An Experimentation Engine for Data-Driven Fashion Systems

Ranjitha Kumar and Kristen Vaccaro
Department of Computer Science
University of Illinois at Urbana-Champaign
{ranjitha,kvaccaro}@illinois.edu

Abstract

Data-driven fashion systems of the future will revolutionize the way consumers shop for clothing and choose outfits: imagine an automated personal stylist that ships clothes straight to your door based on their compatibility with your existing wardrobe, the upcoming events on your calendar, and style trends learned from the web. To build such systems, we must identify the fashion activities that are the largest consumer pain points, the interventions necessary to alleviate those pains, and the computational models that enable those interventions.

To guide the design of these next-generation tools, we propose an experimentation engine for fashion interfaces: leveraging social media platforms to run multivariate design tests with thousands to millions of users. Social platforms are already home to dedicated communities of fashion enthusiasts, and expose programmable agents — chatbots — that can be used to rapidly prototype data-driven design interfaces. Measuring the number of followers and user engagement amongst these prototypes can inform the design of future standalone fashion systems. At this workshop, we will sketch the design space of fashion experiments, and present preliminary results from deploying our "fashion bots."

Introduction

"The clothes on the hanger do nothing; the clothes on the woman do everything."

– Stephen Breyer

The internet has revolutionized the way consumers purchase clothing by disrupting their reliance on brick-and-mortar stores. However, the way in which consumers *shop* remains relatively unchanged. While data and analytics have permeated the retail experience in logistics and ERP, attempts at user personalization and recommendations have primarily been confined to picking which advertisement to show, and online retailers are mostly relegated to marketplaces from which to buy things one already knows ones wants.

We posit that data-driven fashion systems of the future will revolutionize the way consumers shop for clothing and choose outfits. Imagine an automated personal stylist that ships clothes straight to your door, based on their compatibility with your existing wardrobe, your budget, the upcoming events on your calendar, and style trends learned from the web.

To build these next-generation fashion systems, we must identify the fashion activities that are the largest consumer pain points, the interventions necessary to alleviate those pains, and the computational models that enable those interventions. To do so, we need data at a scale amenable to state-of-the art machine intelligence. Accordingly, we propose an *experimentation engine* for fashion interfaces: leveraging social media platforms to run multivariate design tests with thousands to millions of users.

Social platforms are a natural choice for these efforts, since they are already home to dedicated communities of fashion enthusiasts and expose programmable agents — chatbots — that can be used to prototype data-driven design interfaces. Measuring user interactions with these prototypes can rapidly inform the design of future standalone fashion systems.

In this position paper, we sketch the space of fashion activities that are ripe for technological intervention; discuss a few promising classes of system interventions; describe a set of data-driven models that could power them; identify a set of fashion data sources that might be used to back the models; and propose an experimentation engine that leverages popular social media platforms to rapidly prototype, deploy, and test the next generation of fashion interfaces.

Fashion Activities

Humans have a complex relationship with clothes. Beyond choosing which pair of pants to wear on any given day, people continually grapple with fashion-related problems ranging from "how much will this new jacket extend the versatility of my wardrobe?" to "how should I dress to convey strength and competence in this business meeting?" to "can I achieve the same 'look' as that celebrity with a vastly inferior budget?" To build next-generation fashion systems, we must first understand the greatest user pain points experienced in common fashion activities.

Outfit Creation. The most fundamental fashion activity is deciding what to wear each day. When creating outfits, people optimize for both form and function, balancing practical considerations such as weather, occasion, and budget with

