Engineering Management in Production and Services

Volume 18 • Issue 1 • 2026

sciendo

received: 10 August 2025
accepted: 10 January 2026

pages: 60-72

Evolving AI models: adoption patterns of transformers and diffusers

Paweł Cabała

ABSTRACT
This study investigates the development, adoption, and implications of artificial intelligence (AI) models by analysing a comprehensive dataset of over 316,000 models hosted on the Hugging Face platform. Focusing on two dominant model architectures - transformers and diffusion models - it examines their distribution across tasks, user engagement patterns, and practical applications in domains such as natural language processing, computer vision, audio processing, and generative media. The research highlights the growing prominence of generative AI, the role of open-source platforms in shaping model accessibility, and the divergence in use trends between foundational and emerging AI tools. Drawing on correlations between downloads, likes, citations, and model size, the paper discusses how each library's community-driven dynamics shape their respective strengths. Finally, the paper discusses implications for business strategy and adoption, encompassing practical considerations like infrastructure requirements and ethical challenges, and underscores the potential for these evolving model ecosystems to drive innovative, human-centric AI solutions across diverse sectors.

KEY WORDS
artificial intelligence (AI), transformer models, diffusion models, Hugging Face, open source software, model adoption, technological diffusion, natural language processing (NLP), computer vision (CV), innovation strategy

10.2478/emj-2026-0005

Paweł Cabała
Krakow University of Economics,
Rakowicka 27, 31-105 Kraków, Poland
ORCID 0000-0001-6624-6650
e-mail: cabalap@uek.krakow.pl

Introduction

Generative artificial intelligence (AI) represents a paradigm shift, fundamentally altering how organisations and individuals approach innovation, creativity, and productivity. Transcending their origins in purely analytical tasks, contemporary AI models function as catalysts for novelty, generating original ideas, challenging established paradigms, and informing complex decision-making processes. Across diverse sectors - ranging from industry and education to healthcare and marketing - generative models, particularly large language models (LLMs)

Cabała, P. (2026). Evolving AI models: adoption patterns of transformers and diffusers. *Engineering Management in Production and Services, 18*(1), 60-72. doi: 10.2478/emj-2026-0005

and chatbots, are reshaping business operations by automating routine tasks and enabling highly personalised user experiences. Concurrently, the proliferation of these technologies necessitates critical examination of human-AI collaboration, fostering new perspectives on design thinking and innovation strategy (Sedkaoui & Benaichouba, 2024). Transformer and diffusion models are among the most impactful innovations. Transformers have revolutionised language understanding and text generation capabilities (Gupta, 2023), whereas diffusion models have introduced new frontiers in image and content synthesis, underscoring AI's expanding repertoire across both analytical and creative domains (Ahirwar, 2023).

Against this technological backdrop, the present study investigates the development patterns and distribution dynamics of AI models, with a specific focus on transformer and diffusion architectures hosted on the Hugging Face (HF) platform. As a central and rapidly expanding hub for sharing AI models, datasets, and interactive spaces (Osborne et al., 2024), Hugging Face plays a crucial role in democratising access to cutting-edge AI technologies for diverse fields, including behavioural science (Hussain et al., 2024). Recent empirical studies have begun to analyse the complex dynamics of this ecosystem, examining challenges and benefits of model reuse (Taraghi et al., 2024), validating claims about platform use patterns (Jones et al., 2024), and mapping the structure of development activity and collaboration (Osborne et al., 2024). Situated within this growing body of research, the author's primary objective is to delineate use trends across key AI domains specifically for transformer and diffusion models, analyse associated user engagement metrics, and evaluate the broader implications for innovation, collaborative practices, and the ethical deployment of generative AI technologies hosted on this influential platform.

To achieve this objective, three central research questions guide this investigation. First, the study identifies the predominant model architectures - transformer and diffusion - and examines their distribution across disciplines such as NLP, CV, and audio processing. Second, it analyses how user engagement indicators (e.g., downloads, likes, citations) diverge between these architectural classes and assesses the implications of these differences for practical adoption. Third, it explores overarching patterns in generative AI deployment, with an emphasis on characterising innovation trajectories, discerning user preferences and evolving paradigms of creativity, human-AI interaction, and platform-mediated dissemination.

This research assumes transformer and diffusion models are key AI drivers whose Hugging Face visibility reflects broader trends. Engagement metrics (e.g., downloads, likes) are treated as imperfect proxies for utility and reception. Acknowledging platform limitations, we use HF data to analyse usage patterns and diffusion, assuming these dynamics partially reveal wider ethical, social, and strategic aspects of generative AI.

1. Literature review

1.1. Transformer models

Transformer models represent a significant inflection point in artificial intelligence, having fundamentally altered the processing of sequential data (Patwardhan et al., 2023). Introduced by Vaswani et al. (2017), the transformer architecture departed from traditional recurrent and convolutional neural networks by relying exclusively on attention mechanisms. The cornerstone of the transformer model is the self-attention mechanism, which dynamically assigns importance weights to different elements within an input sequence relative to each other. This capability allows the model to capture long-range dependencies more effectively than earlier recurrent architectures, where information transmission often degraded over extended sequence lengths (Islam et al., 2024). Standard transformer architectures typically comprise encoder and decoder stacks, each incorporating multi-head attention layers, position-wise feed-forward networks, residual connections, and layer normalisation (Vaswani et al., 2017). Architectural variations have emerged, tailored to specific task types: encoder-only models like BERT excel at tasks requiring deep bidirectional context understanding (e.g., sentiment analysis, named entity recognition), while decoder-only models such as the GPT series are optimised for autoregressive text generation (e.g., language modelling and creative writing). Architectures combining both encoder and decoder components - for instance, BART and T5 - demonstrate particular effectiveness in sequence-to-sequence transformations (Radford et al., 2019; Lewis et al., 2020; Raffel et al., 2022; Rothman, 2022).

