75,5 % lägre än hårdvaru-HEVC vid 360p, och 82,7 % respektive 65,4 % lägre vid 540p leverantörsuppgift. Vid 360p motsvarar det ungefär 122 kbps där H.264 behöver omkring 1 Mbps.

FIG. 04 **I samtal slår en neural codec hårdvaru-HEVC med 65–76 %**

Bithastighet som krävs för samma bedömda kvalitet på Microsofts testdata för videokonferenser, H.264 = 100. MLVC körs på NPU:er i bärbara datorer och telefoner med upp till 540p30 och använder högst halva NPU:n.



Källa: Microsoft, lanseringen av MLVC (24 juli 2026): subjektiva tester enligt ITU-T P.910 på företagets Video Conferencing Dataset. HEVC-staplarna är härledda ur de rapporterade besparingarna. Leverantörsuppgifter, inte oberoende mätningar.

Två saker nyanserar rubriksiffran. 88 % gäller mot H.264; mot den hårdvaru-HEVC som samma bärbara datorer redan innehåller är det 65–76 %, fortfarande mycket. Och MLVC gav upp komprimering för att kunna driftsättas: en oberoende analys noterar att en perceptuellt tränad DCVC-RT når 81,9 % på samma mått där MLVC når 75,5 %. Priset för att avkoda identiskt på hårdvara som Microsoft inte äger var ungefär sex procentenheter.

Varför samtal är det naturliga hemmet

Ett samtal är ett slutet system. Båda ändpunkterna kör samma program, förhandlar om vad de klarar i början av samtalet och kan falla tillbaka på en standardcodec mitt i det. Upplösningarna är måttliga, näten är ofta dåliga, och de bärbara datorerna i båda ändar levereras nu med NPU:er, liksom telefoner har gjort i flera år. Microsoft sätter 40 TOPS för NPU:n som ribba för sina AI-datorer, och Apples M4 Neural Engine anges till 38 TOPS. Alla de förhållanden som stoppar en neural codec i uppspelning saknas här.

Motexemplen: Opus och MLow

Alla samtalsplattformar gjorde inte samma val, och skälen är lärorika.

Opus, den öppna ljudcodecen i WebRTC och de flesta samtalsappar, lade till djupinlärning i version 1.5 (mars 2024), första gången nätverk användes för att bearbeta eller generera själva signalen. Men de lades till som tillval: förbättring på avkodarsidan (LACE och NoLACE) som inte kräver någon ändring av bitströmmen, en neural maskering av paketförluster och ett system för djup redundans (DRED); alla är avstängda som förval så att gamla enheter inte belastas. Opus 1.6 lade till en neural bandbreddsutvidgning och en DRED-modell som är tre gånger mindre, och är fortfarande helt kompatibel med standarden från 2012. Angreppssättet är ”inuti en standard, valfritt och bakåtkompatibelt”.

Meta gick längre tillbaka. I juni 2024 lanserade företaget MLow, en talcodec för låg bithastighet byggd på klassisk signalbehandling i stället för ett nätverk, som nu används i alla samtal i Instagram och Messenger och rullas ut i WhatsApp. Metas skäl var dess enheter: mer än 20 % av företagets samtal görs på ARMv7-telefoner, och tiotals

miljoner WhatsApp-samtal om dagen sker på telefoner som är mer än tio år gamla. Neurala codecar, inklusive Metas egen EnCodec, fungerar tillförlitligt bara på dyra telefoner, skrev Meta. MLow ger ungefär dubbelt så hög kvalitet som Opus vid 6 kbps och använder samtidigt 10 % mindre beräkningskraft.

För en tjänst vars användare spänner från bärbara datorer i toppklass till tio år gamla telefoner avgör komplexitetsbudgeten. Neurala codecar används först där den installerade basen är ny och leverantören styr båda ändarna.

Vad det här betyder för företag inom kommunikation

Konferenssystem för företag, telenärvaro, fjärrassistans, contribution-länkar och privata nät är de platser där en helt neural codec är en produkt i dag, förutsatt att båda ändarna är dina. Planera för tre saker från början: en reservcodec, avkodningstester på varje NPU du stöder, och en kvalitetsutvärdering på ditt eget innehåll i stället för på en leverantörs dataset.

KAPITEL 05

GENERATIV AI

05 Generativ AI: den största utrullningen som ingen kallar en codec

Varje röst- och videomodell körs på en inlärd codec

Den största utrullningen av helt neural komprimering finns inte i någon codecstandard. Den finns inuti generativ AI. En talmodell lyssnar inte på vågformer och producerar dem inte direkt: den läser och skriver de diskreta tokens som en inlärd ljudcodec ger. Kyutais Mimi, som släpptes 2024, gör om tal till tokens med 12,5 ramar per sekund och 1,1 kbps, lägre än någon telefonicodec, och kombinerar semantisk och akustisk information i en enda ström. Metas EnCodec, publicerad 2022, är en annan; Metas egna ingenjörer beskrev senare sådana codecar som utmärkta vid låga bithastigheter men för tunga för äldre telefoner (kapitel 4).

Videogenerering fungerar på samma sätt. En diffusions- eller transformermodell genererar i det latenta rummet hos en inlärd videoautokodare (VAE), och VAE:ns avkodare gör om resultatet till pixlar. LTX-Videos VAE komprimerar video 1:192: varje latent token står för ett block om 32×32 pixlar över 8 bildrutor, med 128 kanaler. Dess avkodare utför också det sista brusreduceringssteget, så den genererar delvis detaljer i stället för att bara rekonstruera dem.

Varför detta inte är distributionsformat

De här codecarna körs vid varje anrop till en röst- eller videomodell, men ingen av dem är ett format som någon annan avkodar. Varje modell har sitt eget latenta rum; det finns ingen bithastighetsstyrning, inget konformitetstest och ingen garanti för att en annan avkodare skulle ge samma bild. De är utformade för genereringskvalitet, inte för återgivningstrohet vid en given bithastighet, och de avkodar på GPU:er i datacenter. Deras framgång visar att inlärd komprimering fungerar i stor skala när ett företag äger kodaren, avkodaren och hårdvaran.

Vad det här betyder för företag

Två praktiska punkter. För det första är tokeniserarens effektivitet en direkt kostnadspost för alla som erbjuder generativ video eller röst: för en transformer som beräknar uppmärksamhet över alla tokens växer beräkningsbeho-

vet snabbare än antalet tokens, så en mer komprimerande latent representation sänker genereringskostnaden mer än proportionellt analys. Codecingenjörer som förstår latent design, entropimodeller och avkodareffektivitet är värdefulla för team inom generativ AI. För det andra flyttar sig gränsen: standardiseringsorgan har börjat arbeta med kodning för maskiner i stället för människor. MPEG:s Video Coding for Machines (ISO/IEC 23888-2) nådde stadiet utkast till internationell standard i början av 2026 och planeras bli slutgiltig i slutet av 2026, och JPEG AI utformades från början för att tjäna datorseende lika väl som mänskligt tittande.

KAPITEL 06

SPEL OCH GRAFIK

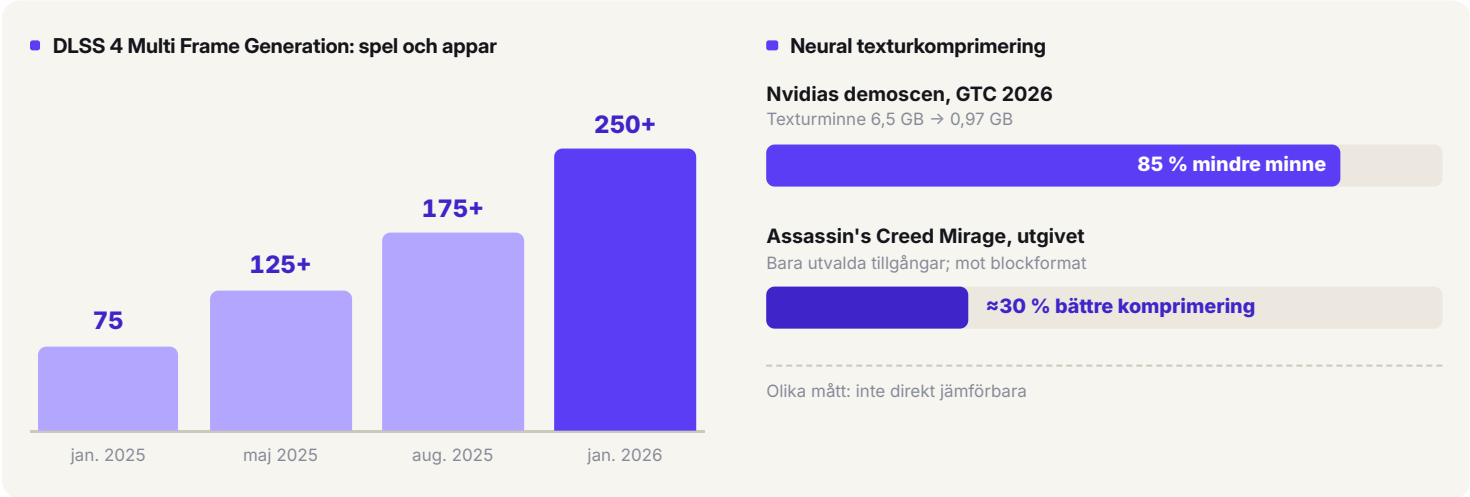
06 Spel och grafik: neural rekonstruktion är etablerad, neurala tillgångar är på väg in