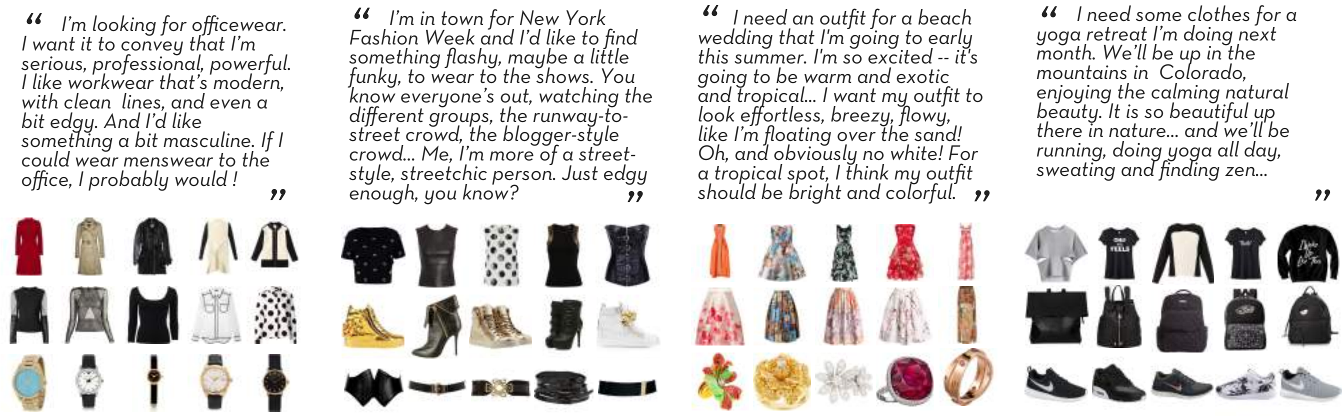

Figure 1: An automated personal stylist taking descriptions of outfit needs in natural language as input to produce item recommendations.

personal preferences around style, silhouettes, and material. To choose an outfit, one must assemble a set of individual clothing items that are compatible with each other and "come together" to form a cohesive concept.

Computation can help find combinations of compatible items that meet constraints. Fashion systems of the future will assist users as they deal with the diverse situations and constraints of daily life, helping them find an outfit for a "business casual dinner," achieve the "same look for less," or create an outfit that transitions from "workwear" to "partywear" with just a few modifications.

Wardrobe Management. In addition to selecting items *from* their wardrobe, users must regularly curate *for* their wardrobe by adding and removing pieces. People add new clothes for a variety of reasons: to complement their existing wardrobe, diversify the types of outfits they create, to reinforce their existing style, or to experiment with a new look or trend.

Computational systems can help people assess whether new items are good investments: versatile in the context of their existing wardrobe, appropriate for a special event, a classic that will never go out of style, or on trend to signal that a user is keeping up with the fashion Joneses.

In addition to adding new items to one's wardrobe, people regularly remove old items because they have worn out, no longer fit, or have gone out of style. In the age of *fast fashion*, where brands produce new styles frequently and inexpensively, "consumers can afford to buy ... in quantities" never seen before (Zarroli 2013). Computation can help users pare down their wardrobes in an optimal manner to make way for new pieces.

In response to fast fashion's rampant consumerism, some fashion-savvy and cost-conscious people have embraced concepts like *capsule collections*, in which one intentionally limits their wardrobe to make dressing simpler, save money, and focus on pieces that evoke strong emotions. Although Steve Jobs' black turtlenecks and blue jeans are an extreme example, these mini-wardrobes typically comprise thirty or forty pieces that are both versatile and well-loved by their owners (Rector 2014). While there are many instructional guides online for capsule neophytes, computational systems could greatly enhance the capsule-building process by identifying highly-compatible collections that minimize the number of pieces while maximizing the number of potential outfits.

Social Feedback. While the affluent can avail themselves to personal stylists for targeted fashion advice, many users look to their social network when they consider buying new items or put together an outfit for a new event. The desire for this sort of social feedback underscores the role of fashion as a vehicle for *social signaling*: people wear clothes to send messages to others[1].

Signaling theory in fashion (Donath 2007) explores the way people use clothing to indicate wealth and status (Nelissen and Meijers 2011), sexual motivations (Grammer, Renninger, and Fischer 2004), and self-defined roles (Piacentini and Mailer 2004). These signals present another opportunity for computational systems to find ways to capture the intent behind certain outfits and to provide flexible platforms for self-expression.