Beyond their origins in NLP, transformers exhibit remarkable versatility across diverse scientific and industrial fields. In computer vision, transformer models have achieved competitive, and often superior, performance compared to convolutional neural networks on large-scale image classification, segmentation, and object detection benchmarks. Transformers also underpin advancements in speech and audio processing, including automatic speech recognition and speaker identification systems (Islam et al., 2024). Furthermore, their capacity to model complex dependencies has led to applications in medicine, aiding the analysis of intricate medical imaging data for disease diagnosis and prognosis, and even in neuroscience and psychiatry for decoding neural signals and modelling cognitive processes (Cong et al., 2024). As research continues to expand its capabilities and efficiency, transformer architectures are poised to remain a central pillar of modern artificial intelligence (Patwardhan et al., 2023).

1.2. Diffusion models

Diffusion models represent a distinct and powerful class of generative methods that learn to synthesise new data by reversing a predefined noise-injection process (Yang et al., 2024). Conceptually, this involves a forward process that gradually adds noise (typically Gaussian) to the original data over a sequence of steps until it approximates pure noise. Subsequently, a learned reverse process iteratively denoises the corrupted data, starting from noise, to generate a realistic sample from the target distribution. Foundational work in score-based generative modelling, which involves learning the gradient of the log-density of the data distribution (the score function), provided key theoretical underpinnings for these approaches (Song & Ermon, 2019).

A significant milestone was the development of Denoising Diffusion Probabilistic Models (DDPMs), which reformulated the objective and demonstrated high-fidelity image generation by training a model to predict the noise added at each step of the forward process (Ho et al., 2020; Gallon et al., 2024). DDPMs and subsequent refinements have established diffusion models as state-of-the-art in various generative tasks. Further innovations, such as Latent Diffusion Models (LDMs), perform the diffusion and denoising operations within a lower-dimensional latent space, significantly reducing computational demands while largely preserving high-quality output synthesis (Po et al., 2024; Rombach et al., 2022).

Diffusion models have achieved exceptional performance across a diverse array of applications, particularly in visual computing. They excel at high-resolution image and video generation, complex image editing, inpainting, super-resolution, and notably, text-to-image synthesis, which has garnered significant popular and research interest (Po et al., 2024; Yang et al., 2024). Their generative capabilities extend beyond 2D images to tasks including 3D shape generation, audio synthesis, and motion generation (Po et al., 2024; Yang et al., 2024). Moreover, researchers are actively exploring the application of diffusion models to challenging interdisciplinary problems, such as molecular generation for drug discovery and material design, highlighting their potential as a versatile and broadly applicable generative AI paradigm (Yang et al., 2024).

1.3. Hugging Face platform

The Hugging Face (HF) platform has emerged as a crucial ecosystem facilitating the development, sharing, and deployment of contemporary AI models, encompassing hundreds of thousands of model, dataset, and interactive space repositories (Osborne et al., 2024). Its open-source libraries significantly lower the barrier to entry, democratising access to advanced AI for researchers and practitioners across various disciplines, including the social and behavioural sciences (Hussain et al., 2024).

The transformers library, in particular, provides standardised interfaces for a vast collection of prominent transformer architectures (e.g., BERT, GPT, and T5), supporting a wide array of NLP, vision, and speech tasks. Empirical evidence confirms the significant preference for this library in the reuse and adaptation (fine-tuning) of pre-trained models (PTMs) within the HF community (Jones et al., 2024). Similarly, the diffusers library provides a modular framework for diffusion models, simplifying experimentation with generative tasks like text-to-image synthesis through pre-configured pipelines, interchangeable noise schedulers, and access to numerous checkpoints.

The platform enables straightforward inference via the HF Hub and provides efficient fine-tuning tools, promoting widespread model accessibility. Beyond basic model use, Hugging Face offers robust support for transfer learning, optimisation, and deployment, allowing models to be adapted and utilised efficiently (https://huggingface.co/docs). This accessibility demonstrably accelerates innovation across numerous fields (Hussain et al., 2024).

However, quantitative analyses reveal highly skewed, Pareto-like distributions for nearly all activity metrics, including downloads, likes, and contributions, indicating that a small fraction of models and users account for the vast majority of engagement. Collaboration patterns also show a core-periphery structure, with most developers working in isolation but a densely connected core driving significant activity (Osborne et al., 2024). While the platform fosters community support, users face challenges related to model understanding, implementation, and documentation quality (Taraghi et al., 2024). Notably, documentation quality (e.g., via model cards) strongly correlates with model popularity (Jones et al., 2024), yet documentation practices remain inconsistent across the Hub (Taraghi et al., 2024). Furthermore, the platform experiences rapid model turnover, higher than traditional software registries (Jones et al., 2024), and lacks licensing information for a majority of hosted artefacts (Osborne et al., 2024).

Despite these complexities - including skewed engagement metrics, collaboration patterns, and documentation challenges - the sheer scale and central role of Hugging Face in the open AI landscape make it an invaluable data source for analysing contemporary AI model trends, adoption patterns, and community dynamics (Jones et al., 2024; Osborne et al., 2024). Understanding the distribution and engagement surrounding key architectures, such as transformers and diffusers, in this specific, influential context provides crucial insights into the trajectory of modern AI development and deployment.

2. Research methods

Data were retrieved between 15 January and 15 March 2025 using Python-based tools to combine API queries with targeted web scraping. The Hugging Face Hub library facilitated direct interactions with official endpoints, while Requests and BeautifulSoup handled custom HTML parsing. dateutil supported date transformations and concurrent.futures enabled asynchronous data retrieval. Collected records were stored and organised in pandas data structures for subsequent filtering and analysis.