Rendera färre pixlar och låta ett nätverk fylla i

I realtidsgrafik rekonstruerar ett nätverk redan det mesta av det spelarna ser. Nvidias DLSS renderar färre pixlar och färre bildrutor och använder nätverk för att ta fram resten. DLSS 4 med Multi Frame Generation lanserades i januari 2025 med stöd i 75 spel och appar; Nvidia räknade till mer än 125 i maj 2025, mer än 175 i augusti och mer än 250 i januari 2026, då företaget presenterade DLSS 4.5 med en transformermodell av andra generationen för superupplösning. Det är en avkodare i allt utom namnet: ”bitströmmen” är en bildruta med låg upplösning plus rörelsedata, och nätverket bygger upp hela bilden igen.

FIG. 05 **När ett företag styr hela stacken sprids neural rekonstruktion snabbt**

Vänster: spel och appar med stöd för Nvidias DLSS 4 Multi Frame Generation. Höger: neural texturkomprimering i en demoscen och i ett utgivet spel, som bara använde den för utvalda tillgångar på grund av kostnaden vid körning.



Källa: Nvidia (januari 2025 till januari 2026; demo på GTC 2026); Ubisoft La Forge (februari 2026). Leverantörsuppgifter; de två texturvärdena mäter olika saker.

Neurala texturer har nått ut, selektivt

Texturer komprimeras i fasta blockformat (BCn) som GPU:er avkodar i hårdvara. Neural texturkomprimering ersätter en uppsättning materialtexturer med en kompakt latent representation och ett litet nätverk som rekonstruerar texlar när de behövs. I februari 2026 beskrev Ubisofts labb La Forge hur tekniken kom med i Assassin's Creed Mirage, ett spel från 2023, med omkring 30 % högre effektiv komprimering än de konventionella formaten vid

samma visuella kvalitet leverantörsuppgift. Den användes inte överallt: ”användningen begränsades av prestandan vid körning”, så Ubisoft tillämpade den på en delmängd av tillgångarna med många instanser och högt tryck på minnet.

Nvidias RTX Neural Texture Compression SDK, i betaversion på GitHub, visar taket. I en demonstration på GTC 2026 behövde en villascen som använde 6,5 GB texturminne med BCn 970 MB med neurala texturer, 85 % mindre leverantörsuppgift. SDK:t erbjuder två vägar: avkodning till BCn vid inläsning, vilket sparar diskutrymme och nedladdningsstorlek men inte videominne, eller avkodning vid sampling, vilket sparar minnet men kostar shadertid i varje bildruta.

API:t är på väg

Grafik-API:er lägger till de matrisoperationer som neural avkodning i shaders behöver. Microsoft släppte Shader Model 6.9 i februari 2026; funktionen cooperative vector, som förhandsvisades 2025, gick upp i en samlad design för linjär algebra i en senare shadermodell, och DirectX Linear Algebra blev en offentlig förhandsversion i april 2026. När det finns i varje drivrutin kan en neural texturavkodare köras på alla tillverkares GPU:er från en enda kodväg.

Vad det här betyder för företag

Spel levererar neural avkodning eftersom villkoren för ett slutet system är uppfyllda: en motor, en tillgångskedja, tensorhårdvara i GPU:n och ett reservformat. Det öppna arbetet finns i verktyg, tillgångskedjor som kodar neurala material, körmiljöer som fungerar mellan tillverkare och framför allt mobilen, där energikostnaden per avkodad texel avgör.

KAPITEL 07



07 SPECIALISERADE LÄNKAR
Specialiserade länkar: satelliter, drönare och robotar

Det här är fallen där helt neural kodning passar bäst på papperet. En ägare styr båda ändarna av länken, bandbredden är knapp och förlustbehäftad, innehållet är tillräckligt avgränsat för att träna på, och en perceptuellt eller till och med semantiskt trogen bild räcker ofta. Förlusttåliga neurala codecar för sådana länkar är ett aktivt forskningsområde, och återställning på serversidan av flöden med låg bithastighet används.

De är också en liten och allt mer reglerad marknad jämfört med konsumentvideo analys. I Kina reviderar tillkännagivande nr 49 från 2026, från handelsministeriet och två andra myndigheter, utfärdat den 19 oktober och i kraft från den 18 november 2026, exportkontrollen för 13 kategorier av drönarrelaterade produkter. Bland teknikposterna finns en som omfattar obemannade markplattformar som delvis definieras av tröskelvärden som 75 % semantisk bildkomprimering och bildrekonstruktion inom 150 millisekunder, enligt fackpressen. Semantisk komprimering har i sig blivit en parameter i exportkontrollen. I USA förde Federal Communications Commission upp utlandstillverkade drönare och deras kritiska komponenter på sin Covered List den 22 december 2025, så att nya utländska modeller inte längre kan få utrustningsgodkännande.

Slutsatsen: en sund nisch för ett specialiserat företag med industri- eller myndighetskunder och kapacitet att hantera exportefterlevnad på båda sidor, men inte en massmarknad, och inte platsen för att ta reda på om neural kodning fungerar i stor skala.

KAPITEL 08

KISEL

08 Kiselperspektivet: energi, beräkning, minne och exakthet

Det här är kapitlet där min egen bakgrund spelar störst roll. Jag har många år av FoU och ledning inom fabless-chipdesign bakom mig, bland annat full-custom-konstruktion av ASIC-datavägar och multimediaprojekt för kamerachip, till exempel acceleratorer för videocodecar, och de frågor ett chipteam ställer om en ny codec handlar inte om BD-rate. De är: hur många joule per pixel, hur många byte DRAM-trafik per pixel, hur många grindar, och kan varje avkodare i världen räkna fram exakt samma bitar? På alla fyra punkterna står en neural avkodare 2026 där hårdvaru-HEVC stod före sina första chip: lovande i mjukvara, dyr i kisel.

