The shift: the evolution of design & architecture under the generative AI influence. Safety and implications.

Arch. Monica-Mihaela Frangulea, Ph.D
CEO Juxta Foundation, Bucharest, Romania
ceo@juxta.ro

Alexandru Lincă, IT engineer
Founder Fondator WPhosting, Bucharest, Romania
alex@wph.ro

Abstract
The evolution of architecture from the peak of modernism in the early 20th century to the present has been a journey from strict, functional minimalism to an eclectic, technology-driven, and sustainable approach. This field has transitioned from the rigid "less is more" philosophy of early modernists to postmodern reactions and, ultimately, to contemporary designs that embrace complex digital forms and ecological responsibility. Nevertheless, while architects and designers often prefer Modernism, Brutalism, or Contemporary styles, these often rank lower in popularity among the general public. According to recent studies, classical, and traditional architectural styles are the ones generally liked by the majority of people, showing a preference for symmetry, balance, and intricate detailing. While modern or contemporary styles are often chosen for new builds due to their efficiency, lower construction costs and easier construction effort, public sentiment strongly favors older, historical detail-rich aesthetics, such as Gothic and Art Nouveau. The introduction of AI-assisted software in architectural practice has brought a dramatic shift in the industry. The shift is due not only to the increased speed of concept image production by generative AI but also to the complexity of details and the multitude of visual references proposed by AI platforms. Generative AI is a subset of artificial intelligence that creates new content by learning patterns from existing data using large-scale machine learning models. These models are trained on massive datasets to understand underlying structures, allowing them to predict and create new, similar data. As the general public shows a clear tendency of favoring older aesthetics, the AI engines are bringing those visual trends into the proposed generated images. This phenomenon has already led to the shift of direction in the architectural style evolution, and we can see this type of aesthetics in awarded projects of major architectural contests, urban planning proposals, and even in real life with many already-built structures around the world. But the interaction with the AI assisted software is not without risks. The European Commission is heavily focused on copyright concerns arising from generative AI and is taking steps to implement and reinforce legislation, with major enforcement actions planned for 2026. Companies like Anthropic are also flagging lately potential risks such as the “zero-day vulnerability”, leaking proprietary Anthropic information, sabotaging code, or steering humans toward wrong decisions. Project Glasswing is a very recent (April 2026) cybersecurity initiative launched by Anthropic, as an initiative to use next-generation AI tools for defensive cybersecurity across critical infrastructure. This initiative can also ensure a secure evolution path for the architecture industry that is shaping the future image of our living environment and cities. It seeks to prevent harmful influences or interventions from various groups or individuals who may prioritize the financial or industrial interests of a select few, while neglecting the aesthetic preferences of the majority of the population.

Keywords: AI, generative design, cybersecurity, GenAI, AI Act

1. Introduction
The evolution of architecture from the peak of modernism in the early 20th century to the present has been a journey from strict, functional minimalism to an eclectic, technology-driven, and sustainable approach. This field has transitioned from the rigid "less is more" philosophy of early modernists to postmodern reactions and, ultimately, to contemporary designs that embrace complex digital forms and ecological responsibility.[1]

1.1. Architecture preferred by the majority VS promoted architecture

Even if architects like Zaha Hadid, Frank Gehry, and Bjarke Ingels push the boundaries of design, creating structures that challenge traditional forms and integrate cutting-edge technology, the general style of promoted and awarded architecture leans toward simplicity, simple shapes and minimalism.

For example, this year, the prestigious Pritzker Architecture Prize was awarded to architects like Liu Jiakun (born in 1956 in Chengdu, People's Republic of China)[2] and Smiljan Radić Clarke, of Santiago, Chile. [3]

The Pritzker Architecture Prize is usually conferred "in acknowledgment of talent, vision, and commitment, which have persistently produced significant contributions to humanity and the built environment through the art of architecture."[2]

The jury citation about the entire work of Liu Jiakun was that "through an outstanding body of work of deep coherence and constant quality, Liu Jiakun imagines and constructs new worlds, free from any aesthetic or stylistic constraint."[2]



Fig. 1. Design Department on new campus, Sichuan Fine Arts Institute designed by architect Liu Jiakun
Source: photo courtesy of Jiakun Architects



Fig. 2. Shuijingfang Museum designed by architect Liu Jiakun
Source: photo courtesy of Arch-Exist

The jury citation about the architectural concept of Smiljan Radić was that "in his designs he works with dimensions of experience that are immediately palpable but escape verbalization—like the perception of time itself: immediately recognizable, yet conceptually evasive."[3]



Fig. 3. Teatro Regional del Biobío designed by architect Smiljan Radić
Source: photo courtesy of Iwan Baan

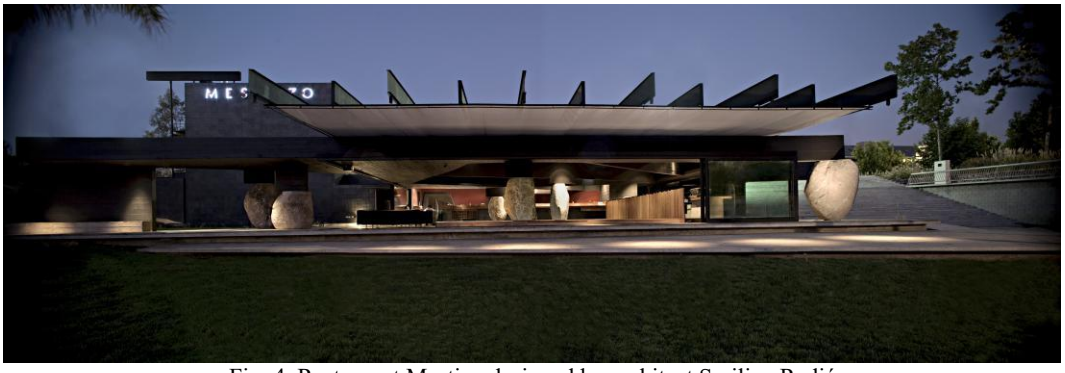


Fig. 4. Restaurant Mestizo designed by architect Smiljan Radić
Source: photo courtesy of Gonzalo Puga

According to recent studies, classical, and traditional architectural styles are the ones generally liked by the majority of people, showing a preference for symmetry, balance, and intricate detailing. While modern or contemporary styles are often chosen for new builds due to their efficiency, lower construction costs and easier construction effort, public sentiment strongly favors older, historical detail-rich aesthetics, such as Gothic, Art Nouveau, Art Deco, Organic and Victorian, which are perceived as having more "character".

