DATA DRIVEN DESIGN: NMETHODS IN ARCHITECTURAL CONCEPT DESIGN

UNDERGROUND CLOUD: DATA CENTER

Sergejs Kopils, Efe Duyan,

Rudolfs Dainis Smits

Sergejs Koplis, MArch, Faculty of Architecture and Design, RISEBA University, Latvia. Efe Duyan, Assoc. Prof. Dr., RISEBA University, Latvia. Rudolfs Dainis Smits, March, RISEBA University, Latvia.

ABSTRACT

This article explores the role of DATA DRIVEN DESIGN Methods in Architectural Concept Design of underground cloud data centers, with a particular focus on the "Underground Cloud" project. As data centers continue to serve as critical nodes for global digital infrastructure, their energy consumption and environmental impact raise pressing challenges. Data-driven design, incorporating advanced computational tools, artificial intelligence (AI), and predictive analytics, offers a solution to optimize spatial layouts, improve energy efficiency, and minimize environmental footprint. This research examines the potential of underground environments, which provide natural thermal stability and enhanced physical security, to support sustainable data center operations. The study analyzes the integration of big data in architectural workflows, highlighting the transition from intuition-based design to data-driven strategies that leverage real-time data inputs (Figure 1). Through case studies of existing underground facilities, the article illustrates best practices in applying AI-driven simulations and Building Information Modeling (BIM) for optimizing environmental performance. The findings suggest that data-driven methodologies can significantly reduce energy consumption, streamline operations, and address spatial constraints while maintaining architectural integrity. Furthermore, the paper presents design recommendations for future underground data centers, emphasizing the importance of interdisciplinary collaboration, adaptive design strategies, and ethical considerations in data usage. The study concludes that the intersection of data-driven design and underground data center architecture represents a promising avenue for addressing the growing demands of the digital era while promoting sustainability and operational resilience. Additionally, the research underscores the potential for underground cloud infrastructure to contribute to urban energy ecosystems by utilizing excess heat for district heating, reducing environmental impact, and enhancing resource efficiency. The study also explores potential challenges and future directions for integrating data-driven methods into architectural design, providing a comprehensive framework for future projects.

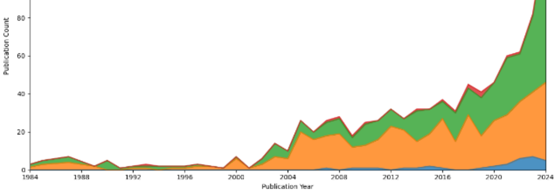


Figure 1. Plotting the dynamics publication by year “Data-Driven-Design”. [diagram] Sergejs Kopils, 2025

KEYWORDS

data-driven design, underground data centers, artificial intelligence in architecture, environmental sustainability, spatial optimization, building information modeling (BIM), digital infrastructure (Figure 2).

INTRODUCTION

The digital age has ushered in an unprecedented growth in data generation, processing, and storage, leading to a significant rise in the demand for efficient and sustainable data centers. As the backbone of modern digital infrastructure, data centers support a wide range of activities, from cloud computing and big data analytics to artificial intelligence (AI) applications and real-time communication systems. However, the rapid proliferation of data-driven technologies presents significant architectural challenges, including spatial constraints, energy consumption, and environmental impact. Traditional above-ground data centers, while effective in meeting operational demands, often struggle with issues related to land use, energy efficiency, and thermal management. As such, architects and engineers are exploring innovative design strategies that leverage data-driven methodologies to address these challenges while ensuring optimal performance and sustainability.

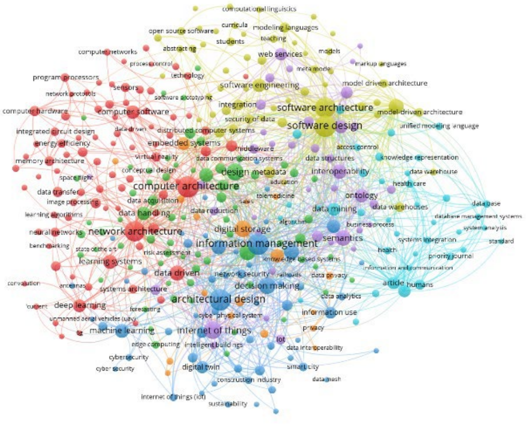


Figure 2. Diagram of connectivity of search words. [diagram] Sergejs Kopils, 2025

This research paper delves into the concept of data-driven design in architecture, with a specific focus on its application to underground cloud data centers. Data-driven design involves the systematic collection, analysis, and application of data to inform architectural decisions, optimizing both spatial organization and operational performance. By utilizing tools such as Building Information Modeling (BIM), Computational Fluid Dynamics (CFD) simulations, and AI-driven design algorithms, architects can create environments that respond dynamically to changing conditions and performance requirements. The shift from intuition-based design processes to data-informed methodologies represents a paradigm shift in the architectural profession, emphasizing precision, adaptability, and sustainability.