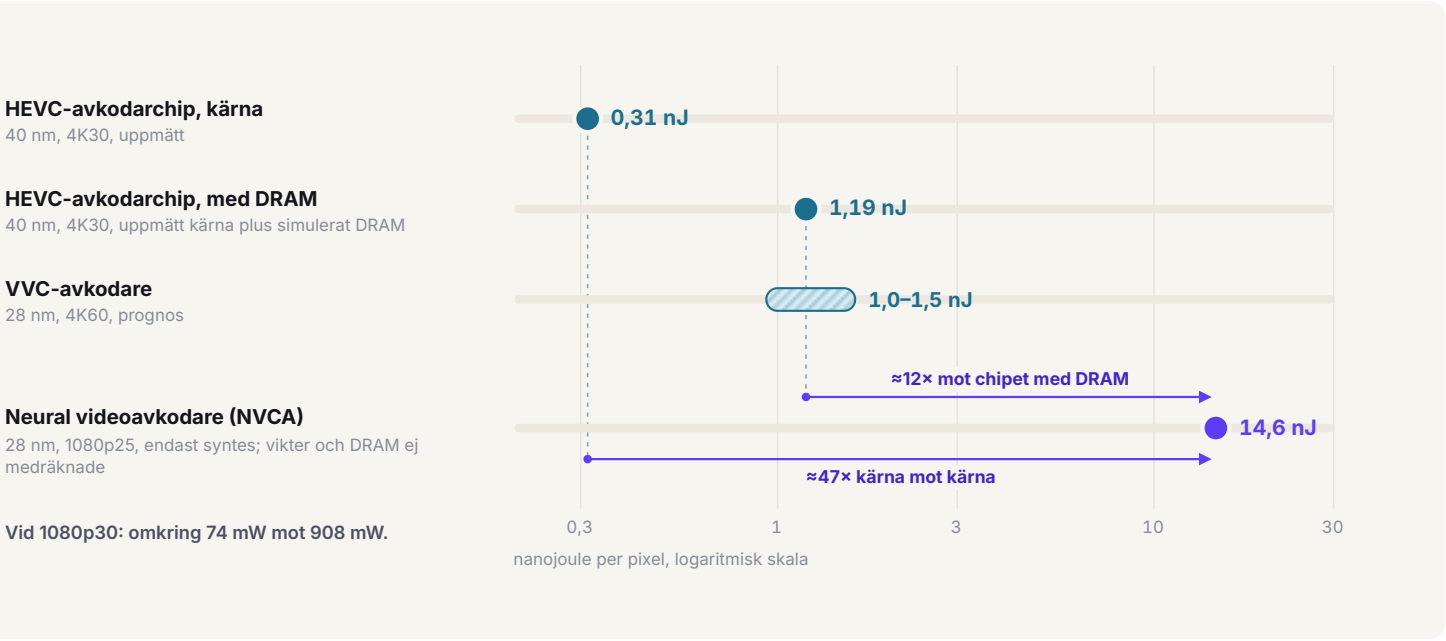
Energi per pixel: en storleksordning isär

Få neurala avkodare har konstruerats som chip, så jämförelsen vilar på en handfull publicerade konstruktioner. Referenspunkten på den klassiska sidan är MIT:s HEVC-avkodarchip (Tikekar et al., ISSCC 2013 och IEEE JSSC 2014): 40 nm, 1,77 mm², avkodning av 4K med 30 bilder per sekund vid 200 MHz. Kärnan förbrukar 78 mW, eller 0,31 nJ per pixel; räknat med simulerad DRAM-effekt behöver systemet 1,19 nJ per pixel. För VVC är produktionssiffrorna proprietära; en analys i Electronic Design 2026, som skalar MIT-chipet till 28 nm och tillämpar VVC:s högre komplexitet, förutser 1–1,5 nJ per pixel för en 4K60-avkodare på omkring en miljon grindar.

På den neurala sidan är den enda publicerade ASIC-konstruktionen av en helt neural videoavkodare, enligt Electronic Designs räkning, NVCA (Zhang et al., DATE 2024), syntetiserad i TSMC 28 nm med 5,01 miljoner grindar, exklusive vikterna, som ligger utanför chipet. Den avkodar 1080p med 25 bilder per sekund vid 760 mW, eller 14,6 nJ per pixel.

FIG. 06 **Energi per avkodad pixel: avkodare med fast funktion ligger en storleksordning före**

Den enda publicerade chipkonstruktionen av en neural videoavkodare behöver omkring 12 gånger så mycket energi per pixel som ett uppmätt HEVC-avkodarchip räknat med sitt DRAM, och omkring 47 gånger så mycket som enbart dess kärna. Den neurala konstruktionens egen trafik för vikter och DRAM är inte medräknad.



Källa: Tikekar et al., IEEE JSSC 2014 (MIT); NVCA, Zhang et al., DATE 2024 (arXiv 2312.10716); VVC-intervallet är en prognos från Electronic Design (2026). HEVC-chipet är tillverkat i en äldre nod (40 nm mot 28 nm), vilket gynnar den neurala konstruktionen.

Det är omkring 12 gånger HEVC-chipet räknat med sitt DRAM, och omkring 47 gånger enbart dess kärna. Jämförelsen gynnar NVCA på tre sätt: HEVC-chipet är tillverkat i en äldre nod, NVCA:s siffra utesluter dess trafik för vikter och DRAM, och NVCA är ett syntesresultat snarare än uppmätt kisel. Vid 1080p30 är skillnaden omkring 74 mW mot ungefär 900 mW: en avkodare som ingen märker, mot en som syns i en telefons batteri- och värmebudget.

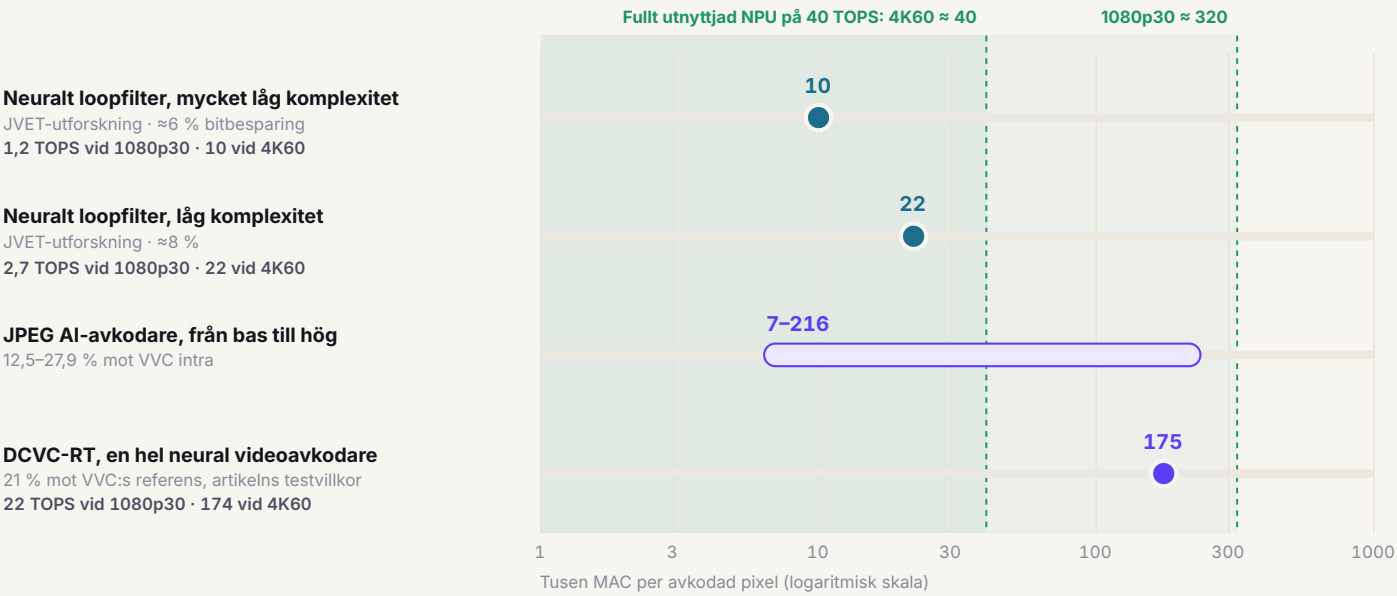
Beräkning per pixel: vad en neural avkodare kräver av en NPU

Den snabbaste publicerade neurala videocodec jag hittade är Microsoft Researchs DCVC-RT (Jia et al., CVPR 2025). Den kodar och avkodar 1080p med 125 respektive 113 bilder per sekund på en Nvidia A100-GPU och sparar samtidigt 21 % bithastighet mot VVC:s referenskodare i artikelns testvillkor, och den använder heltalsaritmetik så att olika enheter avkodar konsekvent. Dess tilläggsmaterial anger komplexiteten för varje modul. När jag lägger ihop de fyra moduler som en avkodare måste köra (avkodare, särdragsextraktor, entropimodell och rekonstruktion) får jag omkring 175 tusen MAC-operationer per avkodad pixel uppskattning.

Vid 1080p30 är det omkring 11 biljoner MAC-operationer per sekund, eller omkring 22 TOPS vid två operationer per MAC, före eventuella förluster i utnyttjandegrad; 4K60 kräver åtta gånger så mycket uppskattning. Som jämförelse definierar Microsoft sina AI-datorer genom en NPU på minst 40 TOPS, och Apple anger M4:s Neural Engine till 38 TOPS. Det är toppvärden, och kapaciteten delas med allt annat enheten gör.

FIG. 07 **Ett neuralt verktyg kostar tiotusentals MAC per pixel; en hel neural avkodare kostar hundratusentals**

Avkodarens MAC-operationer per pixel, logaritmisk skala, jämfört med vad en fullt utnyttjad NPU på 40 TOPS skulle räcka till vid 4K60 och vid 1080p30. Vinsterna är mätta mot olika jämförelsebaser och är inte direkt jämförbara.