So, while architects and designers often prefer Modernism, Brutalism, or Contemporary styles, these often rank lower in popularity among the general public.

According to recent studies, classical, and traditional architectural styles are the ones generally liked by the majority of people, showing a preference for symmetry, balance, and intricate detailing. While modern or contemporary styles are often chosen for new builds due to their efficiency, lower construction costs and easier construction effort, public sentiment strongly favors older, historical detail-rich aesthetics, such as Gothic, Art Nouveau, Art Deco, Organic and Victorian, which are perceived as having more "character".

1.2. The choice of the majority. Research and studies.

In 2025, The Aesthetic City commissioned Ipsos I&O, the leading research institute in the Netherlands, to conduct the first comprehensive national study on Dutch architectural preferences. A representative sample of 770 respondents viewed eight pairs of buildings, with each pair comparing a recent modernist design to a traditional or classical reinterpretation of the same structure. The results were notable: in every comparison, the traditional version was favored, with support levels ranging from 67 to 92 percent.[4]



Fig. 5. Sample of a pair of buildings from "A study by The Aesthetic City, conducted by Ipsos I&O" with the question "Which of the two buildings below appeals to you most?": (a) modern; (b) classical.
Source: https://theaestheticcity.com/research/

The results of a survey of 200 adults from across the U.S. revealed that though a small number of people that are active and vocal agents of historic preservation, a large majority of the public stands behind this movement. On average, people prefer old buildings over new, value historic places, and are generally supportive of historic preservation.[5]

In 2020, an academic study of courthouse architecture found that the study's participants preferred neoclassical designs for courthouses. The study's author, Jack L. Nasar, Academy Professor of City & Regional Planning at Ohio State University, concluded, "The findings agree with consistent findings that architects misjudge public likely public impressions of a design, and that most non-architects dislike 'modern' design and have done so for almost a century." The key findings were that traditional architecture tops modern architecture for Americans. When it comes to the style of architecture Americans favor for U.S. courthouse or federal buildings, traditional is the clear winner across the board with more than 7 in 10 Americans (72%) preferring a traditional look.[6]

In 2007, the American Institute of Architects (AIA) undertook a survey of 2,200 random Americans, also conducted by The Harris Poll. According to the Wall Street Journal, the poll showed that "Americans preferred older buildings that evoke ancient architectural styles such as Gothic, Greek, and Roman traditions."[7]

In 2023, a similar survey was conducted in Romania by PhD architect Iulia-Maria Coțovanu Toader from SDA-UAUIM ("Ion Mincu" University of Architecture and Urbanism - Bucharest) and dr. Corina Ilin, Asociate Professor of Psychology, West

University of Timișoara. The general objective of the study was to see to what extent certain architectural images can create attraction or repulsion. The study revealed some interesting facts, as for example the message of the discourse strongly related to modernism causes a negative response in relation to the respondents' intention to visit the building (the modern glass building shown was a horizontal skyscraper – Vanke Center Shenzen by arhitect Stephen Hall).



Fig. 7. A part of the Horizontal Skyscraper project by Steven Holl Architects, when it was under construction, Vanke Center, Shenzhen, China
Source: https://lebbeuswoods.wordpress.com/2011/06/28/steven-holls-horizontal-skyscraper/

The reaction to the image of the Beijing CCTV Building was similar, the respondents identifying it as rather unvisitable and perceiving it as an object of cold modern architecture.

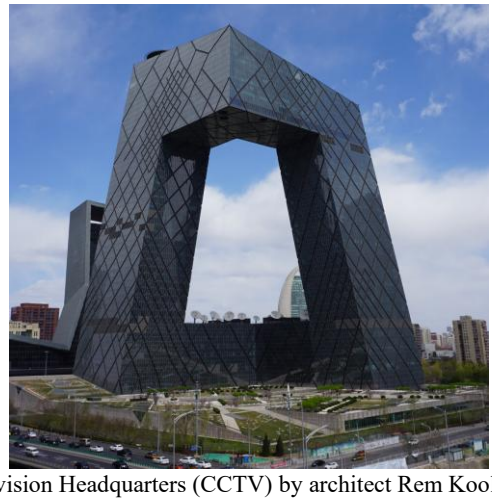


Fig. 8. Central Television Headquarters (CCTV) by architect Rem Koolhaas, Beijing, China
Source: By Dayton12345 - Own work, CC BY-SA 4.0, https://commons.wikimedia.org/w/index.php?curid=141768792

A totally different reaction has been observed related to the image of the ancient Roman Market Gate of Miletus. With a maximum total of 82.7% intention to visit from the respondents, the highest during this survey, this building was a winning image formula to attract visitors to it and generate attachment for the respondents. [8]



Fig. 9. The Gate of Miletus in Pergamon Museum in Berlin, Germany
Source: By Kameister - Own work, CC BY-SA 4.0, https://commons.wikimedia.org/w/index.php?curid=74507606

Using Linkfluence sentiment tool, which crawls data on mentions across the internet from the past 12 months, a new research by comparethemarket.com.au has exposed the architectural with the largest percentages of positive and negative sentiments across Europe, Canada, United States and Australia. Comparing the data across the globe, Gothic is the most positively spoken of architectural design in 13 nations.

On the other hand, Colonial architecture was spoken most negatively about and was the least favoured architectural style in 23 out of the 54 countries analysed, including Greece, Canada and Iceland. Postmodern and Modern architecture also made up the top three styles that sparked the most negative conversations, dividing the globe.[9]

Even major public opinion debates platforms such as Reddit are showing similar conclusions. Based on discussions across Reddit communities like r/ArchitecturalRevival (87,000 weekly visitors, 1,400 weekly contributions) , r/architecture, and r/AskReddit, the most beloved architectural styles tend to be Gothic (including Neo-Gothic/Collegiate Gothic) and Art Deco. Reddit users generally express a strong preference for ornate, historic, or detailed aesthetics over minimalist or "modernist" concrete structures.[10]

1.3. Science Fiction architecture on Earth

Key Preferences in Public Architecture are symmetry and order. People tend to prefer buildings that are organized and balanced, rather than chaotic.