The choice to explore underground cloud data centers stems from their potential to address several critical challenges associated with traditional data centers. Underground environments offer inherent thermal stability, reducing the need for energy-intensive cooling systems and enhancing overall energy efficiency. Additionally, subterranean data centers minimize land-use conflicts, protect critical infrastructure from environmental and human-made threats, and provide opportunities for innovative urban integration through heat recovery systems for district heating networks. Despite these advantages, designing and constructing underground data centers presents unique challenges, including ventilation requirements, structural integrity considerations, and accessibility issues.

This study aims to investigate the principles, methods, and tools of data-driven design in the context of underground cloud data centers. The research focuses on understanding the interaction between environmental, technological, and architectural factors in shaping sustainable and efficient subterranean facilities. Through case studies and empirical analysis, the paper will examine successful implementations of data-driven design, highlighting best practices and identifying potential obstacles. The findings are intended to offer actionable insights for architects, engineers, and policymakers seeking to optimize data center design while minimizing environmental impact.

The structure of this paper is divided into four main sections. The first section provides a theoretical overview of data-driven design principles, tracing their evolution and examining the role of big data, AI, and advanced simulation tools in contemporary architectural practice. The second section focuses on data center design, exploring typologies, environmental considerations, and the specific challenges of underground facilities. The third section presents methodologies for applying data-driven techniques to architectural design, emphasizing the importance of environmental, behavioral, and operational data inputs. The final section discusses empirical findings from case studies, evaluates the effectiveness of the proposed design strategies, and offers recommendations for future research and practice.

THE DATA-DRIVEN DESIGN IN ARCHITECTURE CONCEPT DESING

The Data-driven design in architectural concept design represents a transformative approach that integrates empirical data into the architectural workflow. Historically, architecture has been guided by intuition and traditional practices (Figure 3 and 4), but the advent of big data and computational tools has shifted this paradigm towards evidence-based methodologies.

Herbert A. Simon (1988) emphasized the importance of 'sciences of the artificial' in professional education, highlighting the need for systematic, data-informed processes in design. The rise of big data, as noted by Qabshoqa, Kocaturk, and Kiviniemi (2017), has profoundly influenced architectural practices by providing insights into spatial relationships, environmental conditions, and user behavior.

Big data plays a critical role in contemporary architectural design by facilitating more informed decision-making processes. According to Batty (2013), the integration of urban data (Figure 5 and 6) allows architects to analyze patterns, predict spatial performance, and optimize design solutions. In the context of data centers, big data enables the optimization of server room layouts, cooling system configurations, and energy management strategies. For instance, Wong and Fan (2013) demonstrated how Building Information Modeling (BIM) enhances communication between stakeholders and streamlines project coordination by consolidating diverse data streams into a unified platform.