System Interventions

System interventions address problems encountered during fashion activities. These interventions can range from lightweight "nudges," to completely autonomous action taken by a fashion system. For a fashion experimentation engine to be useful, it should allow researchers to quickly test the efficacy of many different interventions for solving a particular fashion problem.

Nudges. Small nudges from computational systems — interventions that provide information without advocating for specific action — can result in substantive behavioral changes in fashion. For example, many users fall into regular patterns of dressing and fail to explore even the well-matched options already present in their wardrobes. Tsujita

[1]One important social message is "I am not naked," particularly when one is a head of state.

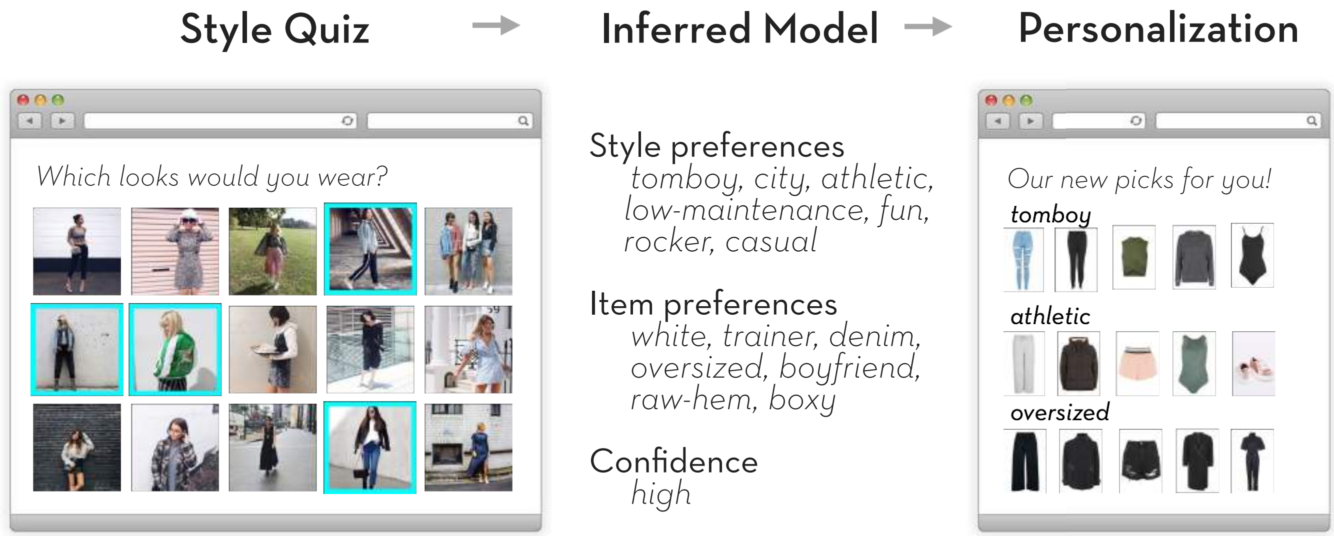

Figure 2: This hypothetical interface illustrates how recommendations based on users' style and item preferences allow systems to personalize shopping experiences and make better recommendations.

et al. suggest a computational system that shows users items from their wardrobe while blurring clothing which has been recently worn (Tsujita et al. 2010). They note that one user "was apt to wear her favorite clothes many times ... [and] didn't wear the clothes stored in the back of her closet," but substantially varied her choices and became more conscious of repetition once using the system.

Suggestions. While nudges can provide users with valuable information even when they are not actively looking for it, users also seek answers to well-formed fashion questions such as "what should I wear to dinner?" Suggestion interfaces — which answer user queries with recommended courses of action — can reduce the cognitive burden on users by helping them make decisions (Fig. 1).

Researchers have proposed a number of computational schemes for making fashion suggestions, either of particular items (McAuley et al. 2015; McAuley, Pandey, and Leskovec 2015; Veit et al. 2015; Di et al. 2013) or of complete outfits (Liu et al. 2012a; Shen, Lieberman, and Lam 2007; Yu et al. 2012; Vartak and Madden 2013). Next generation fashion systems must personalize their recommendations by seamlessly[2] accounting for user preferences, purchase histories, and the contents of their wardrobes.