In cases where publicly documented API calls were insufficient - particularly for restricted or gated models - the research team employed carefully configured web scraping. This dual strategy balanced efficiency (through official endpoints) with data completeness (through supplementary parsing). Error-handling mechanisms included retry loops, pagination checks, and backoff strategies to avoid rate-limit issues, ensuring a robust extraction pipeline. Duplicate entries were addressed by merging

Tab. 1. Data collection stages, tools, and key results

Stage	Tools	Results
1. Organisational data (API queries)	• Hugging Face API • concurrent.futures for parallel requests • retry mechanisms and logging	• queried 207,831 organisational accounts • gathered stats on models, datasets, spaces, creation dates, modifications • identified 6 main organisation types (companies, non-profits, communities, universities, classrooms, government) plus 31,414 unclassified • noted top companies by team size
2. User profiles (web scraping)	• Requests & BeautifulSoup for HTML parsing • dateutil for date handling • follower count extraction from user pages	• mapped 56,019 individual users • captured follower counts, organisational affiliations, and membership • detected popular creators
3. Model & user discovery (pipeline tags)	• multiple sorting/filter approaches • automated merging of pipeline tags • incremental extension of a “master” dataset	• broadened coverage of model IDs, including newly identified ones • ensured minimal duplication by merging metadata for overlapping entries • established foundations for collecting model-level details (library, tasks, licensing, usage stats)
4. Model metadata (API + targeted scraping)	• pandas transformations (drop duplicates, filter outliers) • merging pipeline tags • task-based grouping	• filtered out models with nonpositive storage sizes or missing fields, reducing total from 762,971 to 674,800 • focused on 316,418 models using essential libraries (transformers, diffusers) • yielded final dataset of 263,850 users (organisations + individuals)

metadata tags, and restricted models were flagged separately.

The retrieval pipeline unfolded in four distinct phases, summarised in Table 1. Each phase contributed specific data points (organisational details, user profiles, model identifiers, and metadata) that were ultimately merged into a comprehensive master dataset.

Following data retrieval and filtering, the final dataset captures a broad cross-section of users and models on Hugging Face. Initially, 762,971 models were identified; however, rigorous cleaning and the exclusion of incomplete or invalid entries narrowed this total to 674,800. A subsequent focus on recognised tasks (e.g., text-generation and object-detection) and two key libraries (transformers and diffusers) further reduced the number of models to 316,418. At the user level, 263,850 unique accounts emerged in the dataset, comprising organisational and individual profiles.

Within organisations, companies constituted the largest group-76,210 in total - accounting for 52,384 hosted models and attracting 458,233 followers. Many organisations hosted no models at all, but a small subset maintained large repositories. For instance, Mars Republic stood out with 2,549 members, while DeepSeek commanded the highest follower count at 45,440. Other notable companies included Cognizant, Amazon, Google, IBM, NVIDIA, Intel, Dell Technologies, Cisco, and HUAWEI Noah's Ark Lab, reflecting substantial corporate interest and capacity in AI development.

Individual users varied significantly in both popularity and productivity. While most contributed only a handful of models, a small but prolific group -such as the user "team mradermacher" (mradermacher) with 31,728 repositories - produced a disproportionate share of the platform's outputs. Prominent community figures included Tom Jobbins (TheBloke), followed by Lvmin Zhang (lllyasviel), Merve Noyan (merve), and AK (akhaliq), each drawing thousands of followers. This combination of diverse institutional participation and high-output individual contributors underscores the platform's openness and the collaborative dynamics driving model creation and sharing.

3. Research results

3.1. Overall model distribution

Table 2 presents a high-level overview of the distribution of AI models across major categories - multimodal, computer vision (CV), natural language processing (NLP), audio, tabular, reinforcement learning (RL), and other - categorised by their underlying library (transformers or diffusers).

Aggregating the data reveals a total of 316,418 models within the scope of this analysis, with transformer-based models constituting the vast majority (280,034 models, or approx. 88.5 %), while diffuser-based models represent a smaller but significant segment (36,384 models, approx. 11.5 %). The NLP category exhibits the highest concentration of transformer models (229,875 out of 229,883 total NLP models), reflecting this architecture's established dominance and efficacy in language-centric tasks such as text classification, question answering, and generation. Conversely, diffusers are predominantly utilised within the computer vision domain, where they constitute a substantial proportion (36,342 out of 51,605 CV models, approx. 70.4 %), primarily employed for generative processes, such as text-to-image synthesis and image-to-image translation.

Smaller yet noteworthy domains, such as audio and multimodal, display a more nuanced architectural division. While transformers overwhelmingly dominate the audio category numerically (26,849 vs

Tab. 2. Task-level breakdown of transformers and diffusers on Hugging Face

Category	Transformers	Diffusers	Total
Multimodal	7 129	8	7 137
Computer vision (CV)	15 263	36 342	51 605
Natural language processing (NLP)	229 875	8	229 883
Audio	26 849	20	26 869
Tabular	115	1	116
Reinforcement learning (RL)	787	2	789
Other	16	3	19
Total	280 034	36 384	316 418

Source: author's elaboration on the basis of data retrieved from Hugging Face (March 2025).

20 diffusers), the presence of diffusers, albeit small, points towards emerging applications in generative audio synthesis and augmentation. Similarly, the multimodal category, comprising over 7,000 models, is largely driven by transformers; however, the inclusion of diffusers suggests expanding use cases in complex scenarios, integrating diverse data types like text, images, and potentially video. Collectively, the distribution outlined in Table 2 underscores a key trend in contemporary AI development: established analytical and sequential tasks rely heavily on transformer architectures, whereas the burgeoning field of generative AI, particularly in visual domains, spurs the adoption and development of diffuser models.

3.2. Comparative model metrics

This section analyses the characteristics of models across different libraries and task categories by examining summary statistics for model size, citations, downloads, and likes, as presented in Table 3. These metrics reveal distinct profiles reflecting varying development priorities, user engagement patterns, and application focuses within the Hugging Face ecosystem.