Natural materials such as wood, stone, and brick are generally favored over raw concrete or steel.

Also, contrary to the "less is more" philosophy of modernists, the general public tends to enjoy details, decoration, and ornamentation.

Those visual characteristics were also replicated in science-fiction movies, video games, and virtual reality environments, accessed by a large group of the population.

A recent TV show, "The Peripheral", launched in 2022 by Amazon Prime, depicts an alternative city, London of a virtual future, dominated by classical idols reconfigured as urban landmarks reclaiming a cultural identity.[11]

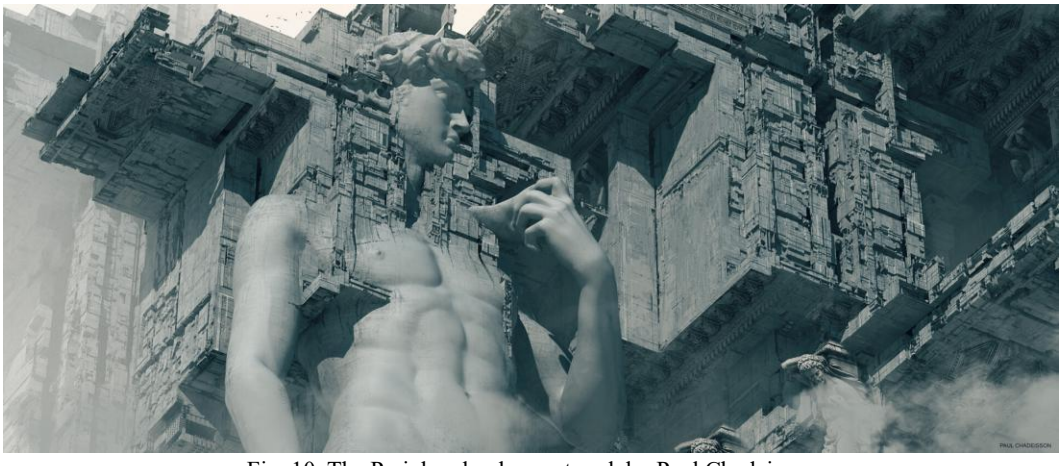


Fig. 10. The Peripheral - show artwork by Paul Chadeisson
Source: https://pao.artstation.com/projects/4XJvA1?album_id=9790482

Paul Chadeisson, the concept artist for "The Peripheral", a renowned French digital artist specializing in industrial and sci-fi environments, noted that the design idea stemmed from a futuristic nostalgia, where survivors brought back historic art as an homage to past civilizations.

The introduction of AI-assisted software in architectural practice has brought a dramatic shift in the industry. The shift is due not only to the increased speed of concept image production by generative AI but also to the complexity of details and the multitude of visual references proposed by AI platforms.

Generative AI is a subset of artificial intelligence that creates new content by learning patterns from existing data using large-scale machine learning models. These models are trained on massive datasets to understand underlying structures, allowing them to predict and create new, similar data.[12]

As the general public shows a clear tendency of favoring older aesthetics, the AI engines are bringing those visual trends into the proposed generated images. This phenomenon has already led to the shift of direction in the architectural style evolution, and we can lately see this type of aesthetics in awarded projects of major architectural contests, urban planning proposals, and even in real life with many already-built structures around the world.

For example, in 2023, for a conference that architect Monica-Mihaela Frangulea had to present at the Romanian National Library in Bucharest, a series of interior design proposals were necessary for the concept of future public libraries. When requested to generate a futuristic public library interior, the DaVinci AI image generating platform did not choose to go towards a simplified brutalist or minimalistic look but on the contrary, it generated some intricate curved lines and organic shaped proposals with gothic influence and/or art nouveau features.



Fig. 11. DaVinci AI platform rendering of a public library interior by architect Monica-Mihaela Frangulea
Source: Monica-Mihaela Frangulea

And lately we start to find such visual similarities in the real world. At the Yuxin Bookstore in Shenzhen, space folds into itself. Designed by X+Living Architecture and Interior Design (based in Shanghai) and completed in 2019, the interior draws on the logic of a Möbius strip, in which floor, wall, and ceiling merge into a single continuous surface.



Fig. 12. a.Yuxin Bookstore in Shenzhen designed by X+Living Architecture and Interior Design
Source: https://www.instagram.com/parametric.architecture/
b. Zhongshuge bookstore in Huai'an designed by X+Living Architecture and Interior Design
Source: https://www.xl-muse.com/html/en/index.php?ac=article&at=read&did=350

Architectural models & AI renderings led by the architect Sera Su Abaç from Zaha Hadid Architects are proposing structures that are following an organic style inspired by the famous Antoni Gaudí architecture from the beginning of the 19th century.


Fig. 13. Architectural Modeling & AI Renderings by architect Sera Su Abaç from Zaha Hadid Architects,
Source: https://www.instagram.com/p/DR5HkZMDJq_/?img_index=3

The Museum of Contemporary Art and Planning Exhibition (MOCAPE) in Shenzhen, China, have turned into reality this style revival trend. Designed by Coop Himmelb(l)au and completed in 2016, this cultural complex's glowing metallic shell and overlapping forms give it the aura of a tech archive or galactic museum. Its cyberpunk aesthetic fits right into Coruscant's entertainment districts or Canto Bight's elite zones.[13]