Autonomous Actions. Sophisticated computational fashion systems could even be trusted to take independent action on behalf of their users. Existing personal stylist services like TrunkClub and Stitchfix mail their subscribers clothing each month, with the goal that users will buy the items that are sent to them. While these services are presently driven by human curation, it is easy to imagine machine learning playing a more prominent role in curation.

Similarly, one could imagine a "magic closet" that lays out an outfit each morning (Liu et al. 2012a) for a user to accept or reject. By engendering a tight feedback loop and correlating fashion choices with holistic data about a user's life (i.e., "I don't want to wear that white skirt because they're serving spaghetti for lunch today"), useful predictive systems could be constructed.

[2]Pun intended.

Data-Driven Models

Data-driven models are the key to realizing many of the system interventions described in this paper. Different types of models can be layered or used separately to support a variety of user interactions: to *understand* fashion trends, *generate* outfit recommendations, *evaluate* whether items match, etc. An experimentation engine should allow researchers to rapidly test how to effectively combine models to support different interactions.

Image parsing. Researchers in computer vision have had some success identifying items in outfits (Yamaguchi, Kiapour, and Berg 2013) and identifying attributes of individual items (Berg, Berg, and Shih 2010; Vittayakorn et al. 2015), leading to innovative search patterns for fashion data (Kovashka, Parikh, and Grauman 2012). They have even been able to evaluate outfit style, both for individuals (Kiapour et al. 2014; Song et al. 2011; Simo-Serra and Ishikawa 2016), groups (Kwak et al. 2013; Murillo et al. 2012), and clothing items (Di et al. 2013; Veit et al. 2015; McAuley et al. 2015). Recent work has measured overall outfit fashionability from images of outfits (Simo-Serra et al. 2015).

Outfit compatibility. Several existing systems measure outfit compatibility or generate compatible outfits, either via low-level hand-annotated features (Liu et al. 2012a; Shen, Lieberman, and Lam 2007; Yu et al. 2012; Vartak and Madden 2013) or higher-order ones generated, for instance, via deep learning (McAuley et al. 2015; McAuley, Pandey, and Leskovec 2015; Veit et al. 2015; Di et al. 2013).