Comparing the two primary libraries, transformer models are characterised by a substantially larger average size (avg. 15.05 GB) and extremely high size variability (std 200.25 GB), likely reflecting the prevalence of large foundational models alongside smaller, specialised ones. They garner significantly higher average downloads (avg. 83k) but show lower average community appreciation via likes (avg. 2.43) compared to diffusers. Diffuser models, conversely, are smaller on average (avg. 6.58 GB) with less, though still considerable, size variability (std 106.03 GB). They exhibit much lower average downloads (avg. 22k) but receive markedly higher average likes (avg. 5.57), suggesting a strong resonance with users, perhaps focused on generative novelty, despite lower overall download volume and very low average citations (avg. 0.06). This contrast highlights a potential divergence between broad utility (transformers) and focused community enthusiasm for specific capabilities (diffusers).

Examining the major task categories reveals further nuances. Natural language processing (NLP) models mirror the overall transformer profile with large average size (avg. 16.01 GB) and immense size variability (std 219.10 GB), moderate average downloads (avg. 65k), and relatively low average likes (avg. 2.47). Computer vision (CV) models are smaller on average (avg. 6.55 GB) but attract higher average downloads (avg. 90k) and notably more likes (avg. 4.49) than NLP models, despite having the lowest average citation count among major tasks (avg. 0.10). Audio models stand out for achieving the highest average downloads (avg. 160k), coupled with extreme download variability (std 15089k), yet they receive the lowest average likes (avg. 1.17) and relatively few citations (avg. 0.14).

Specialised and emerging domains present distinct statistical profiles. Multimodal models, despite lower average downloads (avg. 29k), boast the highest average citation count (avg. 0.75) and like count (avg. 7.21), alongside a large average size (avg. 17.61 GB), suggesting strong academic interest and community appreciation for models that bridge different data types. Tabular models demonstrate an exceptional profile: minimal average size (avg. 0.56 GB) but

Tab. 3. Summary of model statistics across libraries and tasks

Category		Size (GB)		Citations		Downloads (000s)		Likes	
avg.		std	avg.	std	avg.	std	avg.	std	
Library	transformers	15.05	200.25	0.29	0.92	83	6968	2.43	46.61
	diffusers	6.58	106.03	0.06	0.35	22	731	5.57	93.85
Tasks	Multimodal	17.61	70.70	0.75	2.83	29	420	7.21	70.55
	CV	6.55	90.55	0.10	0.41	90	5139	4.49	80.4
	NLP	16.01	219.10	0.30	0.86	65	5225	2.47	47.87
	Audio	11.66	73.79	0.14	0.37	160	15089	1.17	34.86
	Tabular	0.56	1.90	0.28	0.61	1570	10425	5.29	18.15
	RL	2.05	9.99	0.05	0.23	2	20	2.44	30.33
	Other	5.29	8.65	0.62	0.50	958	3720	11	19.73
All models		14.07	191.81	0.26	0.88	76	6560	2.80	54.18

Source: author's elaboration on the basis of data retrieved from Hugging Face (March 2025).

extraordinarily high average downloads (avg. 1570k) with significant variability (std 10425k), indicating immense popularity for specific, likely efficient, tabular data solutions. Reinforcement learning (RL) appears as a niche field, characterised by a small average size (avg. 2.05 GB), very low average downloads (avg. 2k), and the lowest average citations (avg. 0.05). The "Other" category, though small, shows surprisingly high engagement across downloads (avg. 958k), citations (avg. 0.62), and likes (avg. 11.00), potentially representing highly impactful or novel models outside standard classifications. Overall, these diverse statistical footprints underscore the varied nature of development focus, user needs, and community engagement across different AI application domains on the platform.

3.3. Top models by engagement

Among the most downloaded transformer models on Hugging Face, the leader is ast-finetuned-audioset (10-10-0.4593 version), developed by the Massachusetts Institute of Technology (MIT) and released in November 2022, with over 2.1 billion downloads. It is followed by bert-base-uncased, released by Google in March 2022, which has surpassed 2 billion downloads. Other highly downloaded models include wav2vec2-large-xlsr-53-english by Jonatas Grosman and clip-vit-large-patch14 by OpenAI, both released in March 2022, with over 1 billion and 860 million downloads, respectively. The list also features gpt2 by OpenAI community, resnet-50 by Microsoft, and xlm-roberta-large by FacebookAI, all of which were introduced in March 2022 (Fig. 1).

This ranking reflects the cumulative adoption of key transformer models since their respective release dates. While many of the top models were introduced in early 2022 - such as those from Google, OpenAI, Microsoft, and FacebookAI - their high download counts may reflect longer availability rather than current usage trends. Overall, these patterns highlight the long-term utility of foundational models and the rapid uptake of high-impact architectures, especially those supporting speech, vision, and language understanding.

For diffusers, the most downloaded models highlight the rapid growth of generative image modelling in recent years. Leading the group is stable-diffusion-xl-base-1.0, released by Stability AI in July 2023, with over 81 million downloads. It is followed by AnimateDiff-Lightning, developed by ByteDance and introduced in March 2024, reflecting the increasing interest in generative video tasks such as text-to-video synthesis. Other popular models from Stability AI include stable-diffusion-2-1 (released in December 2022) and sdxl-img2img-1.0 (July 2023), both of which focus on text-to-image and image-to-image generation, respectively.