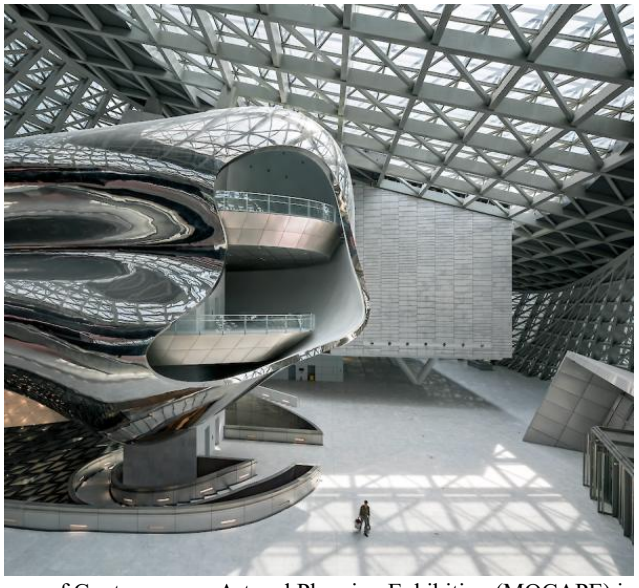


Fig. 14. The Museum of Contemporary Art and Planning Exhibition (MOCAPE) in Shenzhen, China
Source: https://coop-himmelblau.at/projects/museum-of-contemporary-art-planning-exhibition-mocape/

This year, a pavilion made with 3D-printed ceramics was created by Synthetic Architecture for a design week exhibition. In both traditional and modern ceramic patterns, the designers considered that blue and its shades often bring a feeling of calm and quiet and also carry a soft, poetic link to origins and roots. Synthetic Architecture is a design studio with a conceptual language where AI merges a collective design intelligence, weaving the future through a dialogue between innovation and the vernacular.

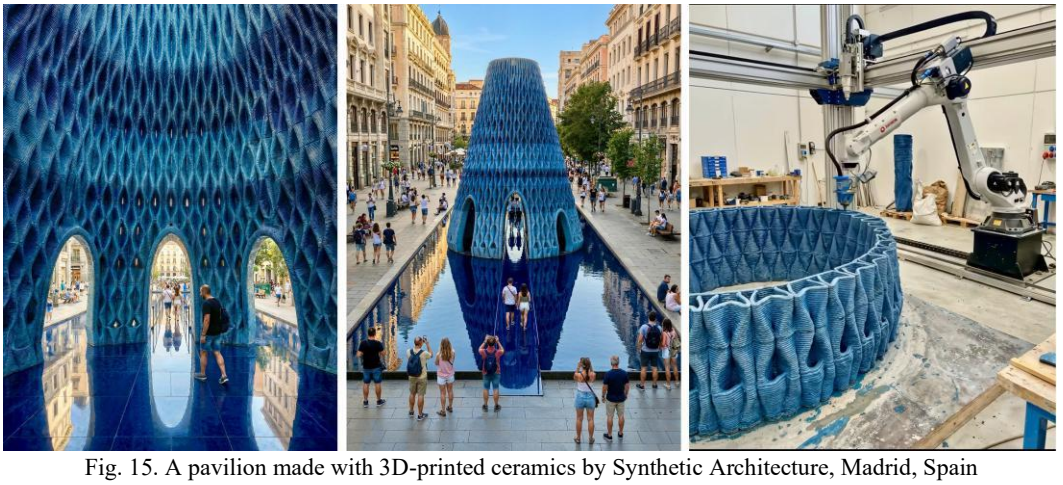


Fig. 15. A pavilion made with 3D-printed ceramics by Synthetic Architecture, Madrid, Spain
Source: https://www.instagram.com/syntheticarchitecture/

1.4. A.I. involvement in enhancing efficiency within the design and construction industry

New and complex AI assisted systems are being created in order to bring higher designing process speed and better efficiency in costs and production processes.

Urban narratives, visual references, and contextual data related to density, climate, energy, and spatial organization are curated into datasets that inform AI model training and structured prompting strategies. These systems generate architectural massing and typological variations without direct formal intervention.

Exporting Aesthetics examines how artificial intelligence shifts architectural authorship from drawing buildings to designing systems for the development of one of the most compressed and rapidly evolving urban regions in the world, the Greater Bay Area in Guangzhou (China), characterized by extreme population density, infrastructural intensity, and continuous redevelopment. Presented at the 2025 Hong Kong–Shenzhen Bi-City Biennale of Urbanism and Architecture, the exhibition features a series of large-scale, AI-generated tower models developed through fully automated workflows. The building structures (the towers) emerge from datasets, prompts, and recursive machine logic. Regional narratives of the area are translated into speculative architectural forms through computation, rather than direct human design authorship. A key goal of Exporting Aesthetics is to question traditional notions of architectural authorship. Human involvement is deliberately restricted. Rather than creating forms through design intuition, the team establishes datasets, prompt structures, and selection criteria. Consequently, architecture arises as a byproduct of system behavior.

The project poses a new and important question: when architecture is generated by machines trained on cultural, spatial, and economic information, where does authorship reside, and what kinds of aesthetics are produced? [14]

2. Risk and challenges of using AI in architecture and construction industry

In the implementation framework of AI-assisted technologies, a great challenge is cultural integration monitoring, as the establishment of feedback mechanisms to track cultural alignment throughout AI implementation phases ensures technological and cultural development remain synchronized, allowing for continuous adaptation and alignment between technological advances and cultural expectations.[15]

Recently, the AI research company Anthropic just published a major report on how AI could affect jobs. One common approach is to compare outcomes between more or less AI-exposed workers, firms, or industries in order to isolate the effect of AI from confounding forces. This study shows that the architecture and engineering industry faces the highest levels of risk related to loss of human intervention, which may lead to increased unemployment in those sectors.

Our measure qualitatively captures several aspects of AI usage that we think are predictive of job impacts. A job's exposure is higher if:

- Its tasks are theoretically possible with AI;

- Its tasks see significant usage in the Anthropic Economic Index;
- Its tasks are performed in work-related contexts;
- It has a relatively higher share of automated use patterns or API implementation;
- Its AI-impacted tasks make up a larger share of the overall role. [16]