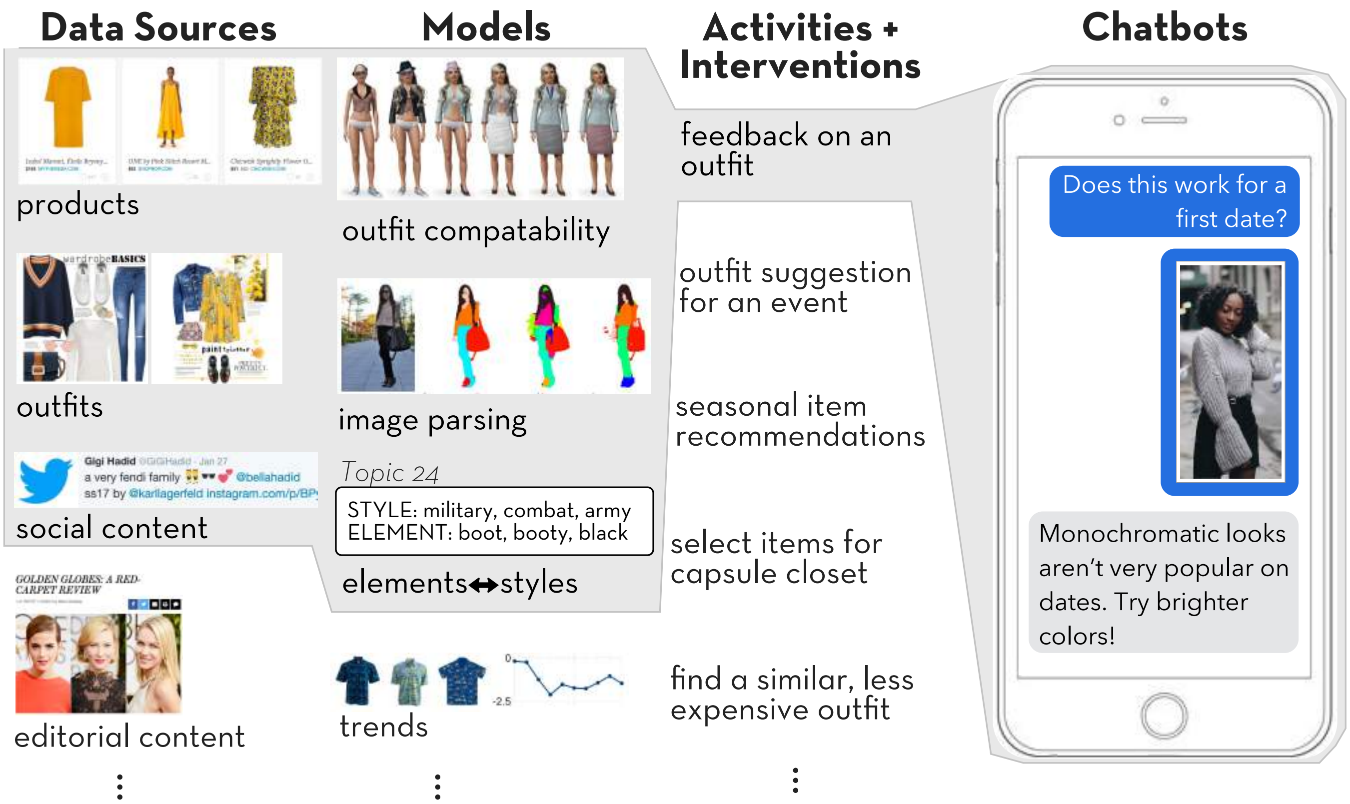

Figure 3: Architecture of a fashion experimentation engine: data sources and models are used to build chatbots that support a variety of activities and interventions. Images representing models for outfit compatibility (Yu et al. 2012), image parsing (Yamaguchi, Kiapour, and Berg 2013), elements to styles (Vaccaro et al. 2016), and trends (He and McAuley 2015) are drawn from their respective papers.

Styles. Existing fashion systems have also described items or outfits in terms of their styles (Kiapour et al. 2014; Song et al. 2011; Simo-Serra and Ishikawa 2016; Veit et al. 2015; Di et al. 2013; McAuley et al. 2015; Yu et al. 2012; Vaccaro et al. 2016). Simiarly, many systems that generate outfits do so with style constraints (Liu et al. 2012a; McAuley et al. 2015; Shen, Lieberman, and Lam 2007; Yu et al. 2012).

Trends. The evolution of fashion over time has attracted a great deal of attention from fashion researchers (Au, Choi, and Yu 2008; Alon, Qi, and Sadowski 2001; Hidayati et al. 2014; He and McAuley 2015; Lin, Zhou, and Xu 2015; Vittayakorn et al. 2015). Trickle-down theories of fashion suggest that the middle-class adopted trends from the rich emulating them and signaling their wealth, while more recent trickle-up and trickle-across theories posit that trends come “from the street” (English 2007). The ability to identify emerging trends and accurately predict their life-cycles would empower next-generation fashion systems (Trufelman 2016).

Personalization. Modeling users’ style and item preferences allows systems to personalize shopping experiences and make better recommendations (Figure 2). User preference data can be collected directly (i.e., style quizzes) or inferred from purchase history, browsing patterns, saved products, and product reviews. To be effective, preference models must be sensitive to how tastes and wardrobe requirements change over time: for example, there can be a significant shift in wardrobe composition when a student graduates from college and enters the workforce.

Fashion Data

Many kinds of data can inform data-driven models of fashion: product information, user data, social and editorial content, and more. For example, a fashion trends model could be trained by combining streams of editorial, social, and ecommerce content. Training models by combining data from multiple sources can be challenging, however, due to disparities in data formats and update rates.

