Computer Vision and Image Understanding 260 (2025) 104454

Contents lists available at ScienceDirect

Computer Vision and Image Understanding

journal homepage: www.elsevier.com/locate/cviu

Made-In: An immersive human-in-the-loop analytics platform for enhancing creative processes in fashion

Emanuele Balloni [a,*], Rocco Pietrini [a], Michele Sasso [b], Emanuele Frontoni [c], Marina Paolanti [c]

[a] Università Politecnica delle Marche, VRAI - Vision Robotics and Artificial Intelligence Lab, Dipartimento di Ingegneria dell'Informazione (DII), 60131 Ancona, Italy

[b] Ubisive srl, 62012 Civitanova Marche (MC), Italy

[c] Department of Political Sciences, Communication and International Relations, University of Macerata, Via Don Minzoni 22/A, 62100 Macerata, Italy

ARTICLE INFO

Keywords:
Human-in-the-loop systems
Immersive analytics
Fashion design
Sustainability
Artificial Intelligence

ABSTRACT

The fashion industry is undergoing a digital transformation, driven by growing demands for sustainability, personalization and immersive experiences. In this paper, we present Made-In (Multimodal and Collaborative Artificial Intelligence for the Design of Inclusive and Sustainable Fashion): an immersive, human-in-the-loop analytics system designed to support fashion professionals in exploring, comparing and contextualizing product data across digital and social platforms. Unlike generative or simulation-based approaches, Made-In provides creative decision support by aggregating real-world data from luxury brand websites and social media. This enables designers and merchandisers to make informed, context-aware choices. The system comprises three core modules: a 3D configurator for visualizing product assortments; a collection grid interface for the comparative analysis of e-commerce data; and a social media trend detector based on deep learning pipelines for image classification, object detection and color clustering. Two curated datasets, one derived from Instagram and the other from fashion e-tailers, provide the system with analytics. A user study with domain experts confirms the platform's usability and relevance for trend forecasting, sustainability evaluation and visual merchandising strategy. The results demonstrate that Made-In effectively bridges the gap between data analytics and human creativity in fashion, offering a scalable solution that aligns with EU goals for digital sustainability and inclusivity.

1. Introduction

The European Union (EU) Strategy for Sustainable and Circular Textiles[1] is an important milestone in addressing the environmental and social impacts of the fashion industry. Fashion is an integral part of Europe's cultural heritage and expertise and is one of the most vibrant and creative sectors in Europe. Fashion is not just clothing; it is an expression of human personality, an art form and a philosophy of form that embraces well-being and inclusion (Olive, 2023; Githapradana et al., 2024). These industries form complex and interlinked value chains from the design and manufacture of fashion goods — such as textiles, clothing, footwear, leather, fur, jewelry and accessories — to their distribution and retailing (Karaosman et al., 2017; Karnad and Udiaver, 2022). However, the environmental impact of the fashion industry is profound and growing. EU average consumption of textiles has the fourth highest environmental and climate change impact after food, housing and mobility. It is also the third largest consumer of water and land, and the fifth largest consumer of primary raw materials and greenhouse gas emissions (Napolano et al., 2025). Fast fashion has a number of negative environmental impacts. The fashion industry is the second largest consumer of water, with current textile dyeing and treatment processes using approximately 93 billion cubic meters of water per year — enough to supply five million people. These processes also account for 20% of the world's wastewater (Niinimäki et al., 2020). As a result, more than 5.8 million tonnes of clothing are discarded each year in the EU, equivalent to around 11.3 kg per person (Wojciechowska, 2021). Alarmingly, only 1% of this discarded material is recycled into new clothing, while the remaining textiles are either dumped in European landfills or exported to Africa and Asia, according to the European Environment Agency (de Oliveira et al., 2021). In response to these challenges, the Circular Economy

* Corresponding author.
E-mail address: e.balloni@pm.univpm.it (E. Balloni).

[1] https://environment.ec.europa.eu/publications/textiles-strategy_en
[2] https://environment.ec.europa.eu/strategy/circular-economy-action-plan_en
[3] https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/european-green-deal_en

https://doi.org/10.1016/j.cviu.2025.104454
Received 16 May 2025; Received in revised form 14 July 2025; Accepted 23 July 2025
Available online 6 August 2025
1077-3142/

Action Plan[2] proposes new regulations for the fashion industry, including a ban on green washing, garment destruction, deforestation and other harmful practices. These regulations aim to overhaul the industry as part of the European Green Deal,[3] ensuring that fashion production and consumption becomes more sustainable and ethical. The fashion industry is made up of a wide range of participants, including designers, manufacturers, retailers, consumers and policymakers, each contributing to the overall system (Vehmas et al., 2018). Each stakeholder plays a different role in shaping the framework of the industry. Designers showcase their creativity through innovative and adaptable designs, while manufacturers turn these ideas into reality through skilled craftsmanship and technology (Lee et al., 2018). Retailers act as intermediaries between producers and consumers, curating and distributing fashion trends to a wide audience (Burnes and Towers, 2016; Khelladi et al., 2024). They are tasked with displaying and selling garments, providing consumers with access to the latest styles and trends. Consumers, in turn, wield considerable influence, driving demand and shaping market preferences through their behavior and purchasing decisions (Anjorin et al., 2024). Policymakers also play a crucial role, implementing policies and regulations to maintain industry standards, ensure compliance with health and safety laws, and address ethical concerns within the fashion sector (Karaman et al., 2016). Taken together, these diverse stakeholders create a dynamic ecosystem that fosters the continued evolution and creativity of the fashion industry, while also contributing to sustainability and circularity efforts to mitigate its environmental and social impacts.

One of the most promising developments supporting this transformation is the emergence of fashion products that are native to the digital world. These items are conceived and experienced entirely within virtual environments, only being produced physically upon confirmed demand. By decoupling design from immediate manufacturing, this approach significantly reduces overproduction, material waste and the carbon footprint associated with traditional retail. Digital-native fashion allows designers to experiment with innovative styles and evaluate market response at minimal ecological cost. It also offers consumers new ways to express themselves and customize products within digital spaces. These products are central not only to the growing integration of fashion and technology, but also to a scalable model that aligns creativity, sustainability and consumer engagement.

In this context, we introduce Made-In (Multimodal and Collaborative Artificial Intelligence for the Design of Inclusive and Sustainable Fashion), an AI-based system designed to promote sustainability and creativity in fashion through immersive, data-driven analytics. Made-In goes beyond providing a virtual shopping experience to serve as a real-time decision support system, integrating diverse data sources into an interactive environment for fashion analysis. It is an immersive analytics system designed to support the sustainable and inclusive design, merchandising and retail strategy of fashion through the integration of multimodal data sources, AI-driven processing and interactive visualization. It combines data scraped from social media platforms (Instagram) and e-commerce websites (e-tailers of luxury brands) to generate real-time, context-aware insights. Through its interactive interface, which includes collection comparison grids, geolocated trend maps and 3D product arrangement tools, Made-In enables users to explore the dynamic relationships between fashion products, consumer behavior and market trends. The system enhances decision-making for designers, merchandisers and brand strategists while fostering alignment with sustainability objectives by anticipating demand, optimizing assortments and monitoring competitor behavior across regions and seasons.

All system modules are supported by two curated datasets. The first is a collection of Instagram posts paired with fashion-related captions, which is used for trend and sentiment analysis. The second is a dataset extracted from luxury brand websites, containing structured product metadata such as price, availability and materials. Although the current implementation focuses on handbags, the system architecture can be adapted to other fashion product categories. A user study was conducted to evaluate the system's usability and effectiveness in supporting informed, responsible and personalized fashion exploration. The results demonstrate the value of embedding analytics within immersive digital environments, confirming Made-In's potential to transform online fashion experiences in line with the goals of the European Green Deal.

The main contributions of this work are as follows: (i) The development of an immersive, AI-based analytics system that enhances digital fashion exploration by integrating multimodal data analysis, sustainability metrics and social media-driven trend insights within an interactive environment; (ii) The construction of two curated datasets: the first consists of image-caption pairs scraped from Instagram using fashion-related hashtags to support trend detection and sentiment analysis, while the second is a structured dataset extracted from luxury brand websites containing detailed product metadata such as availability, price and materials; (iii) The implementation of an interactive interface supporting real-time product comparison, collection alignment, competitor monitoring and geolocated trend analysis, thereby facilitating data-informed decision-making for consumers and fashion professionals alike.

The paper is structured as follows: Section 2 reviews the relevant literature on AI applications in fashion, immersive analytics, and social media-driven trend analysis. Section 3 presents the architecture of the Made-In system, detailing its core components: the 3D configurator, collection grid module, and social media analysis pipeline. Section 4 reports on the experimental evaluation, including the user study design, key findings, and discussion of technical and ethical considerations. Finally, Section 5 concludes the paper and outlines directions for future work.

2. Related works

This section provides an overview of the key research areas that are relevant to AI in Fashion domain. These include developments in generative adversarial networks (GANs), text-based image synthesis and manipulation, and computer vision applications in fashion. For a more exhaustive review of these topics, readers are referred to recent survey papers such as Xia et al. (2022), Wu et al. (2017) and Cheng et al. (2021).