Källa: JVET via ordföranden för SC 29 (april 2026); JPEG-kommittén (2024); DCVC-RT, tilläggsmaterial till CVPR 2025, författarens summa av avkodarmodulerna; Microsofts definition av Copilot+ PC. Författarens beräkning: två operationer per MAC, fullt utnyttjande.

Stegen i figur 7 är det användbara sättet att läsa området. Neurala verktyg som kostar tiotusentals MAC per pixel, som JVET:s loopfilter (omkring 6 % besparing vid 10 tusen, 8 % vid 22 tusen), är rimliga inuti en framtida standardavkodare. JPEG AI spänner över hela intervallet, från 7 till 216 tusen MAC per pixel för 12,5–27,9 % jämfört med VVC intra. En hel neural videoavkodare på ett par hundra tusen MAC per pixel hör hemma på en NPU, med måttlig upplösning, i ett slutet system. Det är precis där MLVC befinner sig: 540p30, med högst halva NPU:n.

Minnestrafik, inte MAC, är ofta den verkliga gränsen

Författarna bakom DCVC-RT fann att det som begränsade hastigheten inte var aritmetik utan ”operationskostnad”: minnesin- och utmatning och antalet funktionsanrop. De tog bort explicita rörelsemoduler och bearbetning i flera skalor till stor del av det skälet. NVCA:s konstruktörer byggde ett särskilt dataflöde för att minska minnestrafiken utanför chipet från rörelse- och residualsärdrag. En chipkonstruktör känner igen mönstret. I MIT:s HEVC-chip står DRAM för omkring tre fjärdedelar av systemets energi (0,88 av 1,19 nJ per pixel), även efter en cache som minskade DRAM-bandbredden med 67 % analys. En neural avkodare som behåller stora särdragskartor med många kanaler mellan bildrutor gör minnesproblemet värre, inte bättre.

Exakthet: varje avkodare måste räkna fram samma bitar

Entropikodning fungerar bara om avkodarens sannolikhetsmodell stämmer exakt med kodarens. I en klassisk codec är den modellen en tabell eller en liten uppsättning heltalsregler. I en neural codec är den ett nätverk, och en flyttalsskillnad mellan två enheter, från en annan beräkningskärna, en annan ackumuleringsordning eller en annan NPU, kan få avkodaren att läsa fel symboler och tappa strömmen från den punkten och framåt.

Alla som har levererat eller standardiserat en neural codec har fått konstruera bort det här. DCVC-RT konverterar sin modell till heltal för konsekvent kodning mellan enheter. MLVC använder en hyperprior-konstruktion med delad skala för att klara avvikelser mellan plattformar, vilket sannolikt är en del av förklaringen till att den ger upp

komprimering jämfört med DCVC-RT analys. JVET:s utforskning av neurala nätverk kör varje nätverk genom SADL, ett litet bibliotek med inferens i 16-bitars heltal. JPEG AI definierar konformitetspunkter i den latenta domänen, residualdomänen och pixeldomänen. För ett chipteam är det samma disciplin som HEVC:s heltalstransformer: standarden måste definiera aritmetiken exakt, vilket begränsar nätverkskonstruktionen och utesluter leverantörsspecifika genvägar.

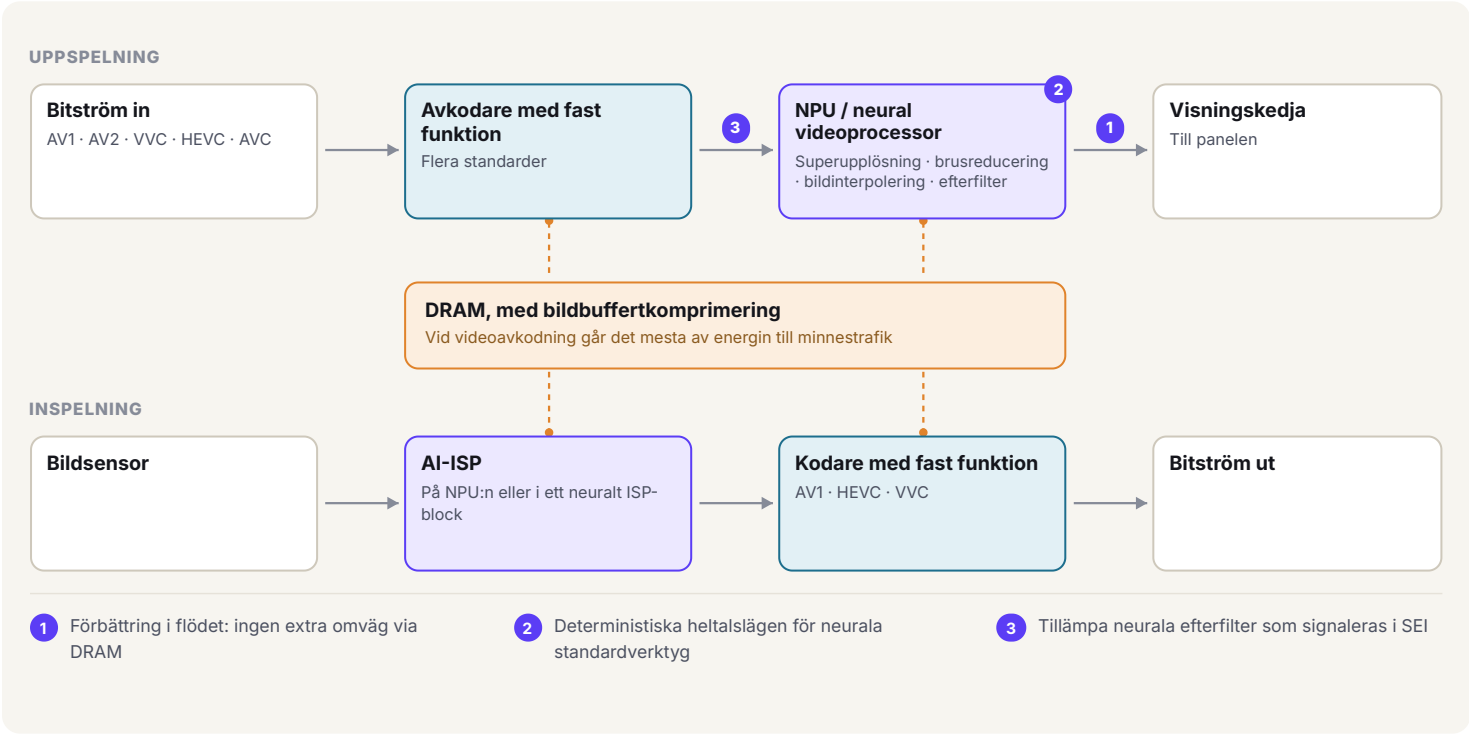
Kiselytan är inte hindret; verifiering och livslängd är det

NVCA:s 5 miljoner grindar mot omkring en miljon för en prognostiserad VVC-avkodare ser mycket ut, men på en modern SoC för telefon eller tv som redan har en NPU på flera TOPS är några miljoner grindar inte det som avgör konstruktionen. Det som kostar verkliga pengar är verifiering mot en konformitetssvit, att hålla en programmerbar avkodare korrekt genom mjukvaruuppdateringar under en enhets hela livslängd, och effekten vid 4K60 HDR. De kostnaderna är motiverade först när en standard har fastställt vad som ska byggas.

Arkitekturen som vinner nu: en codec med fast funktion plus en NPU

FIG. 08 **Uppdelningen i dagens produkter: hårdvara med fast funktion för bitströmmen, en neural motor för pixlarna**

En medie-SoC från 2026. Codecblocket avkodar alla standarder för en bråkdel av en nanojoule per pixel; NPU:n gör det som nätverk är bra på och som standarder inte begränsar. DRAM-trafiken är den kostnad som avgör.



Källa: Författarens bild av arkitekturen, byggd på produktuppgifter från Tencent (Canghai V2), Chips&Media, Allegro DVT och Samsung (2026).

Kislet som säljer 2026 kombinerar en codec med fast funktion för flera standarder med en neural motor bredvid. Codecblocket hanterar AV1, AV2, VVC, HEVC och AVC för en bråkdel av en nanojoule per pixel. Den neurala motorn hanterar det som nätverk är bra på och som standarder inte begränsar: brusreducering och ISP vid inspelning, superupplösning, bildinterpolering och förbättring vid visning. De två delar bildbuffertkomprimering och DRAM-schemat, eftersom minnestrafiken är den kostnad som avgör.