Product Information. Several data-driven fashion systems have focused on characterizing products (Berg, Berg, and Shih 2010; Di et al. 2013; Veit et al. 2015; McAuley et al. 2015). These models are trained on diverse sets of product information such as images; prices; and text descriptions capturing elements such as designer, color, material, and silhouette. Sometimes, product datasets include other related products which can be used as substitutes or complements. These datasets can be drawn from popular aggregative shopping websites (e.g., Amazon, Modcloth), individual brand websites, and social networks with product data (e.g., Polyvore, TheHunt).

User Data. Personalization models are powered by user data. Customer purchase history, browsing patterns, and reviews are all examples of user data leveraged by ecommerce platforms to personalize shopping experiences. In the future, data generated through interactions with conversational agents can also be used to model users.

Social Content. Fashion work that seeks to understand clothing *as it is worn* often uses social content, particularly images from fashion-related social networks like Lookbook and Chictopia (Yamaguchi, Kiapour, and Berg 2013; Vittayakorn et al. 2015; Simo-Serra and Ishikawa 2016; Yu et al. 2012). Researchers also draw images from more general social networks, like Flickr and Instagram, often filtering to fashion related content with queries such as "street shot" (Liu et al. 2012b). Systems may also leverage social content such as text, indicators of popularity (e.g., likes, shares), and the social network structure itself (Lin et al. 2015).

Editorial Content. Researchers leverage editorial content such as magazine articles and fashion blogs to study trends and trendsetters: how ideas spread in a fashion network and how designers, celebrities, and retailers influence each other (Vittayakorn et al. 2015; Lin, Zhou, and Xu 2015). For example, the genre of street style blogs feature fashion photographers who seek out fashionable people on the street (e.g. The Sartorialist, Street Peeper). Magazine articles include content like runway reviews and critiques of trends and new pieces.

Building an Experimentation Engine

Even after describing the space of fashion activities, system interventions, data-driven models, and data sources, designing the next generation of fashion systems remains a complex task. *A priori*, it is impossible to predict which user pain points are the greatest, which system interventions provide the best user experiences, and which data models are the most suitable for a given task. Therefore, we propose to develop an *experimentation engine* for fashion systems that allows researchers to rapidly prototype, deploy, and test design variations (Figure 3).

Fashion Chatbots. To build an engine that can support multivariate design tests with thousands to millions of users, we turn to social media platforms. Platforms such as Facebook, Twitter, and Instagram are already home to dedicated communities of fashion enthusiasts, and expose programmable agents — chatbots — that can be used to rapidly prototype natural data-driven design interfaces. Large companies are increasingly focused on developing domain-specific conversational agents to solve a constrained set of user problems (Mortensen 2016; Radziszewski 2016). Moreover, startups have begin to explore the potential of "fashionbots" to replace traditional websites by providing shopping experiences that directly solve user problems (Sharon 2017). Instead of typing keywords into a search box or navigating hierarchical menus to find items, users can simply type commands such as "white denim," "more like this," or "new arrivals" to quickly explore catalogs of products.

Staged Automation. Similar to the chatbot work of Huang et al. (Huang et al. 2016), fashionbots can be built out in stages. Initially, fashionbots can be backed by humans instead of computational systems, and wizard-of-oz experiments can be run to understand which pain points and types of system interventions elicit the greatest user engagement. These human-powered systems can reveal the psychological motivations that govern dressing and purchasing behavior, helping to prioritize the set of systems and models that should be implemented. Perhaps users would be incentivized to take daily selfies after getting dressed, if a bot would then offer an opinion of the style and affect their outfit conveyed. Once and only once such an interaction is determined to be *useful*, researchers could develop the image parsing and style data models to support it.