2.1. Generative adversarial networks for fashion

GANs have emerged as a leading approach for generating realistic images without the need for conditional inputs. They achieve high-quality image synthesis with relatively efficient training, even on standard hardware. The introduction of DCGAN (Radford et al., 2015) marked a significant shift, incorporating convolutional layers and offering architectural insights that promoted stable training. ProGAN (Karras et al., 2017) advanced the field by adopting a progressive training strategy that enables the generation of images at megapixel resolution. StyleGAN (Karras et al., 2019) introduced a novel generator architecture based on a nonlinear mapping of the latent space, drawing inspiration from style transfer methods. StyleGAN2 [39] refined this architecture by eliminating artefacts and incorporating path-length regularization to achieve more consistent training. StyleGAN2-ADA (Karras et al., 2020) incorporates data augmentation techniques to enhance performance on smaller datasets.

2.2. Text-guided image generation and editing

Models that generate or edit images based on natural language descriptions have become increasingly effective. These systems aim to produce images that visually match textual descriptions, or to make edits that align with new textual inputs while preserving unrelated content. Early work by Reed et al. (2016) introduced a GAN architecture in which text features were injected into both the generator and the discriminator. StackGAN (Zhang et al., 2017) and its successor, StackGAN++ (Zhang et al., 2018), adopted a multi-stage generation pipeline to incrementally improve image resolution. AttnGAN (Xu et al., 2018) integrated an attention mechanism to align specific words with their corresponding visual elements. MirrorGAN (Qiao et al., 2019) introduced a cyclic structure to enforce semantic consistency between text and generated images. More recent large-scale models, such as DALL· E (Ramesh et al., 2022), have leveraged massive datasets and transformer-based architectures to significantly outperform earlier systems in zero-shot settings. DALL· E 2 (Ramesh et al., 2022) further improved the quality of its outputs through diffusion-based generation, while Imagen (Saharia et al., 2022) demonstrated that larger text encoders enhance both image quality and semantic alignment.

Text-driven image manipulation techniques began with encoder–decoder GANs, as demonstrated by Dong et al. (2017). Nam et al. (2018) introduced localized word-level discriminators for more nuanced edits. ManiGAN (Li et al., 2020) contributed a refinement module to improve visual fidelity by better integrating image and text representations. Another class of methods involves converting an image into a latent representation (GAN inversion), manipulating the latent code, and decoding it back to produce the edited image. InterFaceGAN (Shen et al., 2020) identified interpretable directions in latent space that correspond to specific semantic attributes. Image2StyleGAN (Abdal et al., 2019), (Abdal et al., 2020) demonstrated facial image editing via inversion techniques. StyleCLIP (Patashnik et al., 2021) fused the text-to-image matching capabilities of CLIP (Radford et al., 2021) with the generative power of StyleGAN for text-guided editing. TediGAN applied style mixing mechanisms to guide text-based edits, while Baykal et al. (2023) enhanced editing flexibility with adapter modules and refinement layers within the latent space of StyleGAN2.

2.3. Computer vision for fashion applications

Within the fashion domain, computer vision has been widely used in Virtual Try-On (VTO) systems, which aim to render garments realistically on human figures while maintaining their original pose and appearance. Previous studies have addressed the 3D modeling of garments and human bodies (Zhang et al., 2023; Xiu et al., 2023; Cao et al., 2023; Song et al., 2023; Zhao et al., 2023; Jafarian et al., 2023; Jain et al., 2023; Zhu et al., 2023; Qiu et al., 2023; Zou et al., 2023; Grigorev et al., 2023; Wang et al., 2023; De Luigi et al., 2023) and image-based garment transfer (Han et al., 2018; Fele et al., 2022; Ge et al., 2021a,b; Yang et al., 2020; Zhu et al., 2017; He et al., 2022). Text-based modalities have also been adopted for fashion-oriented image generation and editing tasks. FashionGAN (Cui et al., 2018) uses a combination of text inputs, segmentation maps and categorical labels within an encoder–decoder GAN framework. The Fashion-Gen dataset (Rostamzadeh et al., 2018) provides a large-scale benchmark for evaluating text-to-image generation models in fashion. More recently, Text2Human (Jiang et al., 2022) was proposed for synthesizing human figures based on textual garment descriptions; however, it has limited textual flexibility due to a closed set of attribute categories. Other recent fashion-specific image manipulation models, such as Fang et al. (2023) and Zhang et al. (2022), explore the further integration of generative and linguistic capabilities.

While existing research primarily focuses on generative and text-conditioned fashion synthesis, virtual try-on and 3D modeling, our work takes a fundamentally different approach. Rather than generating or altering fashion images, Made-In provides an immersive analytics environment designed to support human creativity and sustainable decision-making by augmenting the exploration of real-world fashion data with AI. While previous studies have used GANs and diffusion models to simulate garments or manipulate visual appearance, our system combines scraped product metadata from luxury e-tailers with user-generated social media content to create interactive tools such as 3D configurators, comparative collection grids and geo-localized trend dashboards. Rather than generating images, these modules facilitate creative ideation, assortment planning and visual merchandising, all of which are based on real product offerings and consumer behavior. Furthermore, Made-In focuses on interpretability, user interaction, and integrating sustainability assessments. This bridges the gap between fashion AI research and practical design procedures.

3. System overview

This section presents the overall architecture of the Made-In system, illustrating how its components interact to support human-in-the-loop decision-making in the fashion domain. Made-In is a modular, immersive analytics platform which integrates various data sources, such as social media content, product metadata from luxury e-commerce platforms and reconstructed 3D assets, to enable informed, creative and sustainable fashion exploration. The system architecture comprises three core modules: (i) a 3D configurator for the spatial visualization of product assortments; (ii) collection grids for the comparative analysis of brand offerings; and (iii) a social media analysis pipeline for the detection and localization of trends. Each of these components is supported by a unified backend infrastructure that handles data acquisition, enrichment, classification and real-time visualization. The user-facing interface provides interactive dashboards and analytics environments designed to enhance the workflows of designers, merchandisers and decision makers. Fig. 1 illustrates the workflow of the Made-In system, showing the full pipeline from data collection to interactive exploration. Further details will be provided in the following subsections.

3.1. Virtual tour interface for immersive interaction

The virtual tour provides an immersive introduction to the Made-In system, allowing users to explore digital fashion environments that replicate the spatial and experiential qualities of a physical retail setting. Built using WebXR 3D technologies and rendered as a navigable scene, the tour interface simulates a high-end concept store. Users can interact with fashion products placed within a stylized environment. Each product in the tour has dynamic overlays that provide real-time access to structured metadata. When a user selects an item, the system retrieves information on the brand, material, availability, pricing and sustainability. Additionally, links to similar items based on either AI-generated style similarity or social media engagement are displayed in context. This interface transforms traditional browsing into an exploratory, data-rich experience that supports informed and intentional consumption. A custom photogrammetric model of a retail space was developed and integrated into a 3D web framework to build the virtual environment. Hotspot zones were defined to enable contextual product retrieval and are linked in real time to both the collection grid and the social analytics modules. Consequently, users can transition seamlessly from visual inspection to comparative analysis and social trend exploration. The tour is optimized for web use and is compatible with desktop, tablet and mobile devices to ensure broad accessibility. The interface is also designed to be modular and extensible so that future iterations can incorporate new product categories, spaces or personalization layers, such as user-based product arrangement or adaptive storytelling. Some examples can be found in Fig. 2.

3.2. 3D configurator for assortment visualization

The 3D configurator module provides an interactive, spatial interface for visualizing product assortments, supporting creative ideation and operational planning in visual merchandising. Unlike traditional 2D dashboards or static catalogue views, the configurator allows users, such as designers, buyers and merchandisers, to create and assess curated selections of fashion products in an immersive, customizable virtual environment. To achieve photorealistic, real-time rendering

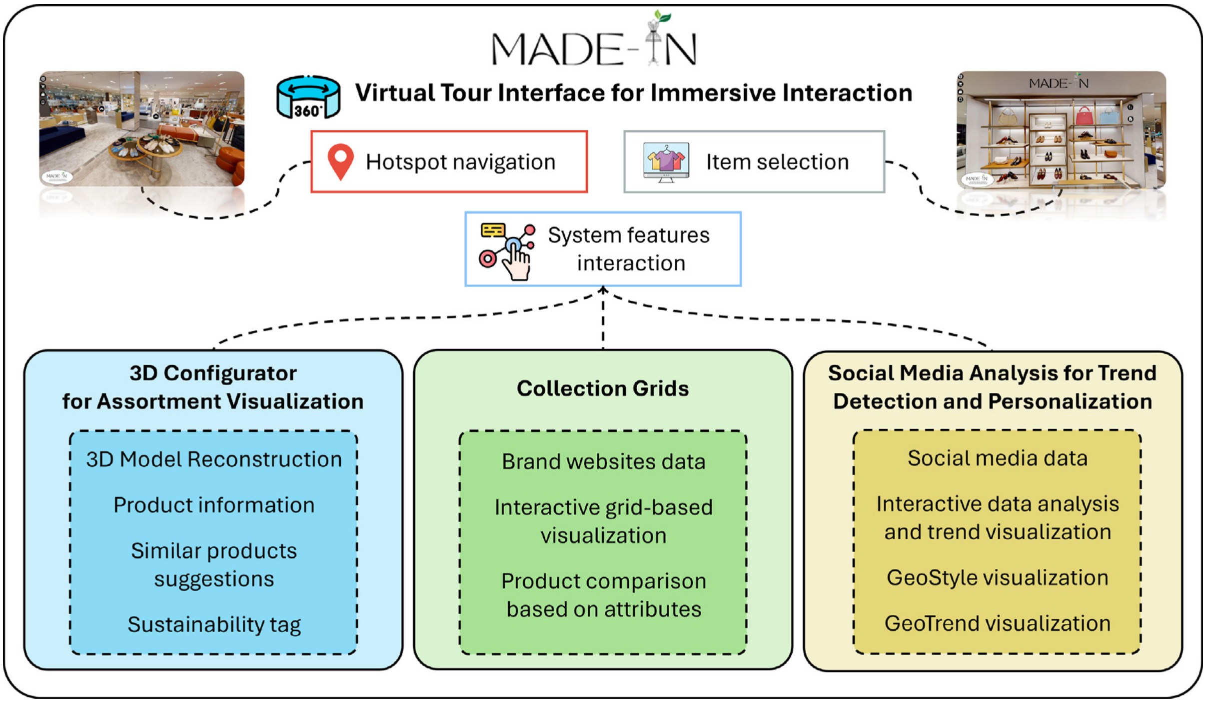


Fig. 1. Overview of the Made-In system.