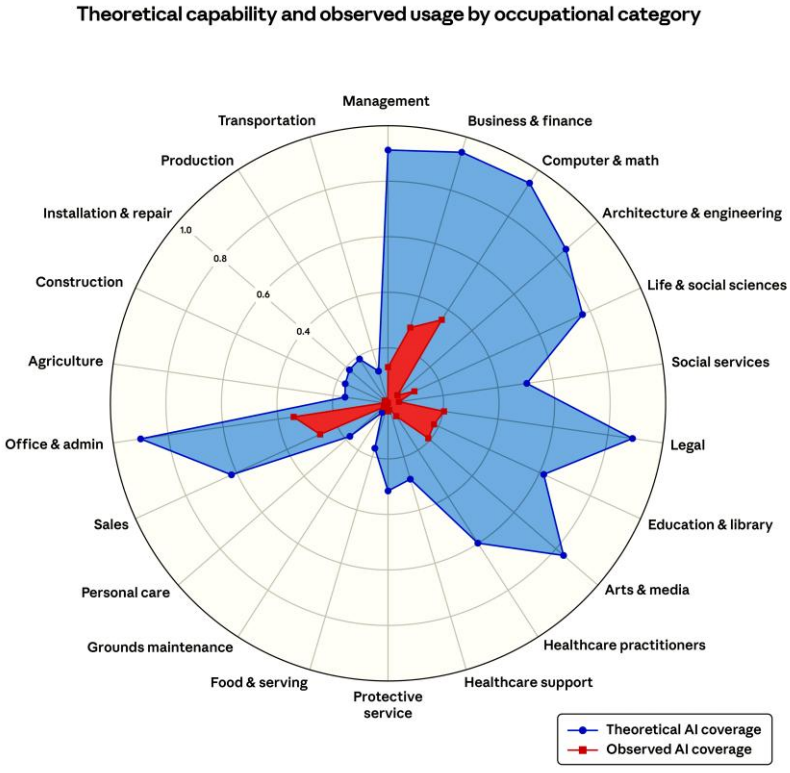


Fig. 16. Theoretical capability and observed exposure by occupational category.
Share of job tasks that LLMs could theoretically perform (blue area) and our own job coverage measure derived from usage data (red area).
Source: https://www.anthropic.com/research/labor-market-impacts

One of the most concerning aspects is that AI models can already automate significant portions of work today. However, legal constraints and slow adoption by companies are hindering progress. Thus, the issue is not solely related to skills; it is also about adoption.

The European Parliament has signaled that it wants to shift from a "consultative" approach to strict enforcement to ensure that AI innovation does not come at the expense of European creators.

The European Commission is also heavily focused on copyright concerns arising from generative AI and is taking steps to implement and reinforce legislation, with major enforcement actions planned for 2026. The key initiative in this area is the EU AI Act, which entered into force in August 2024 and introduces mandatory transparency and copyright compliance for generative AI models.[17]

Those measures are taken in order to insure creators protection, to avoid copyright infringements and protect intellectual property.

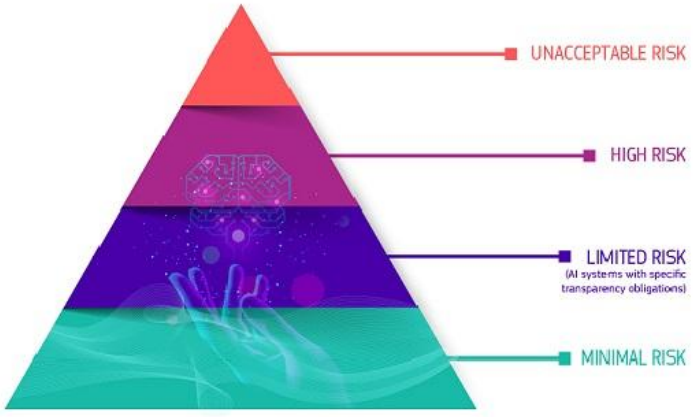


Fig. 17. The AI Act 4 levels of risk defined for AI systems
Source: European Commission

Key areas of concern and legislative action include:

- Transparency and Training Data: Under the AI Act, providers of General Purpose AI (GPAI) models must disclose a detailed summary of the copyrighted content used to train their models;
- Copyrighted Opt-Outs: AI developers are required to respect rights reservations (opt-outs) made by creators under the 2019 Digital Single Market (DSM) Directive, meaning creators can block their work from being used in AI training;
- Enforcement Timeline: While the AI Act is already in force, enforcement of these copyright provisions by the European Commission's AI Office will start in mid-2026;
- Mandatory Labeling: Starting August 2, 2026, AI-generated content must be clearly labeled to prevent deception and protect intellectual property;
- Parliamentary Pressure: In March 2026, the European Parliament passed a resolution calling for even stronger measures, including a "rebuttable presumption" of infringement if AI providers fail to be transparent;
- Upcoming Review: The Commission is required to conduct an overall assessment of the 2019 DSM Directive no sooner than June 7, 2026, which may lead to further revisions. [17]

In conclusion, starting in 2026, the EU AI Act will require every AI company to disclose training data sources, respect copyright opt-outs, and label AI-generated content.

These laws are not just regulatory hurdles. They mark a strategic turning point for the AI industry.

Companies that prepare early will have higher trust from users and enterprise clients, access to EU markets without legal barriers and a stronger compliance reputation with investors

and regulators. It also means that those who delay can face last-minute compliance costs and possible bans from the European market.[18]

3. Generative AI challenges and limitations

Generative AI faces several challenges and limitations, such as:

- Resource requirements. Generative AI models demand substantial computational power and energy to run, making them costly and environmentally impactful to train and use;
- Lack of transparency. Lack of transparency in the internal processes and decision-making pathways of models makes it hard to understand how outputs are generated. This questions the ability to hold AI accountable;
- Complexity of creativity. Because AI-generated content is often the result of learned patterns rather than true innovation, generative AI still falls short of capturing the complexity of human creativity;
- Hallucinations. Generative AI models can produce plausible but false information, leading to potential misinformation. This poses a significant risk to critical fields such as journalism, healthcare, and education;
- Biases. Training data can contain inherent biases, resulting in outputs that perpetuate stereotypes, further marginalizing certain groups;
- Questions regarding intellectual property. Ethical concerns regarding copyright, authorship, ownership, and accountability are an ongoing debate regarding the intense use of AI, signaling a need for responsible AI practices and frameworks;
- The potential for misuse. Because of its ability to create extremely realistic deepfakes, generative AI can be used for malicious purposes, such as misleading content, propaganda or scams (flagged by the Organisation for Economic Co-operation and Development as the highest risk presented by generative AI in achieving national and regional goals - see fig. 17).