Platform-specific Affordances. Different social platforms naturally support different types of interactions. Understanding platform-specific affordances for fashion activities is critical for building effective bots. Facebook messenger bots are well-suited for private conversations similar to those one might have with a personal stylist, while Twitter bots could rapidly disseminate content like fashion editorials, using strategic hashtags to gain visibility and encourage broader community participation and discussion around fashion topics.

Evaluation Metrics. Finally, an experimentation engine manifest in chatbots provides direct measures of success. For example, measuring the number of followers or user engagement — "average time on bot" — amongst these prototypes can identify interactions and models that are more useful than others. Ultimately, if a subset of the chatbot applications go viral, this is a strong signal that it is worth the time and resources to develop them into stand-alone applications.

Ranjitha Kumar is an Assistant Professor in the Department of Computer Science at the University of Illinois at Urbana-Champaign. Her research has received best paper awards and nominations at premiere conferences in HCI, and been recognized by the machine learning community through invited papers at IJCAI and ICML. She received her PhD from the Computer Science Department at Stanford University in 2014. She was formerly the Chief Scientist at Apropose, Inc., a data-driven design company she founded, backed by Andreessen Horowitz and New Enterprise Associates.

Kristen Vaccaro is a PhD student in the Department of Computer Science at the University of Illinois at Urbana-Champaign, advised by Ranjitha Kumar and Karrie Karahalios. She holds a BA in Physics from Reed College.

References

Alon, I.; Qi, M.; and Sadowski, R. J. 2001. Forecasting aggregate retail sales:: a comparison of artificial neural networks and traditional methods. *Journal of Retailing and Consumer Services* 8(3):147–156.

Au, K.-F.; Choi, T.-M.; and Yu, Y. 2008. Fashion retail forecasting by evolutionary neural networks. *International Journal of Production Economics* 114(2):615–630.

Berg, T.; Berg, A.; and Shih, J. 2010. Automatic attribute discovery and characterization from noisy web data. In *Proc. ECCV*.

Di, W.; Wah, C.; Bhardwaj, A.; Piramuthu, R.; and Sundaresan, N. 2013. Style finder: fine-grained clothing style detection and retrieval. In *Proc. CVPR*.

Donath, J. 2007. Signals in social supernets. *Journal of Computer-Mediated Communication* 13(1):231–251.

English, B. L. 2007. *A cultural history of fashion in the 20th century: from the catwalk to the sidewalk*. Berg Publishers.

Grammer, K.; Renninger, L.; and Fischer, B. 2004. Disco clothing, female sexual motivation, and relationship status: is she dressed to impress? *Journal of sex research* 41(1):66–74.

He, R., and McAuley, J. 2015. Ups and downs: modeling the visual evolution of fashion trends with one-class collaborative filtering. In *Proc. WWW*.

Hidayati, S. C.; Hua, K.-L.; Cheng, W.-H.; and Sun, S.-W. 2014. What are the fashion trends in new york? In *Proc. MM*.

Huang, T.-H. K.; Lasecki, W. S.; Azaria, A.; and Bigham, J. P. 2016. "Is there anything else I can help you with?": Challenges in Deploying an On-Demand Crowd-Powered Conversational Agent. In *Proc. HCOMP*.

Kiapour, M.; Yamaguchi, K.; Berg, A.; and Berg, T. 2014. Hipster wars: discovering elements of fashion styles. In *Proc. ECCV*.

Kovashka, A.; Parikh, D.; and Grauman, K. 2012. Whittlesearch: image search with relative attribute feedback. In *Proc. CVPR*.

Kwak, I. S.; Murillo, A. C.; Belhumeur, P.; Belongie, S.; and Kriegman, D. 2013. From bikers to surfers: visual recognition of urban tribes. In *Proc. BMVC*.

Lin, Y.; Xu, H.; Zhou, Y.; and Lee, W.-C. 2015. Styles in the fashion social network: an analysis on lookbook.nu. In *Social Computing, Behavioral-Cultural Modeling, and Prediction*. Springer International Publishing.

Lin, Y.; Zhou, Y.; and Xu, H. 2015. Text-generated fashion influence model: an empirical study on style.com. In *Proc. HICSS*.