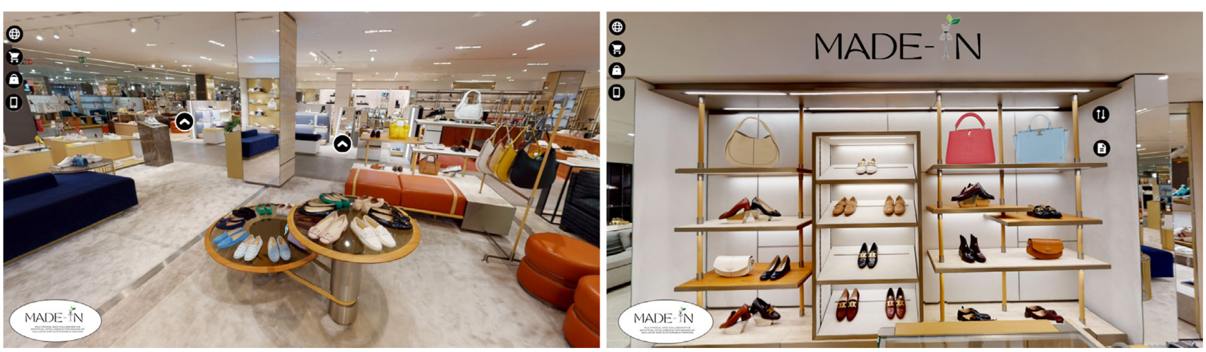


Fig. 2. Example views from the Made-In virtual tour environment. The image on the left illustrates a navigable retail space designed for immersive product exploration, and the image on the right displays a structured product wall designed for interactive browsing and fashion item comparison.

with future compatibility across AR/VR and XR platforms, we used 3D Gaussian Splatting (3DGS) as the neural rendering approach for asset reconstruction, as described in Kerbl et al. (2023). Compared to volumetric NeRF-like methods, 3DGS offers faster training and rendering times while maintaining high fidelity, making it suitable for interactive settings where real-time feedback and smooth scene navigation are essential. This has already been demonstrated in a previous study in the fashion domain Balloni et al. (2025). For the design of the Made-in 3D configurator, we created a process to convert high-resolution product images captured in a controlled showroom environment into 3D point-based representations that are suitable for 3DGS rendering. These images include multiple views of the same item, captured under consistent lighting and background conditions, to facilitate high-quality multi-view reconstruction. Camera poses were estimated using COLMAP (Schönberger and Frahm, 2016), and the resulting data were formatted into the Gaussian Splatting representation. This allows for the efficient neural rendering of each item in its spatial context. The configurator allows users to arrange reconstructed products interactively on virtual shelves, plinths or display zones. Products can be rotated, scaled and positioned within a showroom-like scene to simulate potential in-store or digital layouts. Visual exploration can be carried out using a keyboard and mouse, or optionally via VR controllers in compatible setups. In addition to visual positioning, the system provides metadata on assortments, including brand, material, dominant color, price and sustainability indicators (derived from our collection grid pipeline). This enables users to select items based not just on visual appeal, but also on strategic criteria such as pricing tiers or eco-design objectives. Users can also save and compare different assortment configurations for iterative refinement or team collaboration.

Thus, the 3D configurator serves as both a visualization layer and a decision support interface, enabling fashion professionals to conceptualize, evaluate and optimize product assortments in a spatially and semantically rich digital environment. Its modular architecture enables the seamless integration of new 3D assets and scene templates, facilitating scalability across different fashion categories or seasonal campaigns. Fig. 3 illustrates an example scene from the 3D configurator, showcasing a reconstructed handbag displayed in a virtual merchandising environment. As can be seen, the interface supports both free navigation and metadata overlays, enabling intuitive integration of spatial product visualization with data-driven decision support.

3.2.1. Dataset collection

We began by acquiring a multi-view image dataset for each fashion item using a controlled photo-capture setup. Objects are placed

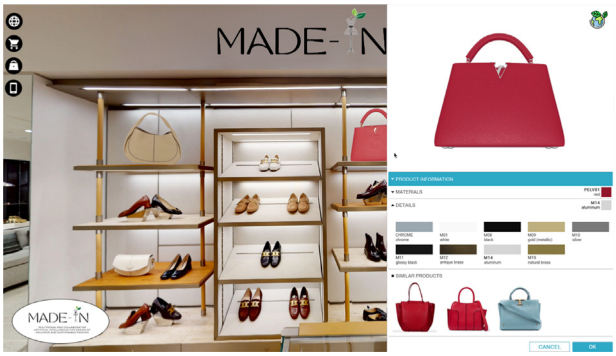


Fig. 3. 3D Configurator Interface of the Made-In system. It shows the 3D model of the selected product, along with information, sustainability flag, details and similar products.

on a motorized turntable with consistent illumination, and a set of 250 high-resolution (6720 × 4480 pixels) images are taken at fixed intervals to provide a full 360° visual profile. To ensure geometric consistency across views, physical calibration markers are used when camera calibration is not pre-specified. After the image acquisition, Structure-from-Motion (SfM) and Multi-View Stereo (MVS) are used to recover camera extrinsics and generate an initial sparse point cloud. For this step, we employ COLMAP,[4] an open-source SfM solution that supports robust feature matching and bundle adjustment, yielding registered camera poses and sparse 3D structure necessary for subsequent reconstruction. The result is a set of high-resolution images with their corresponding camera parameters and the sparse point cloud, which is used as the initialization step for the 3DGS training.

3.2.2. 3DGS training

The core of the 3D generation pipeline involves reconstructing a photo-realistic 3D representation from the image set using 3DGS. We chose 3DGS instead of other approaches, such as Neural Radiance Fields (NeRFs) (Mildenhall et al., 2020), as 3DGS offers lighter and faster real-time rendering and significantly reduced memory usage compared to NeRFs, key features for fashion-focused interactive systems deployed in web browsers.

Starting from the initial point cloud, the reconstruction is performed by leveraging the Splatfacto method from the Nerfstudio framework,[5] which represents each point as an anisotropic Gaussian primitive with spatial, opacity, and radiance properties. The optimization process minimizes the photometric error across all views by adjusting these parameters. Each Gaussian encodes a portion of the surface geometry and appearance, and their aggregation forms a coherent, high-fidelity model of the fashion item. After training, the resulting scene is rendered using a tile-based rasterization pipeline that blends splats based on depth ordering and per-point opacity. Unlike traditional mesh-based approaches, 3DGS models are directly renderable without conversion, making them suitable for a web-based system.

3.2.3. Integration in WebXR platform

The resulting 3DGS model is embedded into the Made-In immersive dashboard via a WebXR-compatible renderer built using Three.js, based on GaussianSplats3D.[6] As 3DGS does not rely on complex ray-marching or neural networks at runtime, the rendering engine achieves high frame rates on desktop, tablet, and mobile devices. The WebXR platform enables real-time manipulation of products within the virtual showroom scene. Users can move and rotate items and apply filters based on metadata (e.g., brand, material). The interface is designed to be modular and extensible, allowing the easy addition of new product categories, environments, or user-defined arrangements. This ensures that the configurator not only supports presentational tasks but also strategic assortment planning and trend-sensitive collection design.

3.3. Collection grids

The Collection Grids module of the Made-In system offers a dynamic, interactive interface for comparative fashion analysis of internal collections and external market offerings. Designed as an AI-augmented human-in-the-loop solution, it enables fashion professionals to explore large amounts of product data in real time by integrating automated crawling, structured scraping and intelligent data enrichment. Its purpose is to facilitate informed decision-making by offering a data-driven evaluation of assortments, competitor positioning, and emerging product trends.

3.3.1. Data collection and target configuration

The data pipeline starts with a configurable interface for defining collection targets. Each target corresponds to a fashion brand or online retailer. For each source, users can specify crawling parameters, such as:

- Crawl Limit: the maximum number of products retrieved per session;
- Concurrency: the number of pages processed in parallel.
- Cycle sleep: the delay between cycles to respect website rate limits.