According to the Organisation for Economic Co-operation and Development (OECD), More than one-third of individuals across the OECD used generative AI tools in 2025, highlighting how rapidly AI is becoming part of everyday life.

Nevertheless, G7 members see disinformation and the associated manipulation of opinions as the biggest risk posed by generative AI. Most G7 members also consider intellectual property right infringement and threats to privacy as major risks.[19]

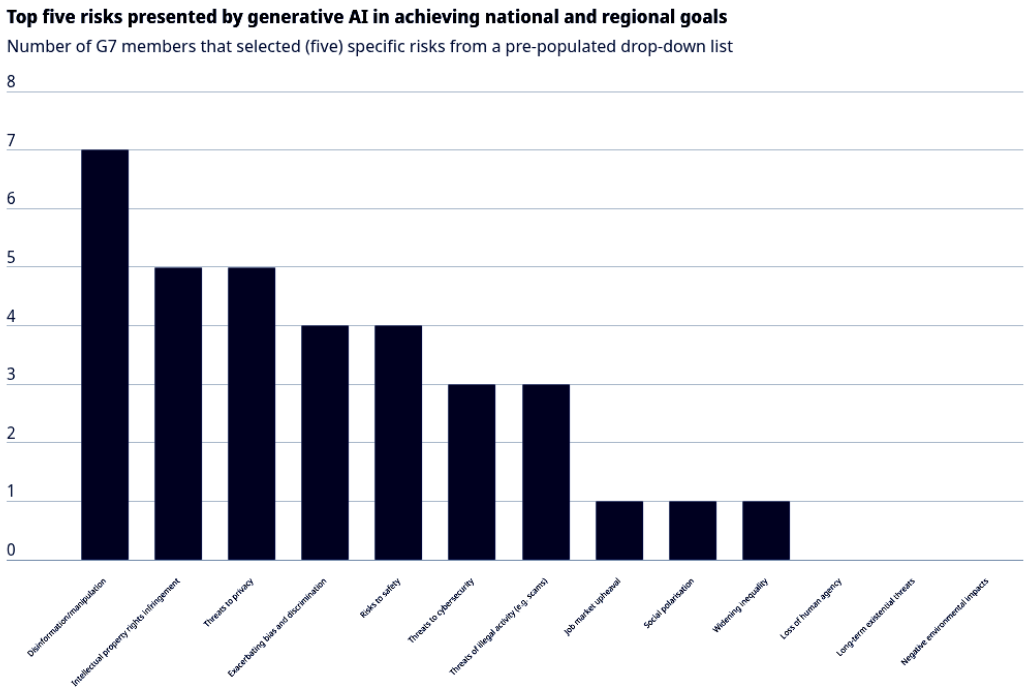


Fig. 18. Top five risks presented by generative AI in achieving national and regional goals
Source: Organisation for Economic Co-operation and Development

4. Cybersecurity risks and potential threats for Generative AI

The main risk that the generative AI engines are facing regarding the visual material (images / videos) proposed / generated for the professionals in the architectural field is the corruption of the AI machine learning algorithm. Such a cybersecurity breach altering the interaction between AI and the user (the user interface) can lead to misleading proposed content, favoring the financial or industrial interests of certain interested entities.

The proposed / generated images can start, for example, to predominantly feature certain types of materials, similarities with certain products (available for purchase in the market - so a masked/hidden advertisement), even certain general features such as design style, population density (extremely crowded structures for public or private use, specific to certain types of economies and construction financial benefits), a.s.o.

Companies like Anthropic are flagging lately potential risks such as the “zero-day vulnerability”, leaking proprietary Anthropic information, sabotaging code, or steering humans toward wrong decisions.

The new Claude Mythos Preview model performs strongly across the board, but it is strikingly capable at computer security tasks.

For example the Project Glasswing is a very recent (April 2026) cybersecurity initiative launched by Anthropic, as an initiative to use next-generation AI tools for defensive cybersecurity across critical infrastructure. It centers around a new frontier AI model called Claude Mythos Preview, which is being deployed specifically for security purposes.[20]

This project brings together major tech companies including Amazon Web Services, Apple, Broadcom, Cisco, CrowdStrike, Google, JPMorgan Chase, the Linux Foundation, Microsoft, NVIDIA, and Palo Alto Networks. Anthropic is also extending access to over 40 additional organizations that build or maintain critical software infrastructure.

Mythos Preview has already been used to identify thousands of previously unknown (zero-day) vulnerabilities in every major operating system and web browser. The model is powerful enough to both find and help fix flaws in important software at scale.

A comprehensive Sabotage Risk Report was released recently by Anthropic, defining "sabotage risk" as the risk of misaligned autonomous actions by an AI model that could contribute significantly to later catastrophic outcomes. The report lays out the existing level of risk rather than claiming a pre-determined safety threshold has been met. Mythos Preview has already been used to identify thousands of previously unknown (zero-day) vulnerabilities in every major operating system and web browser. The model is powerful enough to both find and help fix flaws in important software at scale.[21]

The report covers nine specific risk pathways, including things like leaking proprietary Anthropic information, sabotaging code, or steering humans toward wrong decisions.[22]

The Independent Reviews METR conducted also an external review, broadly agreeing that catastrophic risk from Claude Opus 4.6 is very low but not negligible, while noting some areas where the analytical reasoning could be strengthened — particularly around the sensitivity of the alignment assessment.[23]

The core idea is that an AI capable enough to break software is also the most powerful tool to secure it. Project Glasswing aims to give defenders a durable advantage before attackers can exploit these same AI capabilities. [24]

Hopefully, such cybersecurity initiatives may be capable to ensure a secure evolution path for the architecture industry that is shaping the future image of our living environment and cities, by seeking to prevent harmful influences or interventions from various groups or individuals who may prioritize the financial or industrial interests of a select few, while neglecting the aesthetic preferences of the majority of the population.

5. Advanced Security Risks: Generative Engine Optimization (GEO)

Usually, when limitations of artificial intelligence (AI) are discussed, the focus is on bias, lack of transparency, or the heavy requirements in terms of data. However, a new category

of risk with AI is emerging that is not only harder to recognize but also, in many respects, more difficult to manage.

So, what is Generative Engine Optimization, or GEO?