Marknaden visar uppdelningen. Tencents videocodec-ASIC Canghai (沧海), egenutvecklad sedan 2019, har mer än 100 000 chip av första generationen i drift i Tencent Cloud och företagets egna tjänster för livesändning, video on demand och molnspel; andra generationen, som i maj 2026 presenterades som strömsatt och på väg in i vo-

lymproduktion för drift under andra halvåret, stöder H.265 och H.266 i en och samma arkitektur med mer än 10 % respektive 30 % bättre komprimering än föregångaren, ungefär dubbel genomströmning och bildförbättring på chipet, till exempel superupplösning. Chips&Media säljer codec-IP och en videoinriktad NPU, WAVE-N, sida vid sida. Allegro DVT säljer AV2-avkodare och har lagt till sin första IP för neural videobearbetning. Samsungs tv-processorer kombinerar standardavkodare med hundratals nätverk för uppskalning.

Vad jag skulle bygga, efter roll

SoC- och tv-chipteam. Behåll avkodare med fast funktion för varje standard du måste spela upp, inklusive AV2 nu och VVC där sändningar kräver det. Dimensionera NPU:n för efterbearbetning i skärmens upplösning och bildfrekvens, och utforma vägen från avkodare via NPU till bildskärm så att förbättringen inte lägger till en extra omväg via DRAM. Stöd SEI för neurala efterfilter. Erbjud deterministiska heltalslägen i NPU:n. Behandla H.267 som en konstruktionsstart 2029 eller senare, och behåll viss programmerbarhet för dess neurala verktyg.

IP-leverantörer. Sälj AV2- och VVC-avkodare nu, plus de delar som neural bearbetning behöver runt dem: accelleratorer för efterfilter, bildbuffertkomprimering och konformitetsverktyg.

Videoplattformar. Neural förbehandling vid kodaren lönar sig först, eftersom den körs en gång per titel och inte kräver något av enheterna. Använd superupplösning på klientsidan där enheterna tillåter det. Behåll helt neurala codecar för slutna system som du styr.

KAPITEL 09

STANDARDER OCH PATENT

09 Standarder och patent: där neurala verktyg blir licensierbara

Standardiseringskedjan

Standard	Organ	Neuralt innehåll	Status, november 2026
H.267 (arbetsnamn; bortom VVC)	JVET (ITU-T och ISO/IEC)	Öppet. Klassiska verktyg (ECM) och neurala verktyg (loopfilter, intraprediktion, superupplösning) är båda kandidater	Begäran om förslag juli 2026; svaren utvärderas januari 2027; första testmodell andra halvåret 2027; version 1 slutet av 2029
VVC och HEVC, med VSEI	JVET	Ett neuralt efterfilter kan beskrivas i SEI-metadata; att tillämpa det är valfritt för avkodaren	Gäller; meddelandena utökade i utgåvorna från oktober 2025
AV2	AOMedia	Klassiska verktyg, royaltyfritt	Version 1.0.0 den 28 maj 2026; avkodar-IP licensierad från juni 2026; YouTube-strömmar på webben planerade till början av 2027
JPEG AI	JPEG (ISO/IEC och ITU-T)	Helt inlärd bildcodec, för mänskligt seende och maskinseende	Del 1 publicerad 2025; en andra version diskuteras; inget förvalt stöd i produkter hittat
IEEE 1857.11	IEEE	Bildkodning med neurala nätverk, utformad för NPU:er	Publicerad december 2024

Standard	Organ	Neuralt innehåll	Status, november 2026
Video Coding for Machines	MPEG	Kodning för maskinanalys i stället för tittande	Utkast till internationell standard i början av 2026; slutgiltig planerad till slutet av 2026
Opus ML-tillägg	IETF	Valfri djup redundans, förbättring på avkodarsidan, bandbreddsutvidgning	I libopus 1.5 och 1.6; djup redundans under standardisering

Två saker sticker ut. För det första kommer neuralt innehåll in i standarderna i kanterna före kärnan: som valfria efterfilter, valfri förbättring och bildkodning, som alla lämnar den huvudsakliga avkodningsvägen orörd. För det andra är den enda standard som skulle kunna placera neurala verktyg i avkodningsvägen på varje skärm, H.267, tre år från en första version.

Vem som har positionerat sig i patenten

Standardverktyg blir licensierbara patent, så H.267-processen är också en patentkapplöpning, och de neurala verktygen är en del av den. GreyB:s kartläggning från 2025 av 3 916 patent bortom VVC placerar ByteDance i täten, med 645 patentfamiljer och 251 bidrag till JVET, och noterar samarbeten mellan universitet och industri, till exempel Xiaomi och Xidian University med mer än 20 familjer. MLex rapporterade i mars 2026 att Tencent, Alibaba och ByteDance växer fram som aktiva deltagare i patentpooler för videocodecar. I maj 2026 köpte Sauvegarder Investment Management 126 patenttillgångar för videocodecar från Alibaba som omfattar AV1, AV2 och Video Coding for Machines. InterDigitals köp av Deep Render i oktober 2025 tillförde Deep Renders patent för AI-baserad videokodning till företagets portfölj.

Pengarna per enhet syns tydligt i poolernas prislistor. Access Advances avgifter för VVC (per den 1 juli 2026, för licenstagare som följer villkoren och använder poolens varumärke) är 0,50 dollar per mobiltelefon, surfplatta eller bärbar dator på marknader i Region 1 som USA, EU och Japan, och 0,25 dollar i Region 2; för 4K-tv-apparater 1,00 till 1,50 dollar beroende på pris i Region 1; med ett årligt tak per företag på 75 miljoner dollar. AV1 och AV2 är royaltyfria från AOMedias medlemmar. Vad H.267 än innehåller kommer dess verktyg att prissättas på samma sätt, vilket är skälet till att företag med patent inom neural kodning bidrar till JVET nu analys.

KAPITEL 10

10 VEM SOM FÅR BETALT
Vem som får betalt, och var ett företag kan satsa

Pengarna 2026

FIG. 09 **Vem som får betalt för neural kodning 2026**

Pengarna följer standarderna och plattformarna. Jag hittade inget fristående företag som försörjer sig på att licensiera ut en helt neural videocodec.

TJÄNAR NU

Plattformar

Behåller besparingarna

Netflix neurala nedskalare; superupplösning hos YouTube och Meta; Microsoft släppte MLVC som öppen källkod under MIT

TJÄNAR NU

Enhets- och chiptillverkare

Säljer differentiering

Kamera-ISP:er; uppskalningsprocessorer för tv; Tencents egenutvecklade codec-chip Canghai

TJÄNAR NU

IP-leverantörer

Licensierar varje ny standard

Kunder för AV2-avkodare hos Chips&Media, Allegro DVT och VeriSilicon inom tre veckor efter specifikationen

TJÄNAR NU

GPU-leverantörer

Bygger plattformsinlåsning

DLSS 4 i över 250 spel och appar; SDK för neurala texturer i betaversion

PER ENHET

Patentinnehavare och pooler

Tar betalt per enhet

VVC: 0,50 dollar per telefon, 1,00–1,50 dollar per 4K-tv; H.267-verktyg efter 2029

EXIT GENOM FÖRVÄRV

Startupbolag inom neurala codecar

Uppköpta

WaveOne → Apple (2023); iSIZE → Sony (2023); Deep Render → InterDigital (2025)

Källa: Uppgifter från företag och patentpooler som citeras i rapporten; läget den 12 november 2026. Poolavgifter: Access Advance, VVC, Region 1, för licenstagare som följer villkoren och använder varumärket, per den 1 juli 2026.

Plattformarna behåller besparingarna. Netflix, YouTube, Meta, Microsoft och Tencent bygger neural bearbetning internt och behåller nyttan i form av lägre bandbredd, mindre lagring och bättre kvalitet per bit. När det hjälper spridningen ger de bort koden: Microsoft släppte MLVC under MIT-licensen.

Enhets- och chiptillverkare säljer differentiering. Kamera-ISP:er och bildprocessorer i tv-apparater är där neurala nätverk når varje användare, och de säljs som funktioner: Qualcomms Cognitive ISP, Samsungs uppskalningsprocessorer.