After initialization, catalog pages are manually registered by users to prioritize relevant product categories, seasonal collections or brand-specific lines. Crawling then retrieves product identifiers, links and basic metadata in a structured format for downstream enrichment. The metadata are stored in JSON format, which offers a lightweight and human-readable structure. This enables seamless integration a our downstream pipeline.

3.3.2. Web scraping and metadata normalization

Once the targets are defined, the scraping module extracts standardized product metadata:

- Product titles, SKUs, and brand names,
- Descriptions and taxonomic categories,
- Prices and currencies,
- Product images and links.

The scraping procedure is performed with Python 3.12, using Selenium WebDriver[7] and the BeautifulSoup Python library. The raw data is normalized across sources and cleaned using string-matching and rule-based preprocessing. This ensures compatibility across fashion websites with heterogeneous HTML templates and dynamic content loading. [To enhance scalability and long-term adaptability, we implemented a modular target configuration layer that supports automatic detection of content structure changes in e-commerce templates. The scraper includes a fallback mechanism based on a hybrid approach combining DOM traversal with shallow-learning XPath pattern matching. This allows the system to re-learn key element paths (e.g., price, title, image, material) when minor layout or structure changes occur, without requiring full reconfiguration.]

4 https://github.com/colmap/colmap
5 https://github.com/nerfstudio-project/nerfstudio
6 https://github.com/mkkellogg/GaussianSplats3D
7 https://www.selenium.dev/documentation/webdriver/

Fig. 4. Example collection grids visualizations in the Made-In system.

3.3.3. AI-powered metadata enrichment

Dedicated AI modules are used to produce enriched analytics, each tailored to a key product attribute.

- Dominant Color Detection: Product images are processed using IS-Net (Qin et al., 2022) for background segmentation. The foreground pixels are then clustered using k-means and the dominant colors are identified based on the CIEDE2000 color difference metric (Sharma et al., 2005). Items are then grouped into color categories for comparative analysis.
- Material extraction: Product descriptions are processed using a hybrid approach combining fine-tuned DistilBERT models (Sanh et al., 2019) and domain-specific material taxonomies. The identified materials are then mapped to a controlled vocabulary to enable consistent grouping and filtering.
- Sustainability flagging: Product texts are analyzed using prompt-based LLM calls (Llama 4 (Touvron et al., 2023) in our case), which evaluate sustainability-related information such as recycled content, low-impact processes or certifications. The result is a binary flag or sustainability score that can be used for trend analysis and segmentation.

3.3.4. Interactive grid-based visualization

The enriched dataset is visualized through interactive collection grids, which allow users to explore and compare products across multiple dimensions (Fig. 4). Rows and columns of the grid can be configured using dropdown selectors to visualize attributes such as:

- Brand or source,
- Product category or sub-type,
- Dominant color,
- Material,
- Price tier,
- AI-generated stylistic clusters,
- Sustainability indicators.

Users can apply filters to isolate subsets of interest (e.g., sustainable women's handbags priced above €800, or outerwear from selected brands). The system supports multi-dimensional sorting and real-time visual overlays for highlighting selected features. When hovering over a product tile, a detail view is displayed showing scraped and AI-inferred attributes, including sustainability indicators and material composition. In addition, users can define and save custom presets that specify a time window, brand mix, price range, and product type. These presets enable longitudinal trend monitoring, allowing teams to track how product characteristics evolve across fashion cycles, events (e.g., fashion weeks), or seasonal launches. The Collection Grid module thus serves as the analytical core of Made-In, linking raw product data to strategic, evidence-based decision-making in visual merchandising and assortment design.

3.4. Social media analysis for trend detection and personalization

As done in Balloni et al. (2024), the Social Media Analysis module in Made-In is designed to extract, filter and analyze fashion-related user-generated content from Instagram. This allows emerging trends to be detected and dominant visual attributes to be captured. This supports location-aware personalization. It offers a data-driven view of consumer sentiment and style evolution, helping designers and retailers to adapt their product offerings quickly and in line with changing preferences.

3.4.1. Data collection and preprocessing

We implemented a custom data collection pipeline using the Instaloader[8] library to scrape Instagram posts by crawling predefined, fashion-related hashtags (e.g., shoulderbag, clutch, streetstyle). Each post is stored alongside its image, caption, timestamp, hashtags and geolocation (if available). The pipeline is scheduled to run twice daily to ensure the data remains up to date and aligned with the rapidly evolving fashion discourse. To eliminate irrelevant content, we trained a fashion vs. non-fashion image classifier using the VGG-16 convolutional neural network architecture (Simonyan and Zisserman, 2014). This classifier was then fine-tuned using a balanced, custom-curated dataset of 700,000 manually labeled Instagram images for binary classification. Images that pass this filter are forwarded for further visual and textual analysis.

To increase robustness and automation, the social media acquisition process includes a multi-account rotating authentication mechanism to prevent scraping limits and access throttling. Furthermore, to ensure long-term scalability and compliance, the system is structured to allow easy accommodation of additional social sources beyond Instagram (e.g., Pinterest, TikTok) with minimal integration overhead.

Additionally, a rigorous data anonymization protocol is applied to guarantee that only aggregated and non-personally identifiable information is stored, thereby upholding ethical research standards (Giovanola et al., 2023). This practice reflects a firm commitment to protecting individual privacy and dignity, ensuring that all data handling processes are aligned with principles of ethical responsibility, transparency, and respect for participants' rights, as outlined by institutional review boards and international data protection frameworks (such as GDPR).

3.4.2. Object detection and attribute extraction

Filtered fashion images are processed using YOLOv5, which has been fine-tuned using the "MADAME" (iMage fAshion Dataset sociAl MEdia) dataset (Della Sciucca et al., 2022) to detect and classify handbags. The detected items are then cropped and sent to a K-means clustering module to extract the three most dominant colors. Clustering is optimized using the Elkan variant of K-means and the CIEDE2000 color distance metric to ensure accurate color grouping. The handbag classes detected include backpacks, totes, shoulder bags, clutches and other major subtypes, enabling the system to quantify style distribution and map category frequencies over time and space.

[8] https://github.com/instaloader/instaloader

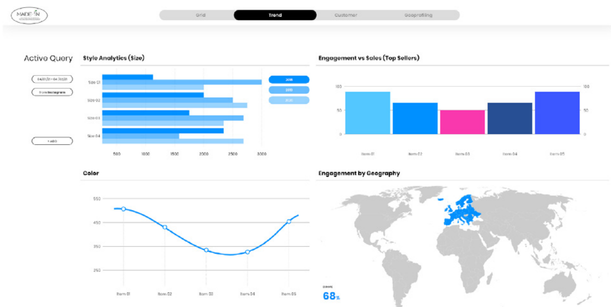


Fig. 5. Analytics and trends visualization from social media in the Made-In system.

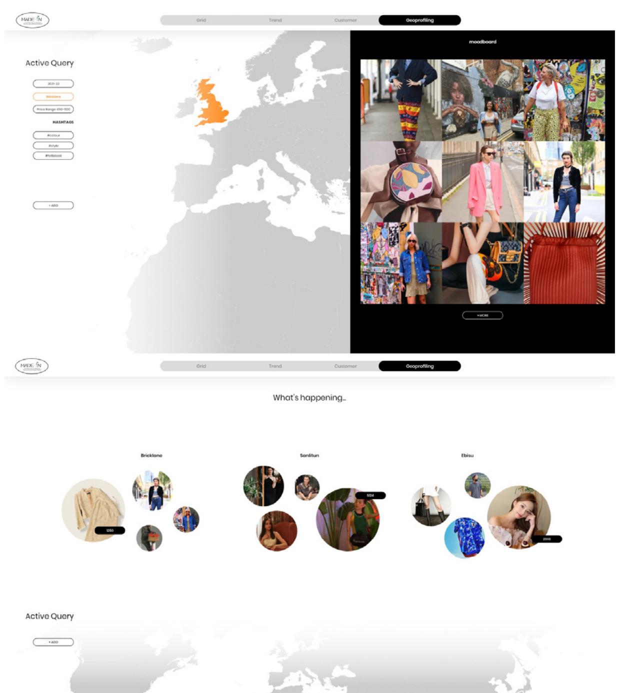


Fig. 6. GeoStyle and GeoTrend visualizations.

3.4.3. *Trend analysis and geospatial modeling*

Caption text is processed using a combination of DistilBERT and Named Entity Recognition (NER) techniques to extract information such as materials, brand mentions and sentiment. Co-occurrence networks and cluster detection algorithms are then used to reveal stylistic patterns, which are mapped using geolocation metadata. We generate real-time geospatial heatmaps using Leaflet.js that visualize where particular colors, materials, or product types are trending. These insights are fed into the GeoStyle module of Made-In, enabling location-specific assortment strategies and supporting campaign planning for targeted markets. A trend analysis dashboard and GeoStyle visualizations are shown in Figs. 5 and 6.

4. Experiments and results

This section presents the outcomes of the empirical evaluation of the Made-In system, covering the usability study design, core performance metrics, and the implications of the findings from both a technical and societal perspective.