Simply put, GEO is the manipulation of the knowledge base of AI systems. As AI becomes architects' partners with their creativity, suggesting designs, recommending materials, or giving technical advice, it is becoming more and more dependent on external sources such as articles, product descriptions, case studies, and even forum discussions. The dilemma is that these tools lack a genuine 'comprehension' of the truthfulness of the content.

5.1. The Consensus Trap

The simplest method to understand GEO is through a phenomenon called “the consensus trap”. If the same concept is repeated in different documents, the AI will consider it to be a trustworthy statement.

In the field of architecture, this could result in certain issues remaining unnoticed but having a major impact in the long term. Picture a material that is constantly being spoken of on the internet as environmentally friendly, long-lasting, and commonly used. Even if these statements are a bit overstated or even untrue, the AI system will still be 'convinced' of the material's virtues, simply because it finds it mentioned frequently in many different sources.

Everything seems to be in order externally. The content is nicely written, the points are logical, and there is nothing that would make one suspect doubleness or falseness. But in truth, the machine is only echoing what it is 'most frequently exposed to'.

This dependence leads to a significant weakness: AI systems do not check the truth like humans do; they just detect patterns and agreement throughout the data they have.

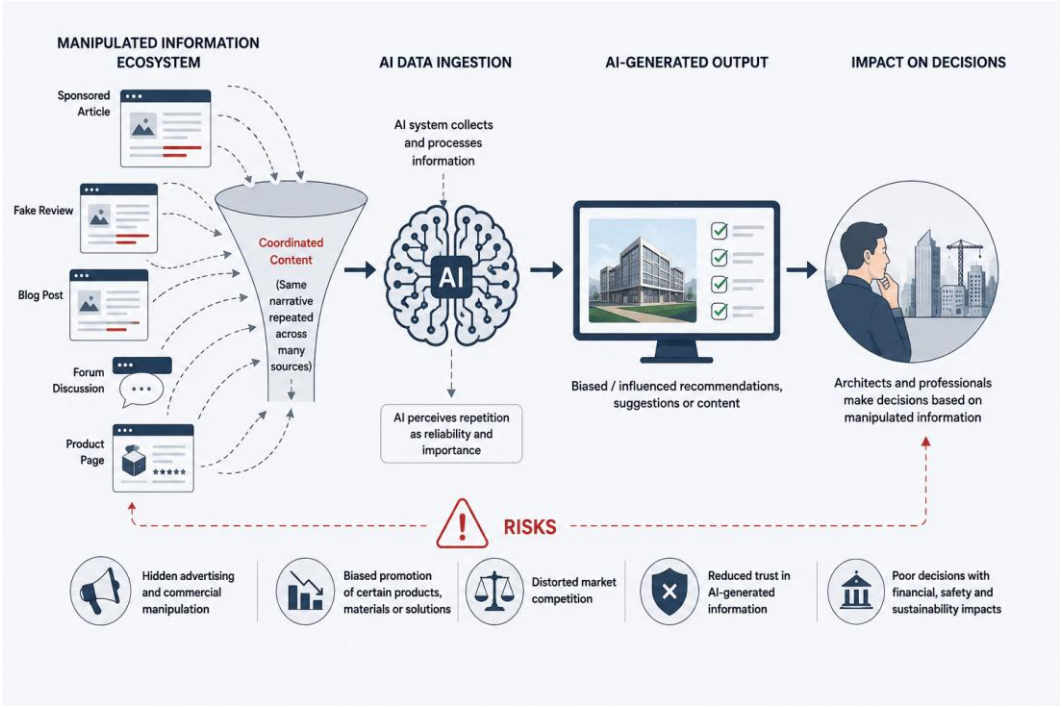


Fig. 19. Generative Engine Optimization (GEO) and its influence on AI-assisted architectural decisions.
Source: Alexandru Lincă

5.2. How the Consensus Trap Works in Architectural Contexts

One of the main things that drive GEO is the trap of consensus. When the same ideas are given again and again in different sources, AI systems usually take this as a sign that the information must be true.

In case of architectural works, this may result in some dangerous actions.

For example, if a certain material is always described in different papers and review as a "very sustainable" or a "leader of the industry".

5.3. Content: How GEO Shapes Architectural AI Systems

Different from a direct attack, GEO doesn't just focus on the one malicious input. Instead, it gradually manipulates the very environment of the data.
Most GEO scenarios entail:

- producing a great deal of technical-sounding material;
- scattering it over a number of platforms (blogs, forums, review sites, etc.);
- pretending that these are independent confirmations;
- letting AI systems learn and give priority to these materials.

Repeatedly being exposed to basically the same stories makes it higher the chance that an AI system will view them as trustworthy and use them in its responses.

In the realm of architecture, it is to say that AI-assisted design-making which is based on statistics rather than on real facts might be the result.

5.4. Implications for Architectural Practice

The influence of GEO is not limited to driving misinformation but can more so impact physical architectural realizations.

Here are some of the risks that are highly relevant:

- Distorted design decisions suggestions made by AI could be based on information that has been tampered with or is not complete;
- Material misselection products that are not appropriate for the specific use may be recommended to architects;
- Commercial bias for a particular vendor or technology could be excessively featured;
- Reduced trust in AI systems as frequent encounters with faulty outputs may result in users losing faith in tools that are AI-assisted.

GEO is so alarming because the outputs generated often look logical, well thought out, and professionally expressed, thus, the manipulation is hard to be identified.

5.5. Why GEO Isso Difficult to Recognize

One of the big challenges with GEO is that it's very subtle. Prompt injections or changing system errors produce obvious anomalies, but GEO will hardly show you any differences. In fact, the information given:

- is grammatically very correct;
- looks technically feasible;
- is very much consistent with different sources.

Therefore, the response of the AI will not be on any angle or situation for a wrongdoing. It is a reflection of a reality which is very subtly altered by coordinated manipulation of data.

5.6. Mitigation and Awareness

In order to fix the GEO problem, we need to completely change how AI systems are created and incorporated in architectural workflows. Instead of working out the model only, one must also consider the quality and origin of the data being fed to the model.
Some of the ways to go about it are:

- effort should be put on the sources that are reliable and can be verified for technical information;
- AI-generated suggestions should be verified against other sources independently;
- keep human intervention in making critical design decisions;
- promote openness of AI systems in the process of choosing information and laying it out.