IP-leverantörer säljer avkodare för varje ny standard, plus NPU:er. Inom tre veckor efter AV2:s specifikation hade Chips&Media, Allegro DVT och VeriSilicon alla presenterat kunder för AV2-avkodare. Samma leverantörer säljer nu block för neural bearbetning vid sidan av codecarna.

GPU-leverantörer bygger plattformar. DLSS och neural texturkomprimering gör Nvidias tensorhårdvara till en del av hur spel skapas.

Patentinnehavare tar betalt per enhet, genom pooler för VVC och HEVC och, efter 2029, H.267.

Startupbolagen inom neurala codecar har köpts upp. WaveOne gick till Apple 2023, iSIZE till Sony Interactive Entertainment 2023 och Deep Render till InterDigital 2025, alla på villkor som inte offentliggjorts. I varje fall köpte

en plattform eller en licensgivare teamet och patenten. Jag hittade inget fristående företag 2026 som försörjer sig på att licensiera ut en helt neural videocodec.

Var ett företag kan satsa

Satsning	Vem köper	Varför nu	Huvudrisk
Neural förbehandling vid kodaren (nedskalning, brusreducering, optimering per titel)	Streaming- och användarvideoplattformar, tv-bolag	Ingen ändring i enheterna krävs; beprövat hos Netflix	Stora plattformar bygger det själva
IP och finjustering för förbättring på skärmsidan (superupplösning, bildinterpolering, bildkvalitet)	Chiptillverkare för tv, digitalboxar och telefoner	NPU:erna finns redan i enheterna; synlig differentiering	Kommoditisering; chiptillverkarnas egna team
Neurala codecar i slutna system (konferenser, fjärrstyrd drift, privata länkar)	Leverantörer inom företagskommunikation, industri och telenärvaro	MLVC visar att det är genomförbart, och koden är öppen	Du måste styra båda ändarna; splittrat NPU-landskap
Kisel eller IP för codec plus NPU (AV2, VVC och efterbearbetning)	SoC-tillverkare	AV2-cykeln börjar	Starka etablerade aktörer
Verktyg för utvärdering, konformitet och deterministisk inferens	Alla som levererar neurala verktyg	Bitexakthet är ett hårt krav	Liten marknad utan paketering
Standardbidrag till H.267	Licensgivare, stora plattformar	Svaren utvärderas i januari 2027	Långt, dyrt, osäkert

För de flesta företag är det i de två första raderna som intäkter finns att hämta 2027. Den tredje är den mest intressanta för ett fokuserat produktföretag. Den sista kräver tålamod och en licensstrategi.

KAPITEL 11

11 PÅSTÅENDEN OCH BELÄGG
Påståenden du kommer att höra, och vad beläggen säger

Påstående	Vad beläggen säger
"Neurala codecar slår redan VVC."	I labbet, ja. DCVC-RT sparar 21 % mot VVC:s referenskodare vid 1080p i artikelns testvillkor; JPEG AI sparar 12,5–27,9 % mot VVC intra. Avkodaren kostar omkring 175 tusen MAC per pixel, eller energin hos en NPU, mot en bråkdel av en nanojoule i hårdvara med fast funktion.
"AI-codecar kommer att ersätta H.26x och AV1 inom några år."	Inte i uppspelning. Den första standarden som kan föra ut neurala verktyg siktar på slutet av 2029, chip följer omkring 2031 och massmarknaden i mitten av 2030-talet uppskattning. Codecen som kommer nu, AV2, är klassisk.
"MLVC minskar bandbredden med 88 %."	Mot H.264 vid 360p, på Microsofts egna testdata. Mot hårdvaru-HEVC är det 75,5 % vid 360p och 65,4 % vid 540p, i samtal där Microsoft styr båda ändarna.
"Neurala texturer minskar videominnet med 85–96 %."	I demonstrationsscener. Det utgivna spelet fick omkring 30 % högre effektiv komprimering och använde neurala texturer bara för utvalda tillgångar på grund av kostnaden vid körning.
"JPEG AI är en ISO-standard, så kameror och webbläsare kommer att införa den."	Jag hittade inget förvalt stöd i produkter. Formatet som webben införde 2026, JPEG XL i Chrome 155, är klassiskt.
"Neural avkodning är för dyr för telefoner."	Delvis sant. MLVC avkodar 540p30 inom en halv NPU; 4K60 med en modell i DCVC-RT-klassen skulle kräva omkring 175 TOPS kontinuerlig beräkningskraft uppskattning.
"Satellit- och drönarlänkar är den stora marknaden för neural video."	Tekniskt bäst lämpade, men marknaden är liten och allt mer exportkontrollerad på båda sidor av Stilla havet.
"Stora plattformar kommer att licensiera en neural codec från ett startupbolag."	De byggde sin egen (Netflix, Microsoft, Tencent) eller köpte startupbolaget (Apple, Sony, InterDigital).

Ett enkelt test skär igenom de flesta säljpresentationer: fråga efter jämförelsebasen, måttet, upplösningen och avkodningskostnaden. Ett påstående som inte kan ange alla fyra är ett forskningsresultat, inte en produkt.

KAPITEL 12

12 HÅLLPUNKTER Att bevaka, 2027 till 2030

När	Hållpunkt	Varför det spelar roll
Januari 2027	JVET utvärderar H.267-svaren i Brisbane	Hur många av de bästa svaren som bygger på neurala verktyg, och till vilken kostnad per pixel
Början av 2027	YouTubes första AV2-strömmar på webben; AV2-avkodare i annonserade SoC:er	Nästa klassiska cykel sätter enheternas budgetar fram till omkring 2030
Andra halvåret 2027	H.267:s första testmodell	Om neurala verktyg finns i kärnkonstruktionen eller bara i utforskningen
2027	MLVC bortom peer-to-peer-samtal i Teams; andra leverantörer som tar till sig den öppna koden	Det första testet av en öppen neural codec utanför ägarens egen miljö
2027–2028	En första produkt som skriver eller läser JPEG AI som förval	Det första neurala formatet som används över företagsgränser
2027–2028	DirectX Linear Algebra i levererade drivrutiner; neurala texturer i fler spel och i mobilen	Neural avkodning i grafik som fungerar mellan tillverkare
Löpande	NPU-budgetar i telefoner och datorer	Vid omkring 100 TOPS kontinuerligt skulle en 1080p30-avkodare i DCVC-RT-klassen ta omkring en femtedel av NPU:n i stället för hälften analys
Slutet av 2029	H.267 version 1	Starten på kiselcykeln för neurala verktyg i uppspelning på massmarknaden

KAPITEL 13

13 ÖPPNA FRÅGOR Öppna frågor

Flera frågor i den här rapporten har ännu inget offentligt svar, och det är dem jag tycker är mest intressanta tekniskt.

Exakthet i stor skala. Hur nära bitexakt aritmetik med enbart heltal kan en neural avkodare komma på NPU:er från olika leverantörer innan den ger upp större delen av sin komprimeringsfördel? MLVC:s avvägning på sex procentenheter är en datapunkt; en standard behöver ett generellt svar.

Minne före MAC. Vad kostar neural efterbearbetning i DRAM-trafik per pixel när den körs i flödet mellan en hårdvaruavkodare och bildskärmen, och hur liten kan dess arbetsmängd göras?

Priset för ett neuralt verktyg i H.267. Vid hur många MAC och nanojoule per pixel slutar ett neuralt loopfilter att vara värt de bitar det sparar i en 4K60-avkodare?

Slutna system bortom samtal. Vilka andra slutna system, från lagring av foton i molnet och spelstreaming till robotar och privata contribution-länkar, kan införa helt neurala codecar innan H.267 kommer?

Avkodare som genererar. När en avkodare genererar detaljer i stället för att rekonstruera dem, som video-VAE:er redan gör, hur bör återgivningstroheten mätas och specificeras?

Det här är områden vi kan utforska tillsammans under de kommande åren, genom tekniskt utbyte och tekniskt samarbete: i@tanshuai.com.