4.1. *Participants*

The evaluation group consisted of 32 participants who were selected to represent the various stakeholder groups within the fashion industry. Inclusion criteria included at least two years of professional experience in a relevant role. Considering our within-subject design, each participant equally participated in all of the experiments (Chheang et al., 2024). Each participant completed an initial form to give informed consent and provide some demographic details. The sample exhibited a balanced demographic profile, with a mean age of 42 years (SD = 4.2) and a gender distribution (66% female, 34% male) that reflected the current composition of the industry. The professional roles encompassed design (n = 12), merchandising (n = 8), trend analysis (n = 7), and academic research (n = 5), thereby ensuring comprehensive coverage of potential use cases. An analysis of educational attainment reveals that 56% of the sample possess Master's degrees in fashion-related disciplines, while 31% hold doctoral qualifications. The range of professional experience varied significantly, with a majority ranging from 2 to 15 years. Following this, participants completed a pre-questionnaire, assessing their familiarity with virtual tours, 3D models, analytics tools, generative AI and LLMs. It is noteworthy that 60% of the respondents reported utilizing digital design tools on a regular basis. While 75% exhibited familiarity with generative AI tools in general (like ChatGPT or MidJourney), only 28% possessed specific experience with fashion applications. Approximately 59% reported experience with 3D modeling tools, though with varying proficiency levels (M = 3.1/5, SD = 1.2), while only 41% had used virtual tour platforms. Familiarity with analytics tools was reported by 44% of participants, but specialized fashion applications were less common (28%).

4.2. *Experimental process*

The experimental process for the Made-In system is shown in Fig. 7. It is designed to encompass both quantitative performance metrics and qualitative user experience data through a multi-phase experimental design. The study employed a within-subjects approach that progressed from controlled, task-based interactions to open-ended exploration. Prior to participation, all subjects underwent an informed consent procedure. This procedure transparently delineated the objectives of the study, the protocols for data management, and the privacy safeguards in place. Subsequent to this, a comprehensive demographic survey was conducted, encompassing professional backgrounds, including years of experience in fashion-related roles, educational qualifications, and previous exposure to digital design technologies, virtual tours and AI.

The core evaluation unfolded through two distinct navigation phases: *Guided Navigation* and *Free navigation*. The initial guided navigation session was conducted in accordance with a protocol design to systematically introduce the system's functionalities in a logical sequence. The participants initiated the training by familiarizing themselves with the fundamental *Virtual Tour Interface*, acquiring proficiency in navigational controls and hotspot interactions through the execution of standardized product selection tasks. This preliminary training established the foundation for subsequent exploration of the *3D Configurator*, where participants engaged in hands-on practice of model manipulation techniques while examining product information displays. This practice entailed a focus on interpreting sustainability certifications and material composition data. Subsequently, the guided session transitioned to more advanced analytical tasks. Participants systematically employed the *Collection Grids* to execute brand comparisons across multiple attribute dimensions, followed by structured exploration of the *Social Media Analysis* platform, where participants were able to view trends based on different parameters (e.g., region, brand, date). Each participant completed the same task set in the same order. The average completion time for the Guided Navigation phase was 18.2 min (SD = 3.6). Subsequent to this guided orientation, participants completed an intermediate assessment questionnaire that captured their impressions of each module's usability, the intuitiveness of various features, and the clarity of information presentation. Afterwards, the *Free Navigation* session began. In this session, participants could explore

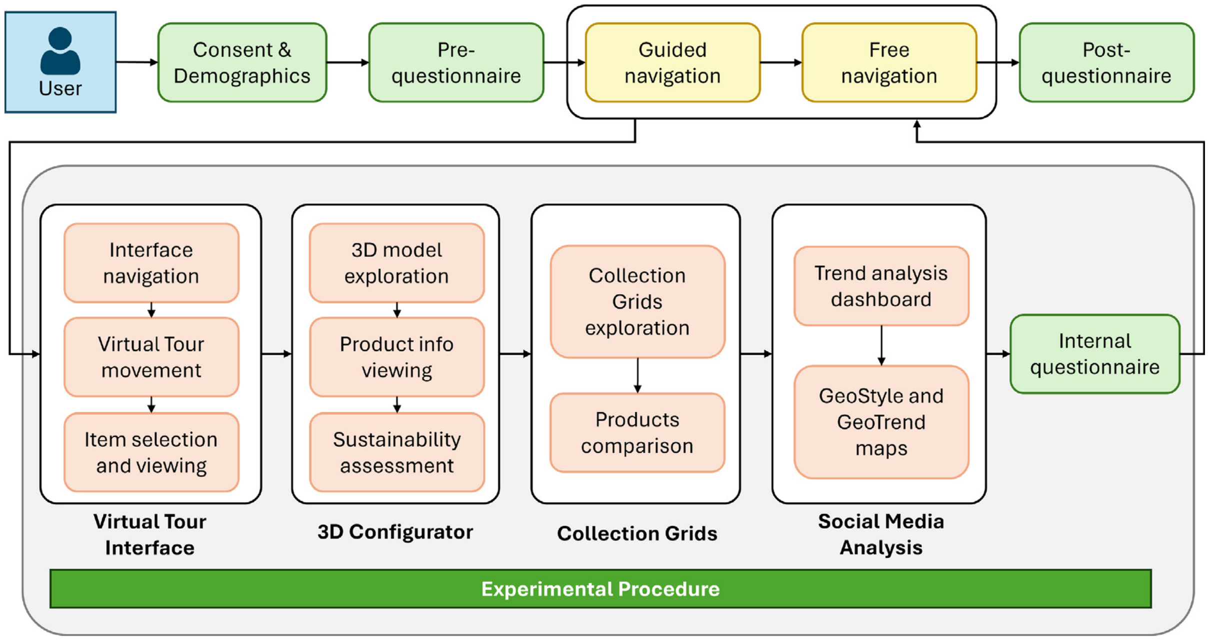


Fig. 7. Experimental process of the user study performed.

and complete the tasks as they pleased. During this phase, participants independently explored the virtual environment, selecting products of personal interest and applying analytical tools according to their own professional curiosity. The average duration of the Free Navigation phase was 24.7 min (SD = 4.1), reflecting both higher engagement and personalized use of system features. A think-aloud protocol was implemented during this session, providing valuable insights regarding real-world usage patterns and spontaneous feature discovery. The study concluded with a comprehensive post-questionnaire that measured several key dimensions of system evaluation, including usability, cognitive load, and adoption willingness of the Made-In system.

4.3. Evaluation metrics

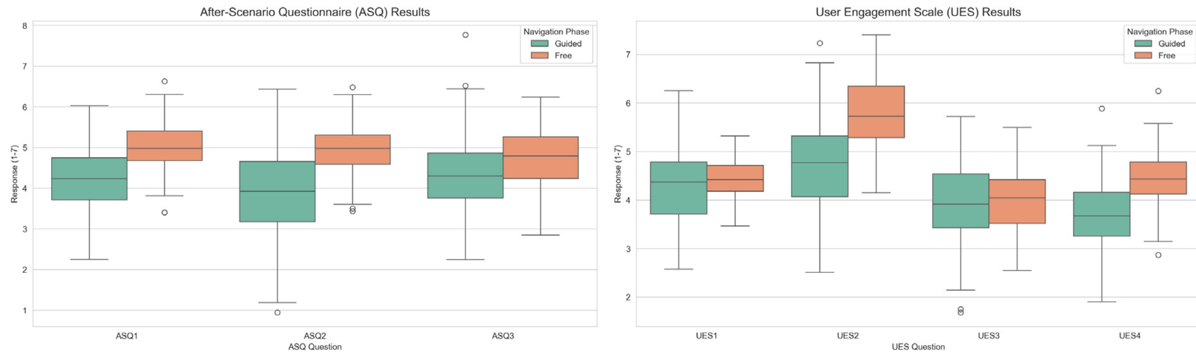
The evaluation of the Made-In system employed a multi-phase assessment methodology designed to capture both immediate user perceptions during interaction and comprehensive post-use evaluations. This approach included the integration of intermediate questionnaires administered after each major interaction phase with post-experiment questionnaires, thereby ensuring a comprehensive understanding of the user experience across different stages of system engagement. Subsequent to both the Guided Navigation and Free Navigation phases, participants completed two standardized intermediate questionnaires. The After-Scenario Questionnaire (ASQ) (Lewis, 1991) served as a focused assessment of task-specific usability, measuring participants' perceptions of task difficulty, time efficiency, and adequacy of support through its three core items (Melo et al., 2022). The ASQ was complemented by the User Engagement Scale (Short Form) (UES-SF) (O'Brien and Toms, 2010; O'Brien et al., 2018), which was administered after each navigation phase to evaluate deeper aspects of the interactive experience. The UES's comprehensive assessment across 4 dimensions, namely focused attention, perceived usability, aesthetic appeal, and reward, enabled tracking of engagement dynamics as users progressed from guided tasks to self-directed exploration (Wang et al., 2024; Vazquez et al., 2021). The post-experiment evaluation employed three well-established instruments to assess overall system perception (Longo, 2017): the System Usability Scale (SUS) (Brooke et al., 1996) provides a reliable global measure of interface quality and system usability; the Technology Acceptance Model (TAM) (Davis et al., 1989) questionnaire examines key factors influencing adoption potential; the NASA Task Load Index (NASA-TLX) (Hart and Staveland, 1988) provides detailed insights into cognitive demands.