When analysing user appreciation through likes (Fig. 2), a different picture emerges. Leading the ranking is DeepSeek-R1, introduced on Hugging Face on 20 January 2025, with over 11,000 likes. Developed by the Chinese company DeepSeek, the model has quickly gained popularity by matching the performance of models like ChatGPT and Gemini while remaining significantly more affordable. As an open-source model under the MIT license, it is available for commercial use and can be run locally, mak-

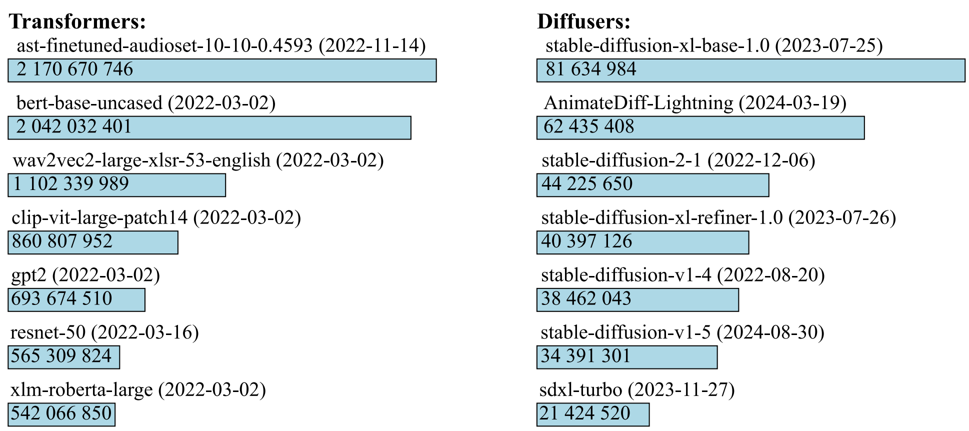

Fig. 1. Top models by downloads

Source: author's elaboration on the basis of data retrieved from Hugging Face (March 2025).

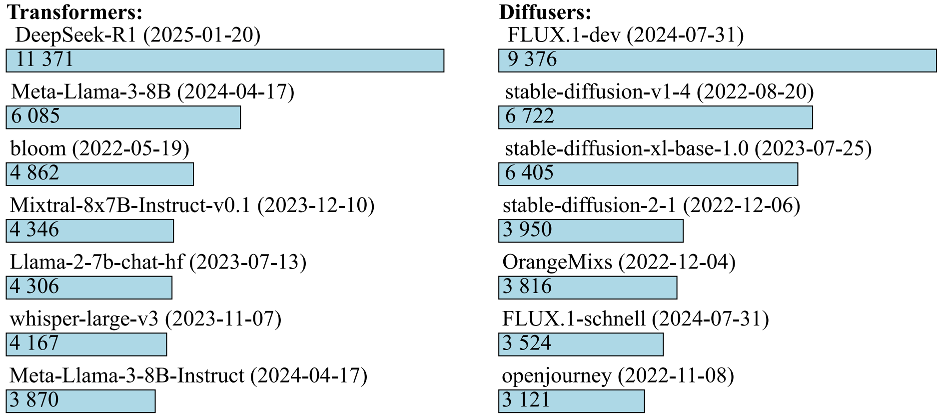

Fig. 2. Top models by likes
Source: author's elaboration on the basis of data retrieved from Hugging Face (March 2025).

ing it highly accessible (Lauridsen, 2025). Following DeepSeek-R are models such as Meta-Llama-3-8B and Meta-Llama-3-70B, both released in April 2024 by Meta Llama, and Mixtral-8x7B-Instruct-v0.1 by the French startup Mistral AI, launched in December 2023. These models, although not always the most downloaded, have earned significant user approval due to their technical sophistication and relevance to cutting-edge research. Their success highlights a clear trend: Hugging Face users gravitate toward models that represent the latest advancements in AI. Notably, OpenAI's whisper-large-v3, a speech recognition model released in November 2023, also ranks highly, reinforcing the preference for state-of-the-art performance across different application domains.

For diffusers, the model FLUX.1-dev by Black Forest Labs, created in July 2024, holds the top spot in likes, demonstrating significant community appreciation despite not being the most downloaded. Also highly rated are OrangeMixs by WarriorMama777 (released in December 2022) and openjourney by PromptHero (from November 2022), both of which are known for artistic and stylised image generation. These models serve more creative or experimental use cases and appear to resonate with users who value innovation, fine-tuned aesthetics, and novel diffusion capabilities. Other top liked models, such as FLUX.1-schnell (July 2024) and the stable-diffusion series, illustrate how the community appreciates both the refinement of established families and the introduction of newer, high-impact variations.

In summary, while download counts often reflect broad utility and cumulative adoption, particularly for foundational models, like counts appear more indicative of perceived innovation, state-of-the-art relevance, or unique capabilities. This distinction highlights both the enduring value of established models and the community's enthusiasm for novel architectures.

3.4. Correlation patterns across metrics

To deepen the understanding of user engagement and adoption dynamics, the author analysed Pearson correlation coefficients between key model-level metrics - downloads, likes, citations, and size - for transformer and diffuser libraries. These correlations help reveal whether popularity, scholarly recognition, or technical attributes tend to move together or reflect distinct use cases and community behaviours. The results, visualised in Fig. 3, uncover markedly different patterns for the two architectures.

For transformers, correlations across metrics are generally weak. The strongest relationship is observed between downloads and likes ($r = 0.16$), while downloads show virtually no correlation with citations ($r = 0.01$). Model size remains largely uncorrelated with other variables. This pattern suggests that adoption within the transformer ecosystem is broad and multifaceted, likely influenced by a wide variety of academic, industrial, and niche applications rather than simple engagement signals.

By contrast, diffusers exhibit notably stronger correlations - most prominently between downloads and likes ($r = 0.59$) - implying that widely used models in this category are also more likely to be positively received by the community. A moderate correlation also emerges between downloads and citations ($r = 0.14$), indicating some alignment between practical use and academic attention. As with transformers, model size remains a weak predictor. The tighter