Begränsningar i den här rapporten

Det här är en skrivbordsstudie baserad på offentliga dokument fram till den 12 november 2026. Jag gjorde inga primärintervjuer och inga oberoende mätningar för den här utgåvan. Flera nyckeltal är leverantörers eller författares påståenden på egna testuppsättningar, som MLVC:s besparingar i bithastighet och kompressionsförhållandena för neural texturkomprimering; de är märkta som sådana. Mina beräkningsuppskattningar för DCVC-RT summerar de värden per modul som författarna publicerar och bortser från förluster i utnyttjandegrad, så den verkliga NPU-belastningen blir högre. Energijämförelsen vilar på några få publicerade konstruktioner i olika processnoder; produktionssiffror för VVC- och AV2-avkodare är proprietära. Standardernas tidsplaner är kommittéernas egna mål. H.267-datumen efter 2029 är mina uppskattningar utifrån historiken för AV1 och VVC.

Om författaren

Tan Shuai är teknikexpert och grundare med mer än tjugo års arbete inom chip, hårdvara och systemmjukvara, och 20 beviljade patent. Han har många år av FoU och ledning inom fabless-chipdesign bakom sig, bland annat full-custom-konstruktion av ASIC-datavägar och multimediaprojekt för kamerachip, till exempel acceleratorer för videocodecar, och det är den här rapportens lins: energi per pixel, minnestrafik och vad en standard låter en avkodare låsa fast i kisel. Hans arbete har tagit honom genom Silicon Valley, New York, Chicago, Taiwan, Greater Bay Area (Shenzhen och Hongkong), Vietnam och Malaysia. I dag arbetar han som AI-forskare och -utvecklare.

Teknisk diskussion: i@tanshuai.com · tanshuai.com

Anmärkningar och ansvarsfriskrivning

Den här rapporten är en oberoende skrivbordsstudie baserad på offentliga dokument som var tillgängliga fram till den 12 november 2026. Den tillhandahålls endast i informationssyfte och utgör inte investerings-, juridisk eller teknisk rådgivning. Produkter, standardernas tidsplaner, patentpoolers avgifter och exportregler ändras; kontrollera allt du förlitar dig på mot primärkällan. Siffror märkta leverantörsuppgift, uppskattning eller analys ska läsas som sådana. Företags- och produktnamn tillhör sina ägare och används endast för identifiering; inget företag har granskat eller sponsrat den här rapporten.

© 2026 Tan Shuai. Du får dela rapporten oförändrad med angivande av upphovsman.

Källor

Samtliga hämtade mellan den 3 och den 12 november 2026. Grupperade efter kapitel.

Trafik, enheter och avgränsning

- Ericsson, *Mobility Report*, juni 2026 (video omkring 75 % av den mobila datatrafiken i slutet av 2025): https://www.ericsson.com/en/reports-and-papers/mobility-report/reports/june-2026 ; sammanfattning: https://telecomlead.com/5g/ericsson-mobility-report-2026-global-mobile-data-traffic-to-reach-515-eb-by-2031-as-india-leads-smartphone-data-consumption-126290
- IDC, globala leveranser av smarttelefoner, Q4 och helåret 2025 (1,26 miljarder), 13 januari 2026: https://www.idc.com/resource-center/press-releases/q425mobilephonetop5/
- TrendForce, prognos för tv-leveranser 2025 (195,71 miljoner): https://iconnect007.com/article/146406/early-shipments-by-tv-brands-push-global-1h25-volume-up-2-but-raise-concerns-for-secondhalf-peak-season/146403/ein ; Omdia, tv-prognos för 2025 (208,7 miljoner): https://iconnect007.com/article/145954/omdia-forecasts-flat-tv-shipments-in-2025-but-mini-led-poised-for-explosive-growth-at-829/145951/ein

Uppspelning och streaming

- Netflix Technology Blog, ”AV1 — Now Powering 30% of Netflix Streaming”, december 2025: https://netflixtechblog.com/av1-now-powering-30-of-netflix-streaming-02f592242d80
- Netflix Technology Blog, ”For your eyes only: improving Netflix video quality with neural networks”, november 2022: https://netflixtechblog.com/for-your-eyes-only-improving-netflix-video-quality-with-neural-networks-5b8d032da09c
- AOMedia, AV2 Bitstream and Decoding Process Specification v1.0.0, 28 maj 2026: https://av2.aomedia.org/v1.0.0/index.html ; ”AV2: What the Final Specification Means for Implementers”: https://aomedia.org/blog%20posts/AV2-What-the-Final-Specification-Means-for-Implementers/
- TV Technology, ”YouTube to Commence AV2 Streaming in 2027” (AV1:s andel av YouTubes katalog): https://www.tvtechnology.com/platform/streaming/youtube-to-commence-av2-streaming-in-2027
- Google Research, ”The in-loop filtering pipeline in AV2”, 2026: https://research.google/pubs/the-in-loop-filtering-pipeline-in-av2/
- Fraunhofer IIS, ”The Future of Broadcast in Brazil” (DTV+ i kommersiell drift, juni 2026): https://mpegh.com/brazil/ ; CartaCapital om mottagare i tv-apparater: https://www.cartacapital.com.br/?p=628337 ; Olhar Digital om VVC i TV 3.0: https://olhardigital.com.br/2026/05/01/pro/em-breve-a-tv-aberta-vai-mudar-no-brasil/
- 9to5Google, ”YouTube adds AI-upscaled ‘Super Resolution’ for low-quality videos”, 29 oktober 2025: https://9to5google.com/2025/10/29/youtube-adds-ai-upscaled-super-resolution-for-low-quality-videos/
- Meta, ”How Meta Deployed Super Resolution at Scale to Transform Video Quality”, @Scale, 16 oktober 2025: https://atscaleconference.com/how-meta-deployed-super-resolution-at-scale-to-transform-video-quality/
- Samsung, 8K TV highlights (NQ8 AI Gen3, 768 neurala nätverk): https://www.samsung.com/us/tvs/8k-tv/highlights/

Fotografi och bilder

- Qualcomm, ”Meet the Snapdragon 8 Gen 2 Super Cameras” (Cognitive ISP): https://www.qualcomm.com/news/onq/2023/07/meet-the-snapdragon-8-gen-2-super-cameras
- Chips&Media och Visionary.ai, fullständig AI-baserad ISP, 5 januari 2026: https://www.edge-ai-vision.com/2026/01/chipsmedia-and-visionary-ai-unveil-the-worlds-first-ai-based-full-image-signal-processor-redefining-the-future-of-image-quality/ ; intervju i PetaPixel: https://petapixel.com/2026/01/18/a-look-at-the-worlds-first-full-ai-based-image-signal-processor
- JPEG-kommittén, pressmeddelande från det 103:e mötet (JPEG AI:s vinster, komplexitet, implementationer i smarttelefoner): https://jpeg.org/items/20240507_press.html
- G. J. Sullivan, introduktion till ISO/IEC JTC 1/SC 29, april 2026 (status för JPEG AI; siffror för ECM och NNVC): https://www.mpeg.org/introduction-to-iso-iec-jtc-1-sc-29-coding-of-audio-picture-multimedia-and-hypermedia-information-the-world-of-jpeg-and-mpeg/
- Arbetsgruppen för IEEE 1857.11 (bildkodning baserad på neurala nätverk): https://sagroups.ieee.org/fvc
- Chrome 155 och JPEG XL, 6 oktober 2026: https://xenospectrum.com/en/chrome-155-jpeg-xl-rust/

Samtal och möten

- Microsoft, ”Satin: Microsoft's latest AI-powered audio codec for real-time communications”, 17 februari 2021: https://techcommunity.microsoft.com/t5/microsoft-teams-blog/satin-microsoft-s-latest-ai-powered-audio-codec-for-real-time/ba-p/2141382
- Microsoft, ”Announcing the Open-Source Release of ML Video Codec (MLVC)”, 24 juli 2026: https://techcommunity.microsoft.com/blog/linuxandopensourceblog/announcing-the-open-source-release-of-ml-video-codec-mlvc/4539875 ; kod: https://github.com/microsoft/mlvc
- Streaming Learning Center, ”Microsoft MLVC: Interoperability, Power Realities, and the Neural Compression Patent Landscape”: https://streaminglearningcenter.com/articles/microsoft_mlvc.html ; analys från iReplay: https://ireplay.tv/blog/mlvc-microsoft-neural-video-codec-open-source-streaming/