All the statistical analysis were performed using Python (v3.8) and the packages Matplotlib (v3.8.4), Pandas (v2.2.1), Pingouin (v0.5.5), Seaborn (v0.13.2), Scipy (v1.12.0), and Numpy (v1.26.4).

4.4. Results

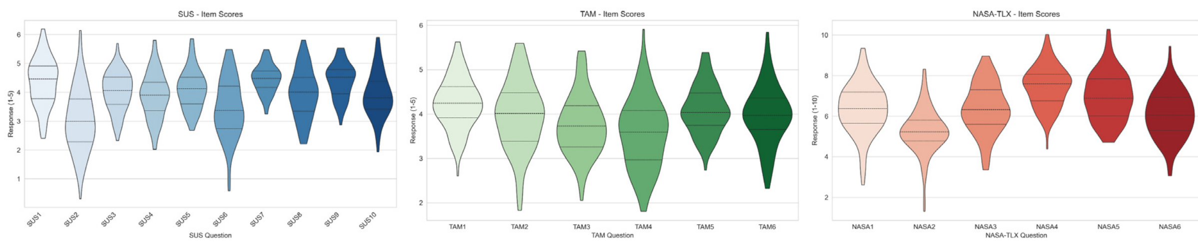
The evaluation of the Made-In system yielded comprehensive insights into user experience and system performance across all administered questionnaires. The results are presented in three main sections: intermediate questionnaire results from the Guided and Free Navigation phases, post-experiment questionnaire outcomes, and comparative analysis across interaction modes.

Intermediate questionnaire results. Results are visually reported in Fig. 8. Analysis of the ASQ (7-point Likert scale) revealed significant improvements between the Guided and Free Navigation phases. During Guided Navigation, participants reported moderate satisfaction with task completion (M = 4.2, SD = 0.8), with particular challenges noted in time efficiency (M = 3.8, SD = 1.1). These scores improved markedly during Free Navigation (task completion: M = 5.1, SD = 0.6; time efficiency: M = 4.9, SD = 0.7), indicating that the structured introduction effectively prepared users for independent exploration. The adequacy of support scores remained consistently high across both phases (Guided: M = 4.5, SD = 0.9; Free: M = 4.7, SD = 0.8), suggesting the interface provided appropriate guidance throughout the interaction. The UES results (5-point Likert scale) demonstrated strong engagement across all subscales. Focused attention scores were consistently high (Guided: M = 4.3, SD = 0.6; Free: M = 4.5, SD = 0.4), confirming the system's ability to maintain user concentration during analytical tasks. The reward subscale showed significant improvement ($p<0.05$) from Guided (M = 3.8, SD = 0.7) to Free Navigation (M = 4.4, SD = 0.5), indicating users derived greater value from the system as they gained proficiency. Novelty scores were highest during Guided Navigation (M = 4.6, SD = 0.5), reflecting the initial appeal of the system's innovative features, while endurability scores peaked during Free Navigation (M = 4.2, SD = 0.6), suggesting sustained engagement over time.



(a) ASQ Responses (b) UES Responses

Fig. 8. Boxplot scores for ASQ and UES.



(a) SUS Responses (b) TAM Responses (c) NASA-TLX Responses

Fig. 9. Violinplot scores for SUS, TAM and NASA-TLX.

Post-experiment questionnaire results. Post-Experiment Questionnaire Results are reported in Fig. 9. The SUS (5-point Likert scale) yielded a comprehensive score of 78.4 (SD = 9.2). High scores were observed for learnability items (M = 4.3, SD = 0.8), though some participants noted challenges with advanced filtering options in the collection grid module (M = 3.2, SD = 1.1). TAM results (5-point Likert scale) indicated strong adoption potential, with perceived usefulness scoring highest (M = 4.2, SD = 0.6 on a 5-point scale). Perceived ease of use showed greater variability (M = 3.7, SD = 0.7), correlating significantly with participants' prior experience with analytical tools (r = 0.42, $p < 0.01$). Behavioral intention scores were uniformly high (M = 4.1, SD = 0.5), suggesting strong willingness to incorporate the system into professional workflows. NASA-TLX results (10-point Likert scale) revealed moderate overall cognitive load (M = 6.3, SD = 1.1 on a 10-point scale). Mental demand was the most prominent factor (M = 6.5, SD = 1.2), particularly during complex tasks like cross-collection comparisons. However, the high performance subscale score (M = 7.3, SD = 1.0) indicated users felt effective when using the system, suggesting the cognitive load was appropriately aligned with task value.

Comparative analysis and discussion. Professional roles significantly influenced results across all measures. Merchandisers reported higher usability (SUS +7.3 points) and lower cognitive load (NASA-TLX −0.8 points) compared to designers, reflecting the system's strengths in assortment planning and competitive benchmarking tasks. These participants also made extensive use of the Collection Grid and sustainability filters, describing them as "a clear step forward compared to manual search or traditional tools". Many reported that the ability to quickly visualize competitive data "helped simulate real decision-making scenarios". Trend analysts showed particularly high engagement during Free Navigation (UES +0.5 points over designers), benefiting from the system's social media analytics capabilities. The phased evaluation revealed an important progression: while initial learning required moderate effort (as shown in Guided Navigation metrics), users quickly achieved proficiency (Free Navigation results) and ultimately found the system highly valuable (post-experiment scores). This pattern was particularly strong for participants with prior fashion analytics experience, who showed steeper learning curves and higher final satisfaction scores. Participants reported high satisfaction in visual exploration, interactivity, and with the integration of AI-powered insights. The collection grids were praised for their clarity and filtering capabilities, while the 3D configurator was defined as "intuitive and useful for a complete experience". The social media analysis platform was especially valued for its geospatial trend maps, which participants described as "crucial for localizing fashion marketing strategies". Participants noted that the sustainability tags and filters allowed for data-informed ethical choices and highlighted the need for richer metadata on second-hand and recycled materials.

These results demonstrate that Made-In successfully balances analytical depth with usability, though opportunities exist to reduce cognitive load in advanced features. The consistently high engagement scores, particularly during free exploration, validate the system's ability to support creative fashion workflows while maintaining usability standards comparable to established professional tools. The findings also highlight the importance of role-specific customization, as different professional groups leveraged distinct system features to achieve their goals.

5. Conclusions and future works

This paper introduces Made-In, an immersive analytics system designed to promote data-driven, inclusive and sustainable practices within the fashion industry. By integrating AI-powered components for trend detection, assortment planning and visual merchandising, Made-In offers fashion professionals a unified platform that enhances their creative and strategic decision-making processes. By combining data from e-commerce platforms and social media, particularly Instagram, with interactive visualization tools and AI-driven metadata enrichment, Made-In provides real-time, evidence-based insights into consumer preferences, product characteristics, and market trends. The results of our user study confirmed the system's usability and relevance across a diverse range of industry roles. Participants emphasized the intuitive nature of the interface, the value of region-specific trend visualizations and the tool's potential to support creative inspiration and operational efficiency. In particular, the ability to compare collections in real time, anticipate sales trends and detect geolocated social signals was considered essential for maintaining competitiveness in a rapidly evolving market. Integrating sustainability criteria such as recycled materials and ethical sourcing also empowers users to make more responsible and transparent decisions, aligning the platform with the EU Green Deal and Circular Economy objectives. From a technical standpoint, Made-In tackles several issues associated with the deployment of AI in the fashion sector, such as data heterogeneity, scalability of real-time inference, and the design of visual interfaces for non-technical users. The architecture has been developed with modularity in mind to support container-based deployment and future integration with enterprise systems. This infrastructure ensures performance and flexibility, and lays the foundation for expanding the system to new fashion categories and application areas.

Looking ahead, the next phase of research will focus on enhancing the system's cognitive and generative capabilities. A key priority will be to demonstrate the adaptability of the Made-In framework to a broader range of fashion product categories, such as footwear, outerwear, and accessories. This will involve configuring the pipeline to accommodate new metadata schemas, visual features, and user interaction patterns, supported by targeted experiments and comparative evaluations. In parallel, we plan to incorporate multimodal foundation models that can anticipate emerging styles by synthesizing diverse inputs, such as text, imagery, and user behavior. These models will improve the system's ability to generate trend forecasts and stylistic recommendations in a more adaptive and personalized way. There will be greater emphasis on explainability, with the introduction of features that improve the transparency of recommendations and the interpretability of AI-generated insights, especially important for fashion professionals making high-stakes design and merchandising decisions. Future iterations of the system will also address the growing relevance of second-hand and circular fashion markets by enabling condition-aware classification, product lifecycle tracking and extended sustainability analytics. Furthermore, cultural diversity and inclusivity will remain foundational principles, guiding the expansion of training datasets to reflect a broader spectrum of body types, regional aesthetics, and fashion practices. To foster reproducibility and collaborative innovation, we plan to release anonymized datasets and provide modular APIs for external integration. These resources will enable researchers, designers, and technologists to build on the Made-In architecture and contribute to the development of a more ethical, inclusive, and sustainable digital fashion ecosystem.

CRediT authorship contribution statement

Emanuele Balloni: Writing – original draft, Methodology, Conceptualization. **Rocco Pietrini:** Methodology, Conceptualization. **Michele Sasso:** Validation. **Emanuele Frontoni:** Supervision, Writing – review & editing. **Marina Paolanti:** Writing – review & editing, Writing – original draft, Supervision, Methodology, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was funded by Tod's Spa. The authors would like to thank Arianna Vecchi and Marco Mameli for their support.