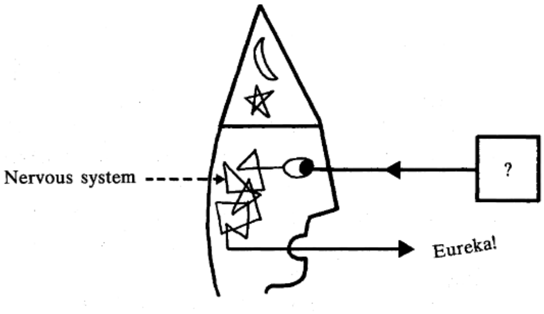


Figure 3. The designer as magician.[diagram]
Herbert, 1988

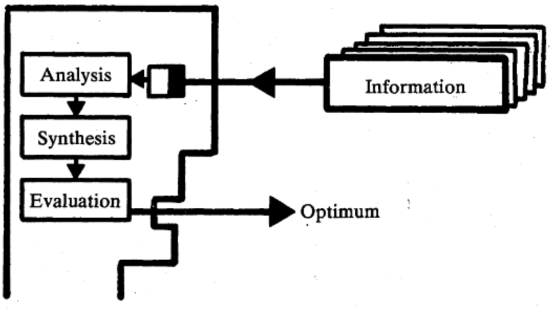
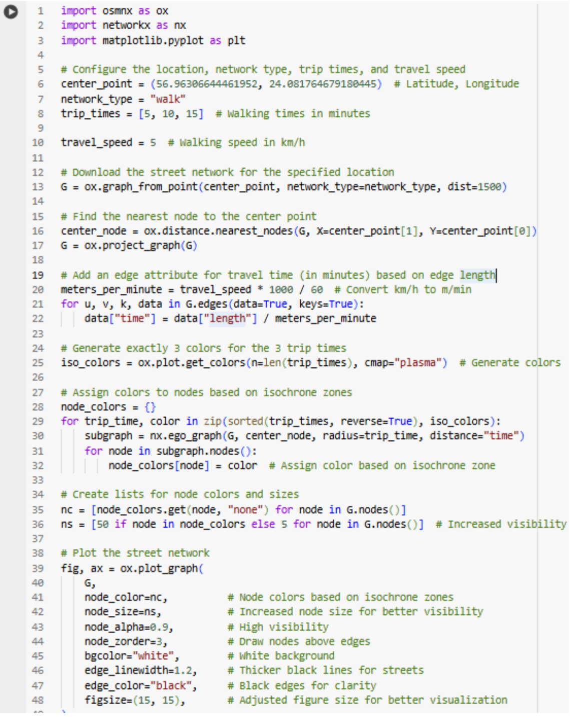


Figure 4. The designer as a glass box. [diagram]
Herbert, 1988



```python
import osmnx as ox
import networkx as nx
import matplotlib.pyplot as plt

# Configure the location, network type, trip times, and travel speed
center_point = (56.96306644461952, 24.081764679180445)  # Latitude, Longitude
network_type = "walk"
trip_times = [5, 10, 15]  # Walking times in minutes

travel_speed = 5  # Walking speed in km/h

# Download the street network for the specified location
G = ox.graph_from_point(center_point, network_type=network_type, dist=1500)

# Find the nearest node to the center point
center_node = ox.distance.nearest_nodes(G, X=center_point[1], Y=center_point[0])
G = ox.project_graph(G)

# Add an edge attribute for travel time (in minutes) based on edge length
meters_per_minute = travel_speed * 1000 / 60  # Convert km/h to m/min
for u, v, k, data in G.edges(data=True, keys=True):
    data["time"] = data["length"] / meters_per_minute

# Generate exactly 3 colors for the 3 trip times
iso_colors = ox.plot.get_colors(n=len(trip_times), cmap="plasma")  # Generate colors

# Assign colors to nodes based on isochrone zones
node_colors = {}
for trip_time, color in zip(sorted(trip_times, reverse=True), iso_colors):
    subgraph = nx.ego_graph(G, center_node, radius=trip_time, distance="time")
    for node in subgraph.nodes():
        node_colors[node] = color  # Assign color based on isochrone zone

# Create lists for node colors and sizes
nc = [node_colors.get(node, "none") for node in G.nodes()]
ns = [50 if node in node_colors else 5 for node in G.nodes()]  # Increased visibility

# Plot the street network
fig, ax = ox.plot_graph(
    G,
    node_color=nc,          # Node colors based on isochrone zones
    node_size=ns,           # Increased node size for better visibility
    node_alpha=0.9,         # High visibility
    node_zorder=3,          # Draw nodes above edges
    bgcolor="white",        # White background
    edge_linewidth=1.2,     # Thicker black lines for streets
    edge_color="black",     # Black edges for clarity
    figsize=(15, 15),       # Adjusted figure size for better visualization
```

Figure 5. Code for loading and analyzing the street network of a selected location using Python and the OSMnx and NetworkX libraries. The code determines the accessibility of zones based on specified walking times and prepares data for isochrone visualization. [Python code] Sergejs Kopils, 2025

The application of artificial intelligence (AI) in architectural design has further expanded the capabilities of data-driven methodologies. Zhang et al. (2016) highlighted how machine learning algorithms can simulate environmental performance, predict energy consumption, and optimize building layouts. AI-driven design tools such as generative adversarial networks (GANs) allow architects to explore multiple design alternatives based on predefined criteria, ensuring that both functional and aesthetic goals are achieved. As noted by Chaillou (2022), these tools facilitate the rapid iteration of design solutions, reducing the time required for manual adjustments and enhancing the precision of architectural decisions.