After all, GEO points out a major change in AI risk: dependability of AI systems is not only about their algorithms, but also about the truthfulness of the information environments from which they draw.

6. Final thought

In architecture related case, decisions often have economic, functional, and safety consequences for a long period of time which means ignoring such a risk cannot be done. However, as a result of AI playing a more critical role in the design process, apart from their assessing technical aspects, it will be necessary to make sure that the informational inputs provided by them are trustworthy.

References

[1] SDH_STUDIO, The Evolution of Contemporary Architecture: A Bold Vision for the Present and Future - Modern vs Contemporary Architecture: A Tale of Two Styles, March 1, 2025, https://sdhstudio.com/what-is-contemporary-architecture/#:~:text=Modern%20architecture%20design%20emerged%20in,is%20a%20dynamic%2C%20evolving%20practice. , date: 10.04.2026.

[2] The Pritzker Architecture Prize, Laureates, https://www.pritzkerprize.com/laureates/liu-jiakun#laureate-page-2726 , date: 13.04.2026.

[3] The Pritzker Architecture Prize, Laureates, https://www.pritzkerprize.com/laureates/smiljan-radic-clarke#laureate-page-2786 , date: 13.04.2026.

[4] The Aesthetic City, Landmark study: Architectural Preference of the Dutch, 2025, https://theaestheticcity.com/research/ , date: 17.04.2026.

[5] Shannon, D. S. (2014). A survey of the public: Preference for old and new buildings, attitudes about historic preservation, and preservation-related engagement, Master's thesis, University of Southern California.

[6] A National Civic Art Society Survey Conducted by The Harris Poll, Americans' Preferred Architecture for Federal Buildings, October 2020, https://www.civicart.org/americans-preferred-architecture-for-federal-buildings , date: 17.04.2026.

[7] Frangos, A. (2007), Americans' Favorite Buildings, The Wall Street Journal, (illustrated sortable list), https://www.wsj.com/public/resources/documents/info-poparch07-sort2.html , date: 17.04.2026.

[8] Coțovanu T. I.-M, (2023), Manipulation through architecture, PhD Thesys , https://www.uauim.ro/doctorat/sustineri/409/ , date: 17.04.2026.

[9] Vivid Maps, The Most Loved Architectural Styles Across the World Mapped, 2022, https://vividmaps.com/most-loved-architectural-styles/ , date: 17.04.2026.

[10] Rredit, r/ArchitecturalRevival, https://www.reddit.com/r/ArchitecturalRevival/ , date: 17.04.2026.

[11] Frangulea, M.-M., Toader, I. (2024), Industrial afterlife: a tale of three cities, Smart Cities International Conference (SCIC) Proceedings. 11, pp. 87–104.

[12] Microsoft, How does generative AI work?, https://www.microsoft.com/en-us/ai/ai-101/how-does-generative-ai-work#:~:text=What%20is%20generative%20AI?%20Generative%20AI%20refers,mirrors%20human%20expression%E2%80%94redefining%20our%20relationship%20to%20technology. , date: 17.04.2026.

[13] Parametric Architecture, Star Wars Architecture on Earth: 10 Buildings From a Galaxy Not so Far Away, https://parametric-architecture.com/star-wars-architecture-on-earth/ , date: 17.04.2026.

[14] Chaudhary, I. (2026), Exporting Aesthetics: AI-Generated Futures of the Greater Bay Area, Parametric Architecture, https://parametric-architecture.com/exporting-aesthetics-ai-futures/ , date: 17.04.2026.

[15] Benshams, R.M. (2025), Cultivating organizational culture for AI integration: A framework for Smart Cities and Regional Development, Smart Cities and Regional Development (SCRD) Journal. 9, 4 , pp. 125–137. DOI:https://doi.org/10.25019/cmstt369.

[16] ANTHROPIC, Labor market impacts of AI: A new measure and early evidence, May 5 2026, https://www.anthropic.com/research/labor-market-impacts , date: 18.04.2026.

[17] European Commission, AI Act, https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai#:~:text=Application%20timeline,August%202026%2C%20with%20some%20exceptions: , date: 19.04.2026.

[18] Scalevise, EU AI Act 2026: What Every AI Company Must Prepare For, https://scalevise.com/resources/eu-ai-act-2026-changes/#:~:text=EU%20AI%20Act%202026:%20What,impact%20is%20anything%20but%20simple. , date: 19.04.2026.

[19] Organisation for Economic Co-operation and Development, Generative AI, https://www.oecd.org/en/topics/generative-ai.html#:~:text=GenAI%20has%20the%20potential%20to:%20*%20Revolutionize,*%20Perpetuating%20biases%20*%20Disinformation%20*%20Deepfakes , date: 20.04.2026.

[20] Carlini, N; Cheng, N; Keane, L, et alii (2026), Assessing Claude Mythos Preview's cybersecurity capabilities, Anthropic, https://red.anthropic.com/2026/mythos-preview/ , date: 20.04.2026.

[21] Sabotage Risk Report: Claude Opus 4.6, Anthropic, https://www-cdn.anthropic.com/f21d93f21602ead5cdbecb8c8e1c765759d9e232.pdf ,date: 20.04.2026.

[22] GradientDissenter, Summary and Comments on Anthropic's Pilot Sabotage Risk Report, Lesswrong, 30 October 2025, https://www.lesswrong.com/posts/wPQaeMvHXq9Ee2tEm/summary-and-comments-on-anthropic-s-pilot-sabotage-risk ,date: 20.04.2026.

[23] Jurkovic, N, Hjalmar W, Beth B, Charles F, Michael C.(2026), Review of the Anthropic Sabotage Risk Report: Claude Opus 4.6, METR, https://metr.org/blog/2026-03-12-sabotage-risk-report-opus-4-6-review/ ,date: 20.04.2026.

[24] Ozarslan, S. (2026), The Glasswing Paradox: The Thing That Can Break Everything Is Also The Thing That Fixes Everything, Picus Security, https://www.picussecurity.com/resource/blog/anthropics-project-glasswing-paradox ,date: 20.04.2026.