Data availability

Data will be made available on request.

References

Abdal, R., Qin, Y., Wonka, P., 2019. Image2stylegan: How to embed images into the stylegan latent space? In: Proceedings of the IEEE/CVF International Conference on Computer Vision. pp. 4432–4441.

Abdal, R., Qin, Y., Wonka, P., 2020. Image2stylegan++: How to edit the embedded images? In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 8296–8305.

Anjorin, K.F., Raji, M.A., Olodo, H.B., 2024. The influence of social media marketing on consumer behavior in the retail industry: A comprehensive review. Int. J. Manag. Entrep. Res. 6 (5), 1547–1580.

Balloni, E., Pietrini, R., Fabiani, M., Frontoni, E., Mancini, A., Paolanti, M., 2024. Social4Fashion: An intelligent expert system for forecasting fashion trends from social media contents. Expert Syst. Appl. 252, 124018.

Balloni, E., Stacchio, L., Mancini, A., Frontoni, E., Zingaretti, P., Paolanti, M., 2025. A Neural Rendering system for fashion design process. Eng. Appl. Artif. Intell. 153, 110773.

Baykal, A.C., Anees, A.B., Ceylan, D., Erdem, E., Erdem, A., Yuret, D., 2023. CLIP-guided StyleGAN inversion for text-driven real image editing. ACM Trans. Graph. 42 (5), 1–18.

Brooke, J., et al., 1996. SUS-a quick and dirty usability scale. Usability Eval. Ind. 189 (194), 4–7.

Burnes, B., Towers, N., 2016. Consumers, clothing retailers and production planning and control in the smart city. Prod. Plan. Control 27 (6), 490–499.

Cao, Y., Han, K., Wong, K.-Y.K., 2023. Sesdf: Self-evolved signed distance field for implicit 3d clothed human reconstruction. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 4647–4657.

Cheng, W.-H., Song, S., Chen, C.-Y., Hidayati, S.C., Liu, J., 2021. Fashion meets computer vision: A survey. ACM Comput. Surv. 54 (4), 1–41.

Chheang, V., Sharmin, S., Márquez-Hernández, R., Patel, M., Rajasekaran, D., Caulfield, G., Kiafar, B., Li, J., Kullu, P., Barmaki, R.L., 2024. Towards anatomy education with generative AI-based virtual assistants in immersive virtual reality environments. In: 2024 IEEE International Conference on Artificial Intelligence and EXtended and Virtual Reality (AIxVR). IEEE, pp. 21–30.

Cui, Y.R., Liu, Q., Gao, C.Y., Su, Z., 2018. FashionGAN: Display your fashion design using conditional generative adversarial nets. In: Computer Graphics Forum. vol. 37, (7), Wiley Online Library, pp. 109–119.

Davis, F.D., et al., 1989. Technology acceptance model: TAM. Al- Suqri, MN, Al- Aufi, AS: Inf. Seek. Behav. Technol. Adopt. 205, 219.

De Luigi, L., Li, R., Guillard, B., Salzmann, M., Fua, P., 2023. Drapenet: Garment generation and self-supervised draping. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 1451–1460.

de Oliveira, C.R.S., da Silva Júnior, A.H., Mulinari, J., Immich, A.P.S., 2021. Textile Re-Engineering: Eco-responsible solutions for a more sustainable industry. Sustain. Prod. Consum. 28, 1232–1248.

Della Sciucca, L., Balloni, E., Mameli, M., Frontoni, E., Zingaretti, P., Paolanti, M., 2022. StyleTrendGAN: A deep learning generative framework for fashion bag generation. In: Image Analysis and Processing. ICIAP 2022 Workshops: ICIAP International Workshops, Lecce, Italy, May 23–27, 2022, Revised Selected Papers, Part II. Springer-Verlag, Berlin, Heidelberg, ISBN: 978-3-031-13323-7, pp. 191–202. http://dx.doi.org/10.1007/978-3-031-13324-4_17.

Dong, H., Yu, S., Wu, C., Guo, Y., 2017. Semantic image synthesis via adversarial learning. In: Proceedings of the IEEE International Conference on Computer Vision. pp. 5706–5714.

Fang, N., Qiu, L., Zhang, S., Wang, Z., Hu, K., Wang, K., 2023. A novel DAGAN for synthesizing garment images based on design attribute disentangled representation. Pattern Recognit. 136, 109248.

Fele, B., Lampe, A., Peer, P., Struc, V., 2022. C-vton: Context-driven image-based virtual try-on network. In: Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision. pp. 3144–3153.

Ge, C., Song, Y., Ge, Y., Yang, H., Liu, W., Luo, P., 2021a. Disentangled cycle consistency for highly-realistic virtual try-on. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 16928–16937.

Ge, Y., Song, Y., Zhang, R., Ge, C., Liu, W., Luo, P., 2021b. Parser-free virtual try-on via distilling appearance flows. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 8485–8493.

Giovanola, B., Tiribelli, S., Frontoni, E., Paolanti, M., 2023. Ethical implication of Artificial Intelligence in the Fashion Industry: A comprehensive analysis. Fash. Highlight (2), 22–28.

Githapradana, W., Raharja, G., Ratna, T., Pebryani, N., 2024. Gender and culture inclusivity in the story of hope fashion collection. ARCHive- SR 8 (1), 43–51.

Grigorev, A., Black, M.J., Hilliges, O., 2023. Hood: Hierarchical graphs for generalized modelling of clothing dynamics. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 16965–16974.

Han, X., Wu, Z., Wu, Z., Yu, R., Davis, L.S., 2018. Viton: An image-based virtual try-on network. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. pp. 7543–7552.

Hart, S.G., Staveland, L.E., 1988. Development of NASA-TLX (task load index): Results of empirical and theoretical research. In: Advances in Psychology. vol. 52, Elsevier, pp. 139–183.

He, S., Song, Y.-Z., Xiang, T., 2022. Style-based global appearance flow for virtual try-on. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 3470–3479.

Jafarian, Y., Wang, T.Y., Ceylan, D., Yang, J., Carr, N., Zhou, Y., Park, H.S., 2023. Normal-guided garment UV prediction for human re-texturing. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 4627–4636.

Jain, R., Singh, K.K., Hemani, M., Lu, J., Sarkar, M., Ceylan, D., Krishnamurthy, B., 2023. Vgflow: Visibility guided flow network for human reposing. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 21088–21097.

Jiang, Y., Yang, S., Qiu, H., Wu, W., Loy, C.C., Liu, Z., 2022. Text2human: Text-driven controllable human image generation. ACM Trans. Graph. 41 (4), 1–11.

Karaman, S., Bagdanov, A.D., Landucci, L., D'amico, G., Ferracani, A., Pezzatini, D., Del Bimbo, A., 2016. Personalized multimedia content delivery on an interactive table by passive observation of museum visitors. Multimedia Tools Appl. 75, 3787–3811.

Karaosman, H., Brun, A., Morales-Alonso, G., 2017. Vogue or vague: sustainability performance appraisal in luxury fashion supply chains. Sustain. Manag. Lux. 301–330.

Karnad, V., Udiaver, V., 2022. Social manufacturing in the fashion industry to generate sustainable fashion value creation. In: Digital Manufacturing Technology for Sustainable Anthropometric Apparel. Elsevier, pp. 49–67.

Karras, T., Aila, T., Laine, S., Lehtinen, J., 2017. Progressive growing of gans for improved quality, stability, and variation. arXiv preprint arXiv:1710.10196.

Karras, T., Aittala, M., Hellsten, J., Laine, S., Lehtinen, J., Aila, T., 2020. Training generative adversarial networks with limited data. Adv. Neural Inf. Process. Syst. 33, 12104–12114.

Karras, T., Laine, S., Aila, T., 2019. A style-based generator architecture for generative adversarial networks. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 4401–4410.

Kerbl, B., Kopanas, G., Leimkühler, T., Drettakis, G., 2023. 3D Gaussian splatting for real-time radiance field rendering. ACM Trans. Graph. 42 (4).

Khelladi, I., Lejealle, C., Rezaee Vessal, S., Castellano, S., Graziano, D., 2024. Why do people buy virtual clothes? J. Consum. Behav. 23 (3), 1389–1405.

Lee, J.H., Ahn, J., Kim, J., 2018. Theoretical competence model of fashion designers in co-designed fashion systems. Fash. Pr. 10 (3), 381–404.

Lewis, J.R., 1991. Psychometric evaluation of an after-scenario questionnaire for computer usability studies: the ASQ. ACM Sigchi Bull. 23 (1), 78–81.

Li, B., Qi, X., Lukasiewicz, T., Torr, P.H., 2020. Manigan: Text-guided image manipulation. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 7880–7889.

Longo, L., 2017. Subjective usability, mental workload assessments and their impact on objective human performance. In: IFIP Conference on Human-Computer Interaction. Springer, pp. 202–223.

Melo, M., Coelho, H., Gonçalves, G., Losada, N., Jorge, F., Teixeira, M.S., Bessa, M., 2022. Immersive multisensory virtual reality technologies for virtual tourism: A study of the user's sense of presence, satisfaction, emotions, and attitudes. Multimedia Syst. 28 (3), 1027–1037.