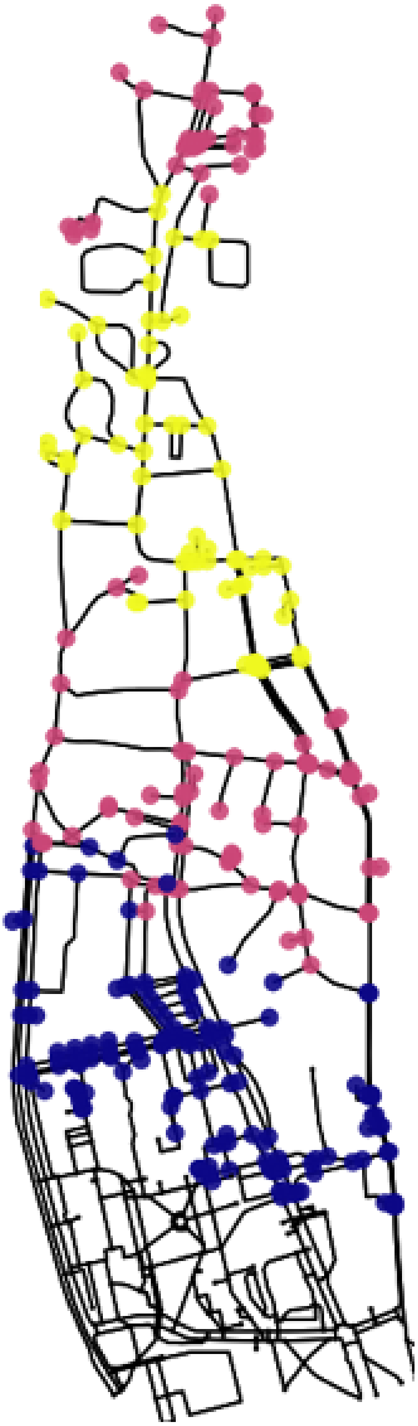


Figure 6. Map visualization of isochrones showing areas accessible within 5, 10, and 15 minutes of walking from the central point. Each zone is colored uniquely according to the time of accessibility. [diagram] Sergejs Kopils, 2025

Environmental sustainability is a key consideration in data-driven design, particularly in the development of underground data centers. Petrova et al. (2019) examined the environmental benefits of subterranean data centers, noting their natural thermal stability and reduced land-use footprint. The integration of real-time environmental data into design processes enables architects to implement passive cooling strategies, optimize ventilation systems, and minimize energy consumption. Velkova (2023) emphasized the potential for underground data centers to contribute to urban heat networks by repurposing excess server-generated heat for district heating.

The transition from traditional to data-driven design processes requires a multidisciplinary approach that incorporates insights from architecture, engineering, computer science, and environmental studies. As argued by Moud et al. (2020), collaboration between these disciplines enhances the effectiveness of data-driven design strategies, particularly in complex projects like underground data centers. The case studies presented in this research illustrate the practical applications of these methodologies, showcasing projects where data analytics, simulation tools, and AI technologies have been successfully integrated to achieve optimal performance and sustainability.

DATA CENTERS – ARCHITECTURAL AND TECHNOLOGICAL PERSPECTIVES

Data centers serve as critical infrastructure in the digital era, supporting cloud computing, big data analytics, and AI-driven applications. The architectural and technological aspects of these facilities are fundamental to their performance, efficiency, and sustainability. The rapid growth of data-driven technologies has intensified the demand for data centers that are both high-performing and environmentally responsible. According to Rashid (2019), data centers have evolved from mere storage facilities into complex systems requiring sophisticated design strategies that balance operational requirements with sustainability considerations.

The role of data centers in modern society is multifaceted, encompassing data storage, processing, and transmission for various digital services. Papaioannou et al. (2017) emphasizes the importance of designing data centers with scalable, adaptable architectures to accommodate future technological advancements. Izadi Moud et al. (2020) highlight the environmental impact of traditional cooling methods and advocate for the adoption of energy-efficient systems, such as liquid cooling and free cooling technologies.

Typologically, data centers can be categorized into enterprise, colocation, cloud, and edge facilities. Each type presents distinct design requirements and operational characteristics. Enterprise data centers are privately owned and operated by single organizations, while colocation facilities host equipment from multiple clients within shared infrastructure (Ajay, 2015). Cloud data centers support distributed computing resources accessed remotely, while edge data centers are strategically located near end-users to reduce latency and enhance performance (Batty, 2013).

The environmental challenges associated with data center operations have prompted architects to adopt innovative design strategies. Velkova (2023) underscores the importance of integrating passive design techniques, such as utilizing underground environments for natural thermal stability. Underground data centers, as studied by Petrova et al. (2019), benefit from consistent subterranean temperatures, reducing the need for mechanical cooling. However, these facilities also face challenges, including ventilation constraints and structural integrity concerns. Addressing these issues requires a multidisciplinary approach, combining architectural expertise with insights from environmental and data science fields.