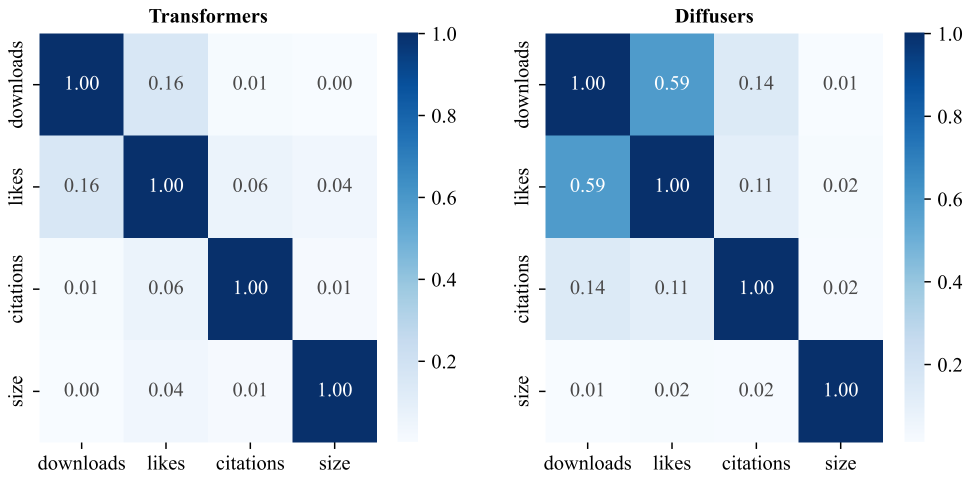

Fig. 3. Pearson correlation coefficients of model metrics for transformers and diffusers

Source: author's elaboration on the basis of data retrieved from Hugging Face (March 2025).

clustering of engagement metrics for diffusers may reflect a more focused and collaborative community, especially around fast-evolving generative capabilities where novelty and utility often go hand in hand.

4. Discussion of the results

4.1. Strategic implications for business adoption

The empirical findings on the distribution and engagement patterns of transformer and diffuser models on the Hugging Face (HF) platform carry direct implications for business adoption and strategic planning. The widespread prevalence of transformer architectures - particularly in NLP and established computer vision (CV) tasks - reflects their maturity and deep integration into applications that enhance operational efficiency and customer interaction. Advanced text-generation models (e.g., GPT-based architectures) are transforming areas such as marketing content creation, customer service automation (via conversational assistants), and internal knowledge management, offering cost savings alongside greater responsiveness and personalisation. Likewise, the high volume and download counts of foundational CV models underscore their practical value in domains such as automated visual inspection, media analysis, and security systems.

At the same time, the rapid growth and strong community engagement surrounding diffuser models signal their emergence as powerful engines of innovation, especially in creative industries. As demonstrated by their dominance in text-to-image tasks and their high average like counts, Diffusers excel in contexts where stylistic novelty and artistic experimentation are essential - such as design, advertising, entertainment, and media. Unlike traditional discriminative models, their ability to synthesise entirely new visual content or transform existing assets opens up new avenues for rapid prototyping (e.g., in fashion or product design), personalised content generation, and scalable aesthetic exploration.

Beyond platform-specific observations, the rise of generative AI is reshaping traditionally human-centred creative fields. In areas like luxury design, for instance, AI-generated outputs can now effectively convey brand identity and evoke targeted emotional responses (Pantano, Serravalle & Priporas, 2024). This evolution highlights AI's potential not only to automate but to genuinely augment human creativity - particularly within collaborative frameworks. However, this potential must be balanced with concerns around aesthetic homogenisation and the marginalisation of human artistry, calling for thoughtful, context-sensitive integration strategies.

The strong community enthusiasm for diffusers - despite their lower overall download volumes compared to transformers - suggests a highly engaged user base driving innovation in generative applications. This dynamic may serve as an early signal of emerging market opportunities, particularly in sectors seeking differentiation through bespoke content creation or AI-augmented creative workflows. Nonetheless, businesses should remain aware of the highly skewed adoption landscape on platforms like HF, where a small number of models accounts for a disproportionate share of engagement (Osborne et al., 2024). As such, initial adoption may concentrate around a handful of well-supported, widely validated models before broadening to more diverse use cases.

4.2. Challenges of enterprise-scale integration

Successfully integrating large-scale AI models, such as advanced transformers or diffusers, into enterprise environments requires careful consideration of several practical challenges. Scalability, encompassing computational resources and infrastructure costs, remains a primary concern. The significant size of many state-of-the-art models places substantial demands on cloud computing resources or on-premise hardware, potentially leading to high operational expenditures and latency issues. This finding aligns with challenges reported within the HF community regarding memory and performance limitations (Taraghi et al., 2024). Organisations must strategically balance model performance with infrastructure costs, often employing techniques, such as model distillation, quantisation, or efficient serving strategies, each involving trade-offs among accuracy, complexity, and maintainability.

Data governance, security, and regulatory compliance represent another critical dimension, particularly in sensitive sectors like healthcare, finance, and public services. Handling proprietary or personal data requires stringent privacy protocols (e.g., GDPR and HIPAA), robust audit trails, and secure deployment mechanisms. While open-source models offer advantages in transparency and the potential for local deployment to enhance data privacy (Hussain et al., 2024), significant gaps exist regarding licensing clarity for a large proportion of models hosted on platforms like HF (Osborne et al., 2024), creating potential legal risks for commercial use. Furthermore, the increasing focus on model security, evidenced by trends such as the adoption of formats like safetensors (Jones et al., 2024), highlights the need for vigilance against vulnerabilities. Failure to implement rigorous data controls and address licensing and security concerns can lead to severe reputational damage, legal liabilities, and erosion of stakeholder trust.

Finally, organisational readiness is paramount. Deploying and managing sophisticated AI models demands specialised expertise spanning data science and domain-specific knowledge. While platforms like Hugging Face provide user-friendly toolkits that lower entry barriers (Hussain et al., 2024), effective enterprise integration requires more than basic technical literacy. Users frequently encounter challenges in understanding model specifics, use nuances, and training pipelines (Taraghi et al., 2024), which underscores the critical importance of high-quality documentation - a factor strongly correlated with model popularity and adoption (Jones et al., 2024) - yet often found lacking or inconsistent (Taraghi et al., 2024). Organisations must invest in targeted talent development, either by hiring specialists or by upskilling existing teams in areas such as AI pipeline management, model lifecycle monitoring, ethical AI practices, and performance tuning. Fostering cross-functional teams that bridge technical depth with business acumen is crucial to unlocking the strategic value of these advanced AI tools.