Liu, S.; Feng, J.; Song, Z.; Zhang, T.; Lu, H.; Xu, C.; and Yan, S. 2012a. Hi, magic closet, tell me what to wear! In *Proc. MM*.

Liu, S.; Song, Z.; Liu, G.; Xu, C.; Lu, H.; and Yan, S. 2012b. Street-to-shop: Cross-scenario clothing retrieval via parts alignment and auxiliary set. In *Proc. CVPR*.

McAuley, J.; Targett, C.; Shi, Q.; and van den Hengel, A. 2015. Image-based recommendations on styles and substitutes. In *Proc. SIGIR*.

McAuley, J. J.; Pandey, R.; and Leskovec, J. 2015. Inferring networks of substitutable and complementary products. In *Proc. KDD*.

Mortensen, D. 2016. Understanding the facebook and microsoft chatbot revolution. `http://bit.ly/2jvqpod`.

Murillo, A. C.; Kwak, I. S.; Bourdev, L.; Kriegman, D.; and Belongie, S. 2012. Urban tribes: analyzing group photos from a social perspective. In *Proc. CVPRW*.

Nelissen, R. M., and Meijers, M. H. 2011. Social benefits of luxury brands as costly signals of wealth and status. *Evolution and Human Behavior* 32(5):343–355.

Piacentini, M., and Mailer, G. 2004. Symbolic consumption in teenagers' clothing choices. *Journal of consumer Behaviour* 3(3):251–262.

Radziszewski, A. 2016. Three challenges you?re going to face when building a chatbot. `http://bit.ly/2j0nYJu`.

Rector, C. 2014. How to Create a Capsule Wardrobe (And why it will change your life). *The Everygirl*.

Sharon, E. 2017. Mode ai. `http://mode.ai/`.

Shen, E.; Lieberman, H.; and Lam, F. 2007. What am I gonna wear?: scenario-oriented recommendation. In *Proc. IUI*.

Simo-Serra, E., and Ishikawa, H. 2016. Fashion style in 128 floats: joint ranking and classification using weak data for feature extraction. In *Proc. CVPR*.

Simo-Serra, E.; Fidler, S.; Moreno-Noguer, F.; and Urtasun, R. 2015. Neuroaesthetics in fashion: modeling the perception of fashionability. In *Proc. CVPR*.

Song, Z.; Wang, M.; Hua, X.-S.; and Yan, S. 2011. Predicting occupation via human clothing and contexts. In *Proc. ICCV*.

Trufelman, A. 2016. The Trend Forecast. `http://99percentinvisible.org/episode/the-trend-forecast/`.

Tsujita, H.; Tsukada, K.; Kambara, K.; and Siio, I. 2010. Complete fashion coordinator: a support system for capturing and selecting daily clothes with social networks. In *Proc. AVI*.

Vaccaro, K.; Shivakumar, S.; Ding, Z.; Karahalios, K.; and Kumar, R. 2016. The elements of fashion style. In *Proc. UIST*.

Vartak, M., and Madden, S. 2013. CHIC: a combination-based recommendation system. In *Proc. SIGMOD*.

Veit, A.; Kovacs, B.; Bell, S.; McAuley, J.; Bala, K.; and Belongie, S. 2015. Learning visual clothing style with heterogeneous dyadic co-occurrences. In *Proc. ICCV*.

Vittayakorn, S.; Yamaguchi, K.; Berg, A.; and Berg, T. 2015. Runway to realway: visual analysis of fashion. In *Proc. WACV*.

Yamaguchi, K.; Kiapour, M. H.; and Berg, T. 2013. Paper doll parsing: retrieving similar styles to parse clothing items. In *Proc. ICCV*.

Yu, L.-F.; Yeung, S.-K.; Terzopoulos, D.; and Chan, T. F. 2012. Dressup! outfit synthesis through automatic optimization. In *Proc. SIGGRAPH Asia*.

Zarroli, J. 2013. In Trendy World Of Fast Fashion, Styles Aren't Made To Last. *NPR*.