The application of data-driven methodologies in data center design has proven effective in optimizing spatial layouts and operational performance. Building Information Modeling (BIM), as described by Wong and Fan (2013), facilitates the integration of diverse data streams, enabling architects to visualize, analyze, and

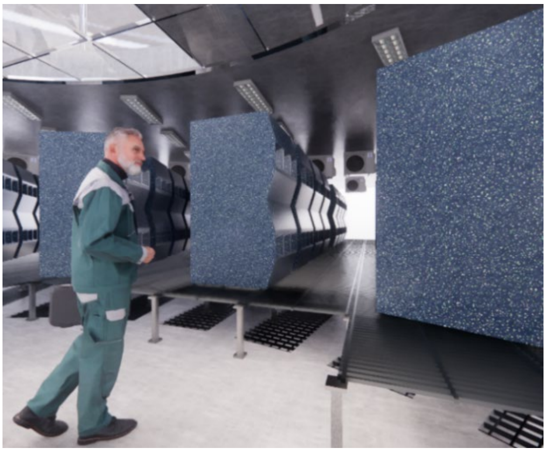


Figure 7. Air circulation within the server room [drawing] Sergejs Kopils, 2025

refine design alternatives. Computational Fluid Dynamics (CFD) simulations further enhance design precision by modeling airflow patterns and predicting thermal behavior within server rooms (Zhang et al., 2016) (Figure 7). AI-powered tools, including generative design algorithms, have been employed to explore innovative configurations that maximize energy efficiency and minimize spatial footprint (Chaillou, 2022).

Case studies illustrate the practical benefits of data-driven design in data center architecture. The Bahnhof Pionen White Mountain Data Center in Stockholm exemplifies the successful adaptation of underground infrastructure for digital applications (Hu, 2015). Housed within a former nuclear bunker, this facility utilizes rock walls as natural insulators, significantly reducing cooling energy requirements. Similarly, the Lefdal Mine Datacenter in Norway leverages cold seawater from nearby fjords to maintain optimal operating temperatures (Wu et al., 2012). These examples highlight the potential of site-specific design strategies that capitalize on local environmental conditions (Figure 8).

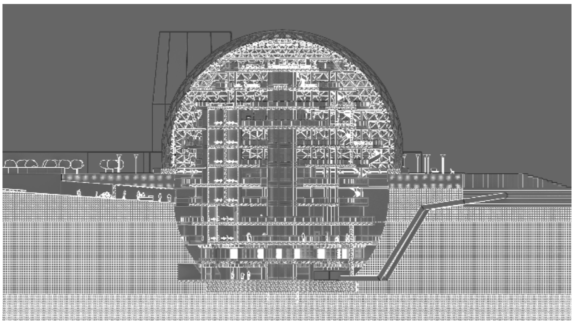


Figure 8. Cooling water sourced from the Daugava River, demonstrating the use of local natural resources for maintaining optimal data center temperatures. [drawing] Sergejs Kopils, 2025

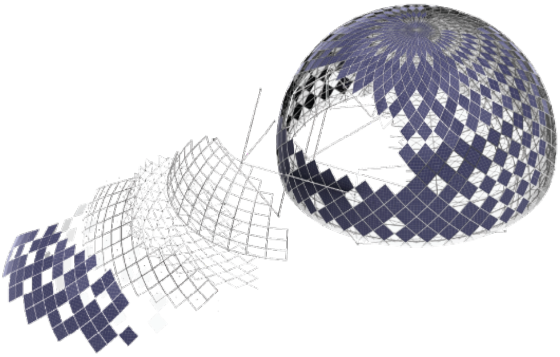


Figure 9. Diagram of solar panels on the façade, showcasing the integration of renewable energy sources to enhance data center sustainability

The future of data center architecture lies in the continued advancement of data-driven design principles. As AI algorithms become more sophisticated, they will enable predictive modeling of energy consumption patterns, equipment performance, and environmental impacts (Moud et al., 2020). Velkova (2023) suggests that digital twins—virtual representations of physical facilities—will become standard tools for real-time monitoring and optimization. Furthermore, the integration of renewable energy sources, such as solar and wind power, into data center designs aligns with broader sustainability goals and regulatory requirements (Figure 9).

METHODOLOGIES FOR DATA-DRIVEN DESIGN

Data-driven design methodologies represent a convergence of architectural theory and digital technologies, utilizing data inputs to inform and optimize design outcomes.

