In the online advertising industry, the process of designing an ad creative (i.e., ad text and image) is mostly a manual process. Typically, each advertiser launches multiple creatives in an A/B testing setup to learn which creatives do better for the target audience, and then refines the creatives in an iterative manner. Due to the manual nature of this process, learning from past creatives, and designing new creatives is time consuming. In this talk, we will cover our recent work on automating the creative design process. This includes: (i) multimodal understanding to infer themes in successful ads, (ii) generation and ranking approaches for recommending ad text and images, and (iii) creative personalization vis-a-vis a consumer’s progress towards a purchase event. Our key contributions include: (i) transformer based cross-modal embeddings for understanding ads, (ii) encoder-decoder models with copy mechanism for ad text generation, and (iii) attention mechanisms for deeper user understanding (better creative personalization). Parts of our work have been accepted as papers at KDD 2019, RecSys 2019, The Web Conference (WWW) 2020, and CIKM 2020. This is joint work with Kapil Thadani, Manisha Verma, Narayan Bhamidipati, and Jelena Gligorijevic at Yahoo Research, and Yichao Zhou and Wei Wang at UCLA.