Mildenhall, B., Srinivasan, P.P., Tancik, M., Barron, J.T., Ramamoorthi, R., Ng, R., 2020. NeRF: Representing scenes as neural radiance fields for view synthesis. In: ECCV.

Nam, S., Kim, Y., Kim, S.J., 2018. Text-adaptive generative adversarial networks: manipulating images with natural language. Adv. Neural Inf. Process. Syst. 31.

Napolano, L., Foschi, J., Caldeira, C., Huygens, D., Sala, S., 2025. Understanding textile value chains: Dynamic Probabilistic Material Flow Analysis of textile in the European Union. Resour. Conserv. Recycl. 212, 107888.

Niinimäki, K., Peters, G., Dahlbo, H., Perry, P., Rissanen, T., Gwilt, A., 2020. The environmental price of fast fashion. Nat. Rev. Earth Environ. 1 (4), 189–200.

O'Brien, H.L., Cairns, P., Hall, M., 2018. A practical approach to measuring user engagement with the refined user engagement scale (UES) and new UES short form. Int. J. Hum.-Comput. Stud. 112, 28–39.

O'Brien, H.L., Toms, E.G., 2010. The development and evaluation of a survey to measure user engagement. J. Am. Soc. Inf. Sci. Technol. 61 (1), 50–69.

Olive, C., 2023. Value, virtue, and vivienne westwood: On the philosophical importance of fashion. Open Philos. 6 (1), 20220250.

Patashnik, O., Wu, Z., Shechtman, E., Cohen-Or, D., Lischinski, D., 2021. Styleclip: Text-driven manipulation of stylegan imagery. In: Proceedings of the IEEE/CVF International Conference on Computer Vision. pp. 2085–2094.

Qiao, T., Zhang, J., Xu, D., Tao, D., 2019. Mirrorgan: Learning text-to-image generation by redescription. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 1505–1514.

Qin, X., Dai, H., Hu, X., Fan, D.-P., Shao, L., Gool, L.V., 2022. Highly accurate dichotomous image segmentation. In: ECCV.

Qiu, L., Chen, G., Zhou, J., Xu, M., Wang, J., Han, X., 2023. Rec-mv: Reconstructing 3d dynamic cloth from monocular videos. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 4637–4646.

Radford, A., Kim, J.W., Hallacy, C., Ramesh, A., Goh, G., Agarwal, S., Sastry, G., Askell, A., Mishkin, P., Clark, J., et al., 2021. Learning transferable visual models from natural language supervision. In: International Conference on Machine Learning. PMLR, pp. 8748–8763.

Radford, A., Metz, L., Chintala, S., 2015. Unsupervised representation learning with deep convolutional generative adversarial networks. arXiv preprint arXiv:1511.06434.

Ramesh, A., Dhariwal, P., Nichol, A., Chu, C., Chen, M., 2022. Hierarchical text-conditional image generation with clip latents. 1, (2), p. 3, arXiv preprint arXiv:2204.06125.

Reed, S., Akata, Z., Yan, X., Logeswaran, L., Schiele, B., Lee, H., 2016. Generative adversarial text to image synthesis. In: International Conference on Machine Learning. PMLR, pp. 1060–1069.

Rostamzadeh, N., Hosseini, S., Boquet, T., Stokowiec, W., Zhang, Y., Jauvin, C., Pal, C., 2018. Fashion-gen: The generative fashion dataset and challenge. arXiv preprint arXiv:1806.08317.

Saharia, C., Chan, W., Saxena, S., Li, L., Whang, J., Denton, E.L., Ghasemipour, K., Gontijo Lopes, R., Karagol Ayan, B., Salimans, T., et al., 2022. Photorealistic text-to-image diffusion models with deep language understanding. Adv. Neural Inf. Process. Syst. 35, 36479–36494.

Sanh, V., Debut, L., Chaumond, J., Wolf, T., 2019. DistilBERT, a distilled version of BERT: smaller, faster, cheaper and lighter. arXiv preprint arXiv:1910.01108.

Schönberger, J.L., Frahm, J.-M., 2016. Structure-from-motion revisited. In: Conference on Computer Vision and Pattern Recognition. CVPR.

Sharma, G., Wu, W., Dalal, E.N., 2005. The CIEDE2000 color-difference formula: Implementation notes, supplementary test data, and mathematical observations. Color. Res. Appl.: Endorsed By Inter- Soc. Color. Counc. Colour Group (Gt. Britain), Can. Soc. Color. Color. Sci. Assoc. Jpn. Dutch Soc. the Study Color. Swed. Colour Cent. Found. Colour Soc. Aust. Cent. Fran c Cais de la Coul. 30 (1), 21–30.

Shen, Y., Yang, C., Tang, X., Zhou, B., 2020. Interfacegan: Interpreting the disentangled face representation learned by gans. IEEE Trans. Pattern Anal. Mach. Intell. 44 (4), 2004–2018.

Simonyan, K., Zisserman, A., 2014. Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556.

Song, D.-Y., Lee, H., Seo, J., Cho, D., 2023. Difu: Depth-guided implicit function for clothed human reconstruction. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 8738–8747.

Touvron, H., Lavril, T., Izacard, G., Martinet, X., Lachaux, M.-A., Lacroix, T., Rozière, B., Goyal, N., Hambro, E., Azhar, F., et al., 2023. Llama: Open and efficient foundation language models. arXiv preprint arXiv:2302.13971.

Vazquez, D., Cheung, J., Nguyen, B., Dennis, C., Kent, A., 2021. Examining the influence of user-generated content on the fashion consumer online experience. J. Fash. Mark. Manag.: Int. J. 25 (3), 528–547.

Vehmas, K., Raudaskoski, A., Heikkilä, P., Harlin, A., Mensonen, A., 2018. Consumer attitudes and communication in circular fashion. J. Fash. Mark. Manag.: Int. J. 22 (3), 286–300.

Wang, Z., Yuan, L.-P., Wang, L., Jiang, B., Zeng, W., 2024. Virtuwander: Enhancing multi-modal interaction for virtual tour guidance through large language models. In: Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems. pp. 1–20.

Wang, K., Zhang, G., Cong, S., Yang, J., 2023. Clothed human performance capture with a double-layer neural radiance fields. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 21098–21107.

Wojciechowska, P., 2021. Fibres and textiles in the circular economy. In: Fundamentals of Natural Fibres and Textiles. Elsevier, pp. 691–717.

Wu, X., Xu, K., Hall, P., 2017. A survey of image synthesis and editing with generative adversarial networks. Tsinghua Sci. Technol. 22 (6), 660–674.

Xia, W., Zhang, Y., Yang, Y., Xue, J.-H., Zhou, B., Yang, M.-H., 2022. Gan inversion: A survey. IEEE Trans. Pattern Anal. Mach. Intell. 45 (3), 3121–3138.

Xiu, Y., Yang, J., Cao, X., Tzionas, D., Black, M.J., 2023. Econ: Explicit clothed humans optimized via normal integration. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 512–523.

Xu, T., Zhang, P., Huang, Q., Zhang, H., Gan, Z., Huang, X., He, X., 2018. Attngan: Fine-grained text to image generation with attentional generative adversarial networks. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. pp. 1316–1324.

Yang, H., Zhang, R., Guo, X., Liu, W., Zuo, W., Luo, P., 2020. Towards photo-realistic virtual try-on by adaptively generating-preserving image content. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 7850–7859.

Zhang, H., Lin, S., Shao, R., Zhang, Y., Zheng, Z., Huang, H., Guo, Y., Liu, Y., 2023. Closet: Modeling clothed humans on continuous surface with explicit template decomposition. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 501–511.

Zhang, H., Xu, T., Li, H., Zhang, S., Wang, X., Huang, X., Metaxas, D.N., 2017. Stackgan: Text to photo-realistic image synthesis with stacked generative adversarial networks. In: Proceedings of the IEEE International Conference on Computer Vision. pp. 5907–5915.

Zhang, H., Xu, T., Li, H., Zhang, S., Wang, X., Huang, X., Metaxas, D.N., 2018. Stackgan++: Realistic image synthesis with stacked generative adversarial networks. IEEE Trans. Pattern Anal. Mach. Intell. 41 (8), 1947–1962.

Zhang, D., Zuo, C., Wu, Q., Fu, L., Xiang, X., 2022. Unabridged adjacent modulation for clothing parsing. Pattern Recognit. 127, 108594.

Zhao, F., Li, Z., Huang, S., Weng, J., Zhou, T., Xie, G.-S., Wang, J., Shan, Y., 2023. Learning anchor transformations for 3d garment animation. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition pp. 491–500.

Zhu, J.-Y., Park, T., Isola, P., Efros, A.A., 2017. Unpaired image-to-image translation using cycle-consistent adversarial networks. In: Proceedings of the IEEE International Conference on Computer Vision. pp. 2223–2232.

Zhu, L., Yang, D., Zhu, T., Reda, F., Chan, W., Saharia, C., Norouzi, M., Kemelmacher-Shlizerman, I., 2023. Tryondiffusion: A tale of two unets. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 4606–4615.

Zou, X., Han, X., Wong, W., 2023. Cloth4d: A dataset for clothed human reconstruction. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 12847–12857.