Identifying Relevant Data Inputs: Data-driven

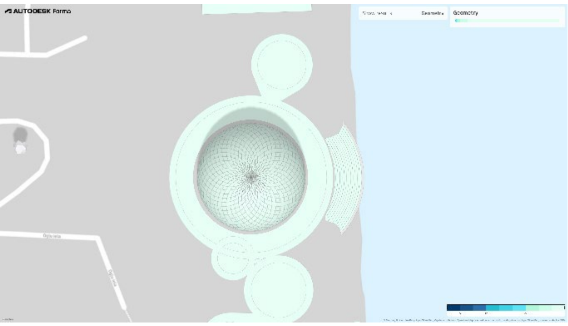


Figure 10. Microclimate analysis from Autodesk Forma, displaying air temperature distribution and thermal comfort indices for the site.

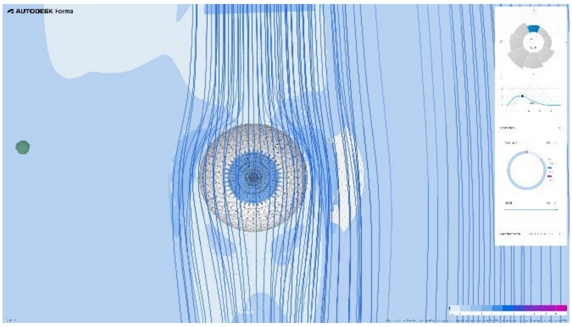


Figure 11. Wind flow analysis in Autodesk Forma, showing airflow patterns and their interaction with the building's geometry to identify natural ventilation opportunities.

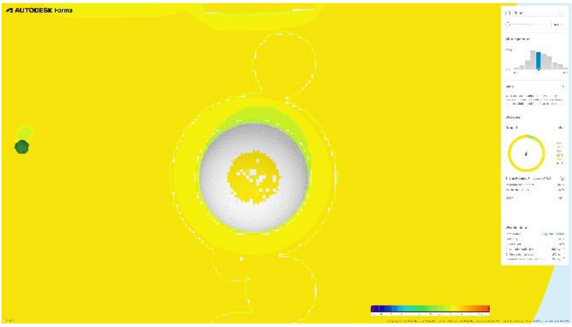


Figure 13. Sun hours analysis in Autodesk Forma, illustrating the total duration of sunlight exposure on the building's surfaces.

design hinges on gathering and analyzing three primary data categories: environmental, behavioral, and operational. Shehabi et al. (2016) emphasize that environmental data, such as temperature, humidity, and airflow patterns, are crucial for optimizing thermal performance in data centers (Figure 10, 11, 12 and 13). Behavioral data, including user interaction and occupancy patterns, guide spatial layout and energy efficiency strategies (Greenberg et al., 2009). Operational data, such as server load distribution and fault detection metrics, enable predictive maintenance and system optimization (Hogan, 2018).

AI-Driven Optimization and Simulation in Space Design: Artificial intelligence (AI) has transformed spatial design processes, enabling architects to explore vast design alternatives and refine layouts through simulations. According to As and Basu (2021), AI-powered generative design tools analyze multiple performance metrics to produce optimized spatial configurations. Petrova et al. (2019) highlight the application of Computational Fluid Dynamics (CFD) simulations to model airflow and thermal behavior, which is critical in reducing energy consumption. Additionally, machine learning models (Figure 14), as noted by Chaillou (2022), help architects predict energy use patterns and adapt designs accordingly.

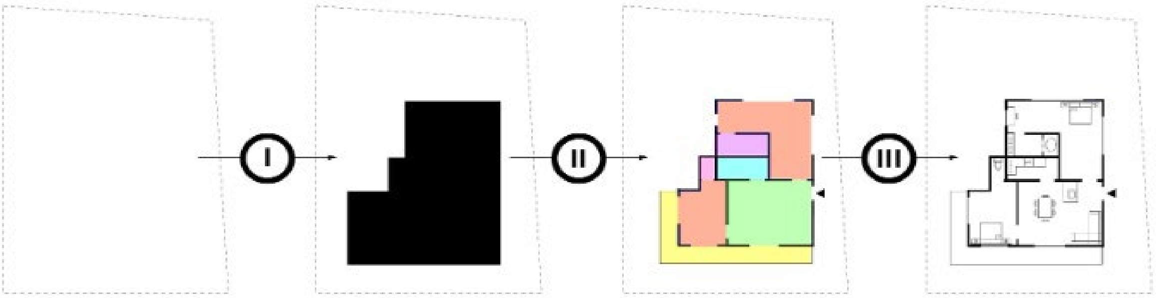


Figure 12. Generation Pipeline, Model I to III [diagram] Chaillou, 2022

Case Studies of Data-Driven Architectural Projects: The Google Hamina Data Center demonstrates the power of data-driven design by employing AI simulations and seawater cooling to optimize energy efficiency (Petrova et al., 2019). Similarly, Facebook’s Luleå Data Center utilizes predictive analytics for waste heat recovery, lowering operational costs (Hogan, 2018). In urban contexts, Helsinki’s underground data centers leverage subterranean environments for natural cooling, exemplifying site-responsive design (Velkova, 2023

Integration of Data Inputs for Comprehensive Design: Successful data-driven design integrates environmental, behavioral, and operational data into a cohesive framework. According to Izadi Moud et al. (2020), Building Information Modeling (BIM) platforms facilitate this integration by consolidating data into a shared model, enhancing collaboration among stakeholders. Additionally, digital twins, as described by Moud et al. (2020), allow for real-time monitoring and continuous optimization throughout a data center's lifecycle.