- Versionsinformation för Opus 1.5: https://opus-codec.org/demo/opus-1.5 ; Opus 1.6: https://opus-codec.org/demo/opus-1.6
- Meta Engineering, ”MLow: Meta's low bitrate audio codec”, 13 juni 2024: https://engineering.fb.com/2024/06/13/web/mlow-metas-low-bitrate-audio-codec/
- Microsoft, Copilot+ PCs (NPU på 40+ TOPS): https://www.microsoft.com/en-us/windows/copilot-plus-pcs ; Apple, presentationen av M4 (Neural Engine på 38 TOPS): https://www.apple.com/newsroom/2024/05/apple-introduces-m4-chip/

Generativ AI

- Kyutai, modellkort för Mimi (12,5 Hz, 1,1 kbps): https://huggingface.co/kyutai/mimi
- Lightricks, ”LTX-Video: Realtime Video Latent Diffusion” (video-VAE med 1:192): https://arxiv.org/abs/2501.00103
- MPEG, pressmeddelande från det 153:e mötet (Video Coding for Machines som DIS): https://www.mpeg.org/wp-content/uploads/2026/03/press_release_153.pdf

Spel och grafik

- NVIDIA, spridningen av DLSS 4: lanseringen och Computex 2025 (75 och 125+): https://www.nvidia.com/en-us/geforce/news/125-dlss-4-multi-frame-gen-games-more-announced-computex-2025/ ; Gamescom 2025 (175+): https://www.nvidia.com/en-us/geforce/news/gamescom-2025-dlss-4-full-ray-tracing-game-announcements/ ; CES 2026 (250+, DLSS 4.5): https://www.nvidia.com/en-us/geforce/news/dlss-4-rtx-path-tracing-game-announcements-ces-2026/
- Ubisoft La Forge, ”Shipping Neural Texture Compression in Assassin's Creed Mirage”, februari 2026: https://www.ubisoft.com/en-us/studio/laforge/news/415slmB3ZGzv8d2LhO9QT/shipping-neural-texture-compression-in-assassins-creed-mirage
- TweakTown om NVIDIA:s demo av neurala texturer på GTC 2026 (6,5 GB till 970 MB): https://www.tweaktown.com/news/110886/nvidias-neural-texture-compression-cuts-vram-usage-from-6-5gb-down-to-970mb/index.html
- Microsoft, DirectX-bloggen, Shader Model 6.9 retail, 26 februari 2026: https://devblogs.microsoft.com/directx/shader-model-6-9-retail-and-more/ ; om cooperative vector: https://devblogs.microsoft.com/directx/shader-model-6-9-and-the-future-of-cooperative-vector/ ; planer från GDC 2026: https://windowsreport.com/microsoft-showcases-neural-rendering-and-dxr-2-0-plans-for-directx-at-gdc-2026/

Specialiserade länkar

- Global Trade Alert om Kinas tillkännagivande nr 49 från 2026: https://globaltradealert.org/intervention/161396-china-government-revises-export-controls-on-drone-related-items ; DroneXL: https://dronexl.co/2026/10/19/china-drone-export-controls-motors-props-november-18/ ; Army Recognition (tröskelvärden för obemannade markplattformar): https://www.armyrecognition.com/news/army-news/2026/china-tightens-drone-export-controls-raising-supply-chain-risks-for-the-us-europe-and-ukraine
- DroneLife, FCC för upp utlandstillverkade drönare och komponenter på Covered List, 22 december 2025: https://dronelife.com/2025/12/22/fcc-adds-foreign-made-drones-and-components-to-covered-list-citing-national-security-risks/

Kisel

- M. Tikekar et al., ”A 249-Mpixel/s HEVC Video-Decoder Chip for 4K Ultra-HD Applications”, IEEE JSSC 2014: https://dspace.mit.edu/handle/1721.1/93876 ; MIT:s sida om chipet (78 mW vid 4K30): https://eecsg.mit.edu/chips/hevc-video-decoder-for-4k-ultra-hd-applications/
- S. Zhang et al., NVCA, DATE 2024: https://arxiv.org/abs/2312.10716
- Electronic Design, ”Why Traditional VVC Will Outlive the Neural-Codec Hype on Silicon”, 2026: https://www.electronicdesign.com/home/article/55414955/why-traditional-vvc-will-outlive-the-neural-codec-hype-on-silicon
- Z. Jia et al., ”Towards Practical Real-Time Neural Video Compression” (DCVC-RT), CVPR 2025: https://openaccess.thecvf.com/content/CVPR2025/html/Jia_Towards_Practical_Real-Time_Neural_Video_Compression_CVPR_2025_paper.html ; tilläggsmaterial (komplexitet per modul): https://openaccess.thecvf.com/content/CVPR2025/supplemental/Jia_Towards_Practical_Real-Time_CVPR_2025_supplemental.pdf
- NNVC och biblioteket SADL för heltalsinferens: https://arxiv.org/abs/2309.05846
- Tencent Canghai V2, 27 maj 2026 (Shanghai Securities News via 10jqka): https://stock.10jqka.com.cn/20260527/c677016020.shtml ; sammanfattning av funktionerna: https://www.besthub.dev/articles/tencent-s-canghai-v2-chip-enters-mass-production-what-the-new-capabilities-mean-for-video-encoding-bcfc18dbfa29
- Chips&Media, avkodar-IP för AV2, 12 maj 2026: https://www.design-reuse.com/news/202530501-chips-media-completes-development-of-next-gen-av2-hw-decoder-ip/ ; AV2-licens, 17 juni 2026: https://www.edge-ai-vision.com/?p=58270
- Allegro DVT, avkodar-IP för AV2, 16 juni 2026: https://www.edge-ai-vision.com/2026/06/allegro-dvts-pulsar-decoder-ip-adds-support-for-av2-video-codec/
- VeriSilicon, VC9800D med AV2, 10 juni 2026: https://www.businesswirechina.com/en/news/63176.html

Standarder och patent

- MPEG, pressmeddelande från det 155:e mötet (gemensam begäran om förslag bortom VVC), 5 augusti 2026: https://www.mpeg.org/wp-content/uploads/2026/08/ag3w254SN26474Press-Release-MPEG155.pdf
- MPEG, pressmeddelande från det 152:a mötet (resultat av begäran om belägg; utgåvor av VVC och VSEI), 30 oktober 2025: https://www.mpeg.org/wp-content/uploads/2025/11/Press-Release-MPEG152.pdf ; MPEG-kolumnen i ACM SIGMM: https://records.sigmm.org/2025/12/09/mpeg-column-152nd-mpeg-meeting/

- Streaming Media, ”AI Helps InterDigital Reach Beyond VVC in Race to Develop Next-Gen Codec”: https://www.streamingmedia.com/Articles/News/Online-Video-News/AI-Helps-InterDigital-Reach-Beyond-VVC-in-Race-to-Develop-Next-Gen-Codec-171971.aspx
- GreyB, patentkartläggning bortom VVC (H.267): https://greyb.com/blog/exclusive-video-codec-landscape/ ; om licensiering av AV2 och försäljningen av Alibabas portfölj: https://greyb.com/blog/av2-licensing/
- MLex, ”Fragmented codec licensing exposes need for clearer rules for Chinese tech majors”, 27 mars 2026: https://www.mlex.com/mlex/articles/2458402/fragmented-codec-licensing-exposes-need-for-clearer-rules-for-chinese-tech-majors
- Access Advance, royaltyavgifter för MCBA och VVC Advance per den 1 juli 2026: https://accessadvance.com/wp-content/uploads/2025/07/MCBA-Program-Overview-September-2026.pdf

Pengar och företag

- InterDigital köper Deep Render, 30 oktober 2025: https://www.advanced-television.com/2025/10/30/interdigital-acquires-deep-render
- TechCrunch, ”Apple acquired a startup using AI to compress videos”, 27 mars 2023: https://techcrunch.com/2023/03/27/apple-acquired-a-startup-using-ai-to-compress-videos
- Neowin, ”Sony acquires AI video startup iSIZE”, november 2023: https://www.neowin.net/news/sony-acquires-ai-video-startup-isize-to-enhance-playstation-streaming/

INLÄRD KOMPRIMERING · UPPSPELNING × KISEL

Neural kodning 2026

Den används runt codecen, inte i stället för den

Varje nyckeltal i rapporten anges med källa och datum. Där författaren själv härleder en siffra är den märkt som uppskattning; där bara säljaren redovisar den är den märkt som leverantörsuppgift.

Tan Shuai

Teknikexpert och grundare · AI-forskare och -utvecklare

tanshuai.com · i@tanshuai.com

ÖPPEN UTGÅVA · Uppgifter per den 12 november 2026