4.3. Community engagement and development patterns

The vibrant user community surrounding platforms such as Hugging Face fosters a dynamic environment of rapid innovation and community-driven model improvement. As observed by Taraghi et al. (2024), the community provides significant benefits through shared expertise, collaborative problem-solving, and the contribution of new model checkpoints, fine-tuning scripts, and application examples. This collective, open approach can accelerate model refinement and performance gains for both transformer and diffuser architectures, allowing businesses to potentially benefit from cutting-edge AI developments without bearing the full cost of proprietary R&D, particularly when leveraging adaptable open-source models (Hussain et al., 2024). However, this collaborative picture is nuanced by findings suggesting a core-periphery structure in development activity, with influence concentrated among a few prolific developers and large organisations (Osborne et al., 2024), serving as a reminder that the open AI ecosystem is not immune to the dynamics seen in traditional open-source software.

The ongoing evolution of generative AI, driven in part by diffusers, is poised to reshape business models, particularly in the creative and digital sectors. As these models become more accessible and efficient, opportunities arise for novel generative services - from automated content creation and personalised advertising to virtual prototyping and environment design. This trend, coupled with lower entry barriers facilitated by platforms like HF, may fuel a "creator economy" dynamic, enabling smaller entities to compete by leveraging powerful generative tools. Furthermore, the high turnover rate observed in models on HF (Jones et al., 2024) suggests a rapidly evolving landscape, in which adaptability and continuous learning are key strategic imperatives for businesses seeking to maintain a competitive edge.

Looking ahead, the convergence of analytical capabilities (often associated with transformers) and generative prowess (increasingly linked to diffusers) likely points towards the development of powerful hybrid architectures. Such models could offer deep domain-specific understanding and fluid, multimodal generation capabilities, further transforming strategic planning, product innovation, and human-AI collaboration across diverse industries. Harnessing the potential of this evolving ecosystem requires not only technical proficiency but also strategic foresight regarding community trends, ethical considerations, and the shifting competitive landscape.

Conclusions

This study sought to illuminate the contemporary landscape of artificial intelligence development and adoption by quantitatively analysing the distribution, use patterns, and community engagement surrounding two pivotal architectures, transformers and diffusers, within the influential Hugging Face (HF) ecosystem. By examining a large-scale dataset derived from the platform, the research aimed to provide empirical insights into current trends, practical implications, and the evolving dynamics shaping the open AI movement.

Addressing the first research question concerning model distribution, the analysis confirms the marked dominance of transformer architectures, particularly within established domains, such as natural language processing (NLP) and audio processing. This underscores their enduring efficacy for tasks involving sequential data and complex language understanding. Conversely, diffusion models demonstrate significant and rapidly growing traction in computer vision (CV), especially for generative tasks such as text-to-image synthesis. This bifurcation highlights a key trend: while transformers remain foundational for many analytical AI applications, diffusers are spearheading advancements in creative content generation, particularly in the visual domain.

Regarding user engagement patterns (the second research question), the findings reveal distinct characteristics for each architecture. Transformer models collectively achieve higher aggregate download volumes, likely reflecting their broad applicability and deep integration into existing research and enterprise workflows. However, diffuser models exhibit stronger correlations between engagement metrics (likes, downloads, and citations), suggesting a potentially more cohesive and actively engaged community where perceived innovation, popularity, and academic relevance are more closely intertwined. The enthusiastic reception of diffusers for novel content creation points towards a dynamic user base focused on exploring capabilities beyond traditional AI tasks.

Examining broader trends and the alignment between model development and evolving work paradigms (the third research question), the data signal accelerating innovation trajectories, especially in creative industries, driven by the proliferation of powerful generative models. Emerging user preferences lean towards multimodal capabilities and accessible, customisable solutions, fostered by a vibrant community-driven ecosystem on platforms like HF that encourages rapid model refinement. The findings depict a shifting landscape where advanced generative architectures are increasingly integrated into design workflows, collaborative platforms, and diverse application frameworks. This emergent synergy between human expertise and AI-driven automation, as evidenced on HF, suggests a future where generative capabilities significantly augment professional roles and catalyse novel commercial opportunities and artistic domains.

This investigation offers valuable empirical insights into the proliferation and adoption dynamics of key AI architectures on the Hugging Face platform, highlighting the differential impact of transformers and diffusers on research agendas and commercial strategies. Nevertheless, the findings should be interpreted cautiously, acknowledging several limitations. The study's exclusive reliance on HF data means it may not fully capture developments within proprietary systems or other open-source communities. Furthermore, inherent platform factors, such as algorithmic visibility bias and potential metadata inconsistencies, could influence engagement metrics, requiring care when equating metrics like downloads directly with real-world utility or impact. Future research incorporating cross-platform analyses, longitudinal studies, and external performance validation is essential for a more comprehensive understanding of the AI landscape.

As generative models become increasingly embedded within organisational and societal functions, the imperative for responsible development and deployment intensifies. Addressing critical concerns related to bias, privacy, fairness, and labour market impacts requires robust ethical frameworks and proactive governance, alongside compliance with evolving regulatory requirements. Successfully navigating these complex challenges necessitates sustained interdisciplinary collaboration, bridging technical, managerial, ethical, and legal expertise. Ultimately, as Sedkaoui and Benaichouba (2024) suggest, realising the transformative potential of AI hinges not solely on technological advancement, but on collective, concerted efforts by all stakeholders - researchers, developers, policymakers, and the public - to foster inclusive, sustainable, and genuinely beneficial innovation for both industry and society at large.

Acknowledgement

The publication presents the results of the project financed from the subsidy granted to the Krakow University of Economics.