THE UNDERGROUND CLOUD DATA CENTER

Criteria for Evaluating Data-Driven Design Outcomes: Evaluating the effectiveness of data-driven design requires analyzing spatial efficiency, energy performance, and environmental impact. Petrova et al. (2019) emphasize that spatial efficiency is crucial in underground facilities, where excavation costs and space limitations dictate compact, optimized layouts. Hensel et al. (2022) advocate for using Building Information Modeling (BIM) and generative design algorithms to simulate and refine spatial configurations, ensuring efficient cable management and emergency accessibility.

Application of Data-Driven Methods in Conceptualization: AI-powered simulations and digital twin models are integral to the design of the Underground Cloud Data Center. Moud et al. (2020) highlight that predictive modeling, such as CFD simulations, optimizes airflow and cooling efficiency. Velkova (2023) notes the environmental benefits of utilizing subterranean thermal stability, reducing reliance on mechanical cooling systems.

Technological Infrastructure and Spatial Design Interaction: The integration of technology and architecture is key to the facility’s performance. Wong and Fan (2013) demonstrate that BIM integrates mechanical, electrical, and architectural systems into a unified model, streamlining design collaboration. Chaillou (2022) adds that generative design tools help balance server density with ventilation pathways, ensuring thermal stability without sacrificing spatial efficiency

Challenges and Limitations: Despite technological advancements, underground data centers face notable challenges. Petrova et al. (2019) cite ventilation and air circulation constraints (Figure 15), requiring sophisticated HVAC solutions. Additionally, Moud et al. (2020) note that limited space for vertical expansion restricts future scalability.

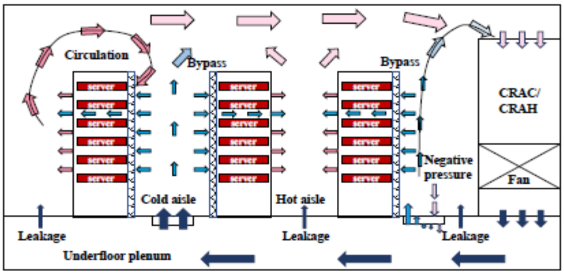


Figure 14. Diagram of room-level air cooling [diagram] Huang, 2023

CONCLUSION

This article has demonstrated how data-driven design methodologies address the complex challenges of underground cloud data center architecture. By integrating advanced computational tools, such as Building Information Modeling (BIM), AI-driven simulations, and digital twins, architects can optimize spatial layouts, reduce energy consumption, and enhance environmental sustainability. Analysis of environmental, behavioral, and operational data, reveals that data-driven design not only enhances performance but also promotes adaptive, sustainable architectural solutions. Case studies from industry leaders, such as Google's Hamina and Facebook's Luleå data centers, exemplify the practical benefits of these methodologies in achieving energy efficiency and operational resilience. Furthermore, the empirical study of the Underground Cloud Data Center highlights the effectiveness of applying these methodologies in subterranean environments, where thermal stability and compact spatial organization are critical. Despite inherent challenges, such as ventilation management and scalability constraints, the findings confirm that data-driven methodologies offer a robust framework for resolving spatial, environmental, and technological complexities in data center design.

PROPOSALS

To advance the application of data-driven design in underground cloud data center architecture, several targeted proposals emerge from this research:

Expand Integration of Digital Twin Technologies: Leverage digital twins for continuous real-time monitoring, predictive maintenance, and iterative performance optimization. This will enhance operational efficiency and provide valuable datasets for future design improvements.

Develop Advanced AI-Driven Simulation Tools: Implement generative design algorithms and AI-powered simulations to optimize spatial layouts, airflow management, and thermal performance. Emphasize the integration of machine learning to refine predictions and improve design adaptability.

Promote Interdisciplinary Collaboration: Establish cross-disciplinary frameworks that unite architects, engineers, data scientists, and sustainability experts to develop holistic design strategies that align technological performance with environmental goals.

Enhance Sustainable Design Practices: Prioritize energy-efficient strategies, such as waste heat reuse for district heating and the use of renewable energy sources. Employ predictive analytics to monitor and reduce carbon footprints throughout the data center lifecycle.

Advance Research on Subterranean Infrastructure: Conduct further studies on ventilation, structural integrity, and material innovation specific to underground data centers to address operational challenges and scalability limitations.

Incorporate Adaptive Design Frameworks: Utilize modular construction techniques and dynamic systems that allow for flexible adjustments in response to evolving technological requirements and increased data storage demands.

Ensure GDPR Compliance and Strengthen Data Privacy: Develop design strategies that integrate secure data management systems aligned with the General Data Protection Regulation (GDPR). Implement encrypted storage solutions and access controls to protect user data, ensuring that data privacy and security are prioritized throughout the design lifecycle.

Strengthen Ethical and Regulatory Compliance: Develop guidelines ensuring data privacy, security, and adherence to sustainable construction standards, aligning technological innovation with societal and environmental responsibilities.

REFERENCES

Ajay, A. (2015) *State Data Center: Top Ten Design Parameters, Towards Next Generation E-Government.*

Batty, M. (2013) 'Big data, smart cities and city planning', *Dialogues in Human Geography*, 3(3), pp. 274–279. Available at: https://doi.org/10.1177/2043820613513390.

Chaillou, S. (2022) *Artificial Intelligence and Architecture: From Research to Practice*. London: Thames & Hudson.

Greenberg, A., Hamilton, J., Maltz, D.A. and Patel, P. (2009) 'The cost of a cloud: research problems in datacenter networks', ACM SIGCOMM Computer Communication Review.

Hensel, M. et al. (2022) *Data-Driven Design for Architecture and Environment.* London: Wiley.

Herbert, S. (1988) *'The Science of Design: Creating the Artificial'*, The MIT Press [Preprint].

Hogan, M. (2018) 'Big Data Ecologies', Ephemera, 18(3), pp. 631–657.

Hu, T.H. (2015) *A Prehistory of the Cloud*. Cambridge, MA: The MIT Press.

Huang, X. et al. (2023) 'Cooling Technologies for Internet Data Center in China: Principle, Energy Efficiency, and Applications', *Energies*, 16(20), p. 7158. Available at: https://doi.org/10.3390/en16207158.

Moud, H.I. et al. (2020) 'Sustainability Assessment of Data Centers Beyond LEED', *in 2020 IEEE Green Technologies Conference (GreenTech)*. Oklahoma City, OK, USA: IEEE, pp. 62–64. Available at: https://doi.org/10.1109/GreenTech46478.2020.9289793.

Papaioannou, T., Nejabati, R. and Simeonidou, D. (2017) 'The Benefits of a Disaggregated Data Centre: A Resource Allocation Approach', *IEEE Global Communications Conference* (GLOBECOM 2016).

Petrova, E. et al. (2019) 'Towards data-driven sustainable design: decision support based on knowledge discovery in disparate building data', *Architectural Engineering and Design Management,* 15(5), pp. 334–356. Available at: https://doi.org/10.1080/17452007.2018.1530092.

Qabshoqa, M., Kocaturk, T. and Kiviniemi, A. (2017) 'A value-driven perspective to understand Data-driven futures in Architecture', in *eCAADe 2017: ShoCK!* – Sharing of Computable Knowledge! Rome, Italy, pp. 407–416. Available at: https://doi.org/10.52842/conf.ecaade.2017.2.407.

Rashid, M.A. (2019) *Data Center Architecture Overview.* National Academy for Planning and Development.

Shehabi, A. et al. (2016) *United States Data Center Energy Usage Report.* Lawrence Berkeley National Laboratory.

Velkova, J. (2023) 'Retrofitting and ruining: Bunkered data centers in and out of time', *New Media & Society,* 25(2), pp. 431–448. Available at: https://doi.org/10.1177/14614448221149946.

Wong, K. and Fan, Q. (2013) 'Building information modelling (BIM) for sustainable building design', *Facilities*, 31(3/4), pp. 138–157. Available at: https://doi.org/10.1108/02632771311299412.

Wu, K., Xiao, J. and Ni, L.M. (2012) 'Rethinking the architecture design of data center networks', *Frontiers of Computer Science* [Preprint]. Available at: https://doi.org/10.1007/s11704-012-1155-6.

Zhang, N. et al. (2016) 'ELM Meets Urban Big Data Analysis: Case Studies', *Computational Intelligence and Neuroscience*, 2016, pp. 1–10. Available at: https://doi.org/10.1155/2016/4970246.