Literature

Ahirwar, K. (2023). *A very short introduction to diffusion models.* Retrieved from https://kailashahirwar.medium.com/a-very-short-introduction-to-diffusion-models-a84235e4e9ae

Chen, M., Mei, S., Fan, J., & Wang, M. (2024). Opportunities and challenges of diffusion models for generative AI. *National Science Review*, *11*(12). doi: 10.1093/nsr/nwae348

Cong, S., Wang, H., Zhou, Y., Wang, Z., Yao, X., & Yang, C. (2023). Comprehensive review of Transformer-based models in neuroscience neurology and psychiatry. *Brain and Behavior, 13*(2). doi: 10.1002/brx2.57

Gallon, D., Jentzen, A,. & von Wurstemberger, P. (2024). An overview of diffusion models for generative artificial intelligence. *ArXiv*, *12*(2024). doi: 10.48550/arXiv.2412.01371

Gupta, P. (2023). *Transformer models: A breakthrough in artificial intelligence.* Retrieved from https://medium.com/%40prashantgupta17/transformer-models-a-breakthrough-in-artificial-intelligence-e3de92d37f8f

Ho, J., Jain, A., & Abbeel, P. (2020). Denoising diffusion probabilistic models. *Advances in neural information processing systems*, *33*, 6840-6851.

Hussain, Z., Mata, R., Binz, M., & Wulff, D. U. (2024). A tutorial on open-source large language models for behavioral science. *Behavior Research Methods*, *56*, 8214-8237.3 doi: 10.3758/s13428-024-02455-8

Islam, S., Elmekki, H., Elsebai, A., Bentahar, J., Drawel, N., Rjoub, G., & Pedrycz, W. (2024). A comprehensive survey on applications of transformers for deep learning tasks. *Expert Systems with Applications*, *213*, 122666. doi: 10.1016/j.eswa.2023.122666

Jones, J., Jiang, W., Synovic, N., Thiruvathukal, G. K., & Davis, J. C. (2024). What do we know about Hugging Face? A systematic literature review and quantitative validation of qualitative claims. In *Proceedings of the 18th ACM / IEEE International Symposium on Empirical Software Engineering and Measurement* (ESEM '24) (pp. 18-14). doi: 10.1145/3674805.3686665

Lauridsen, P. S., (2025). *DeepSeek: Potential and challenges in education*. Retrieved from https://viden.ai/en/deepseek-potential-and-challenges-in-teaching/

Lewis, M., Liu, Y., Goyal, N., Ghazvininejad, M., Mohamed, A., Levy, O., Stoyanov, V., & Zettlemoyer, L. (2020). BART: Denoising sequence-to-sequence pre-training for natural language generation, translation, and comprehension. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics* (pp. 7871-7880). Association for Computational Linguistics.

Osborne, C., Ben Allal, L., & Cihon, P. (2024). The AI community building the future? A quantitative analysis of development activity on Hugging Face Hub. *Journal of Computational Social Science*, *7*, 2067-2105.4 doi: 10.1007/s42001-024-00300-8

Pantano, E., Serravalle, F., & Priporas, C.-V. (2024). The form of AI-driven luxury: How generative AI (GAI) and Large Language Models (LLMs) are transforming the creative process. *Journal of Marketing Management*, *40*(17-18), 1771-1790. doi: 10.1080/0267257X.2024.2436096

Patwardhan, N., Marrone, S., & Sansone, C. (2023). Transformers in the real world: A survey on NLP applications. *Information*, *14*(4), 248. doi: 10.3390/info14040248

Po, R., Yifan, W., Golyanik, V., Aberman, K., Barron, J. T., Bermano, A. H., Chan, E. R., Dekel, T., Holynski, A., Kanazawa, A., Liu, C. K., Liu, L., Mildenhall, B., Nießner, M., Ommer, B., Theobalt, C., Wonka, P., & Wetzstein, G. (2023). *State of the art on diffusion models for visual computing*, *43*(2). doi: 10.48550/arXiv.2310.07204

Radford, A., Wu, J., Child, R., Luan, D., Amodei, D., & Sutskever, I. (2019). *Language models are unsupervised multitask learners*. OpenAI blog, *1*(8), 9.

Raffel, C., Shazeer, N., Roberts, A., Lee, K., Narang, S., Matena, M., Zhou, Y., Li, W., & Liu, P. J. (2020). Exploring the limits of transfer learning with a unified text-to-text transformer. *Journal of Machine Learning Research*, *21*(140), 1-67.

Rombach, R., Blattmann, A., Lorenz, D., Esser, P., & Ommer, B. (2022). High-resolution image synthesis with latent diffusion models. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* (pp. 10684-10695). IEEE. doi: 10.1109/CVPR52688.2022.01042

Rothman, D. (2022). *Transformers for Natural Language Processing (2nd ed.)*. Birmingham, UK: Packt Publishing.

Sedkaoui, S., & Benaichouba, R. (2024). Generative AI as a transformative force for innovation: A review of opportunities, applications and challenges. *European Journal of Innovation Management*. doi: 10.1108/EJIM-02-2024-0129

Song, Y., & Ermon, S. (2019). Generative modeling by estimating gradients of the data distribution. In H. Wallach, H. Larochelle, A. Beygelzimer, F. d'Alché-Buc, E. Fox, & R. Garnett (Eds.), *Advances in Neural Information Processing Systems 32* (pp. 11895-11907). New York, USA: Curran Associates Inc.

Taraghi, M., Dorcelus, G., Foundjem, A., Tambon, F., & Khomh, F. (2024). Deep Learning Model Reuse in the HuggingFace Community: Challenges, Benefit and Trends. In *2024 IEEE International Conference on Software Analysis, Evolution and Reengineering (SANER)* (pp. 512-523). IEEE.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). Attention is all you need. *Advances in Neural Information Processing Systems*, *30*, 5998-6008. doi: 10.48550/arXiv.1706.03762

Yang, L., Zhang, Z., Song, Y., Hong, S., Xu, R., Zhao, Y., Zhang, W., Cui, B., & Yang, M.-H. (2023). Diffusion models: A comprehensive survey of methods and applications. *ACM Computing Surveys*, *55*(6), 1-35. doi: 10.48550/arXiv.2209.00796