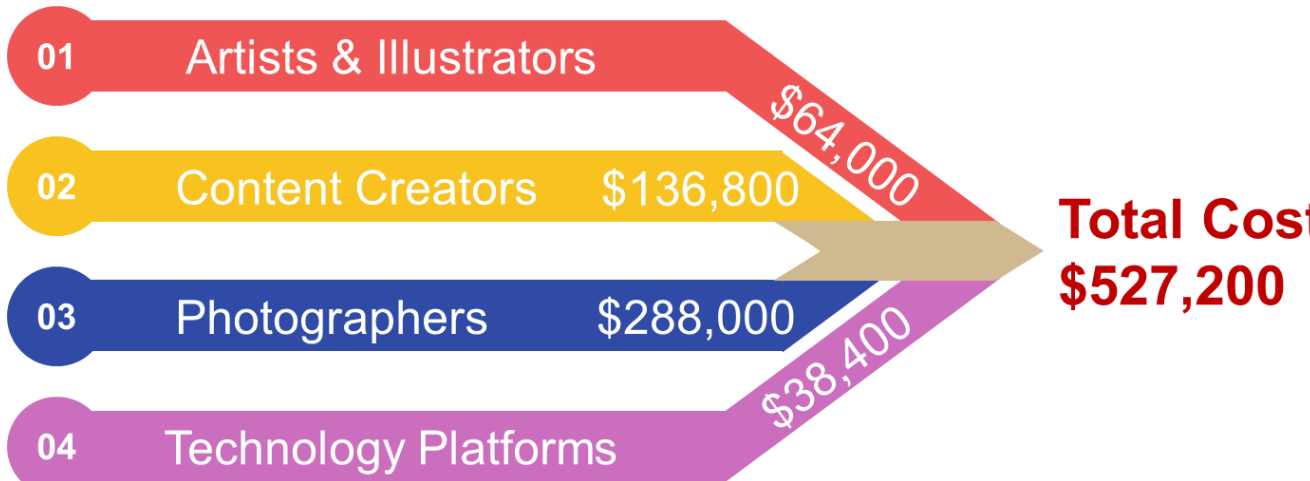


## BUSINESS PROBLEM

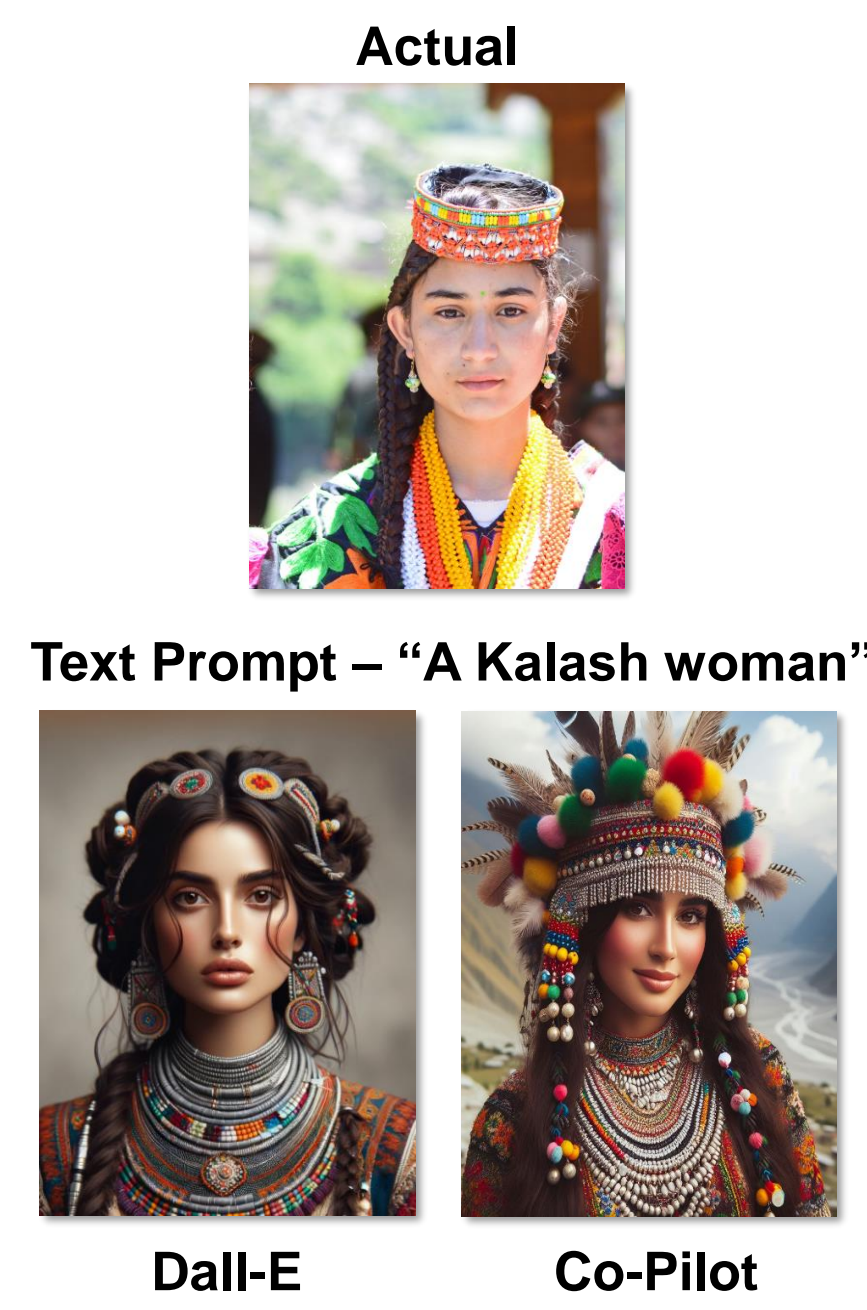
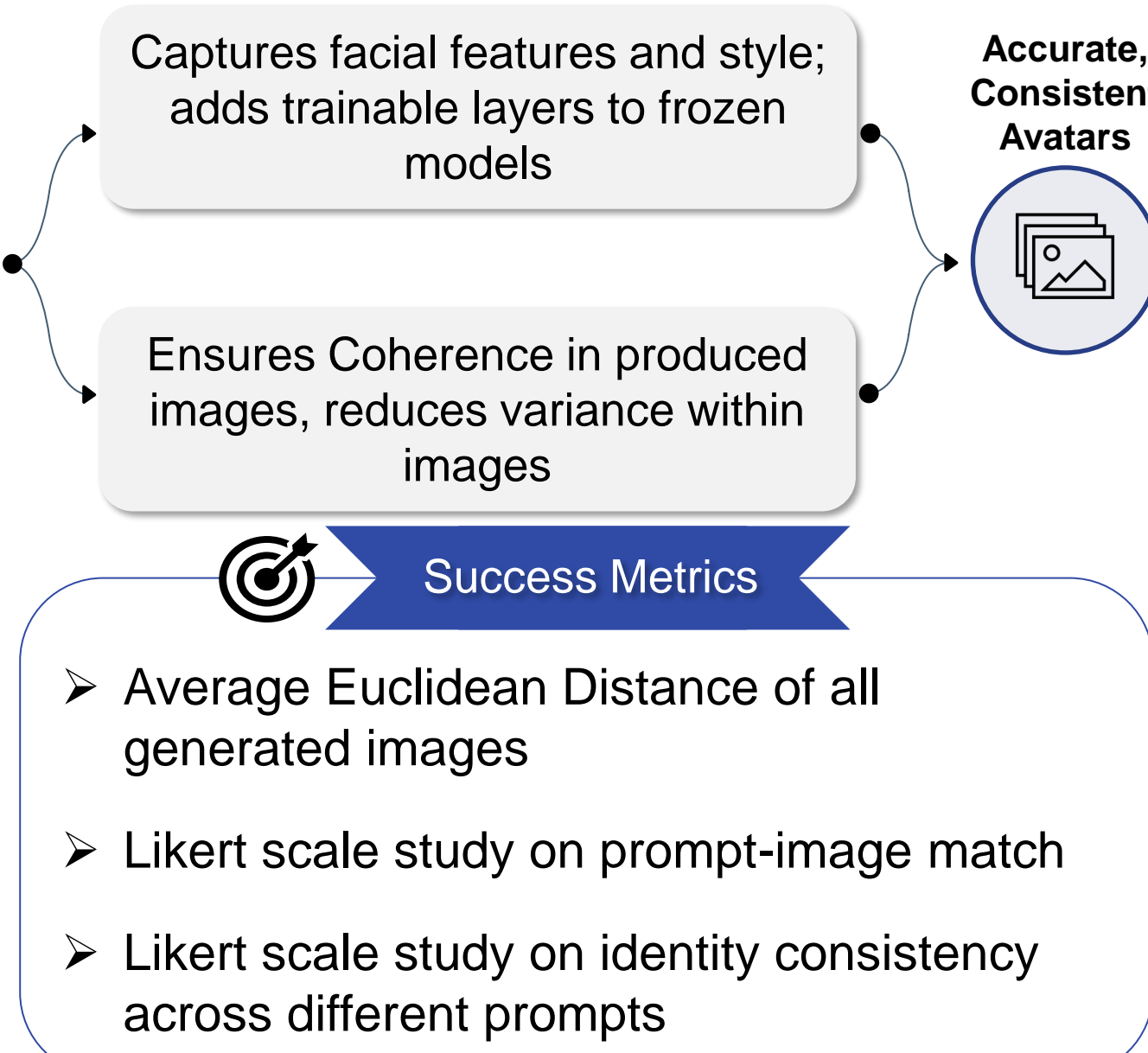
Current Generative AI models do not represent indigenous people accurately. They have a bias towards westerners which stems from the training data used for building these models. This can lead to severe underrepresentation and even misrepresentation of these 476 million people of the world. Covering one community out of the 5,000 indigenous communities using a conventional approach will cost approximately half a million dollars, including expenses for animators, photographers, studios, animation software, and travel time to these communities to create localized content.



This solution can be used for advertisements, educational content, public health messages, game characters, comic books, movie characters etc.

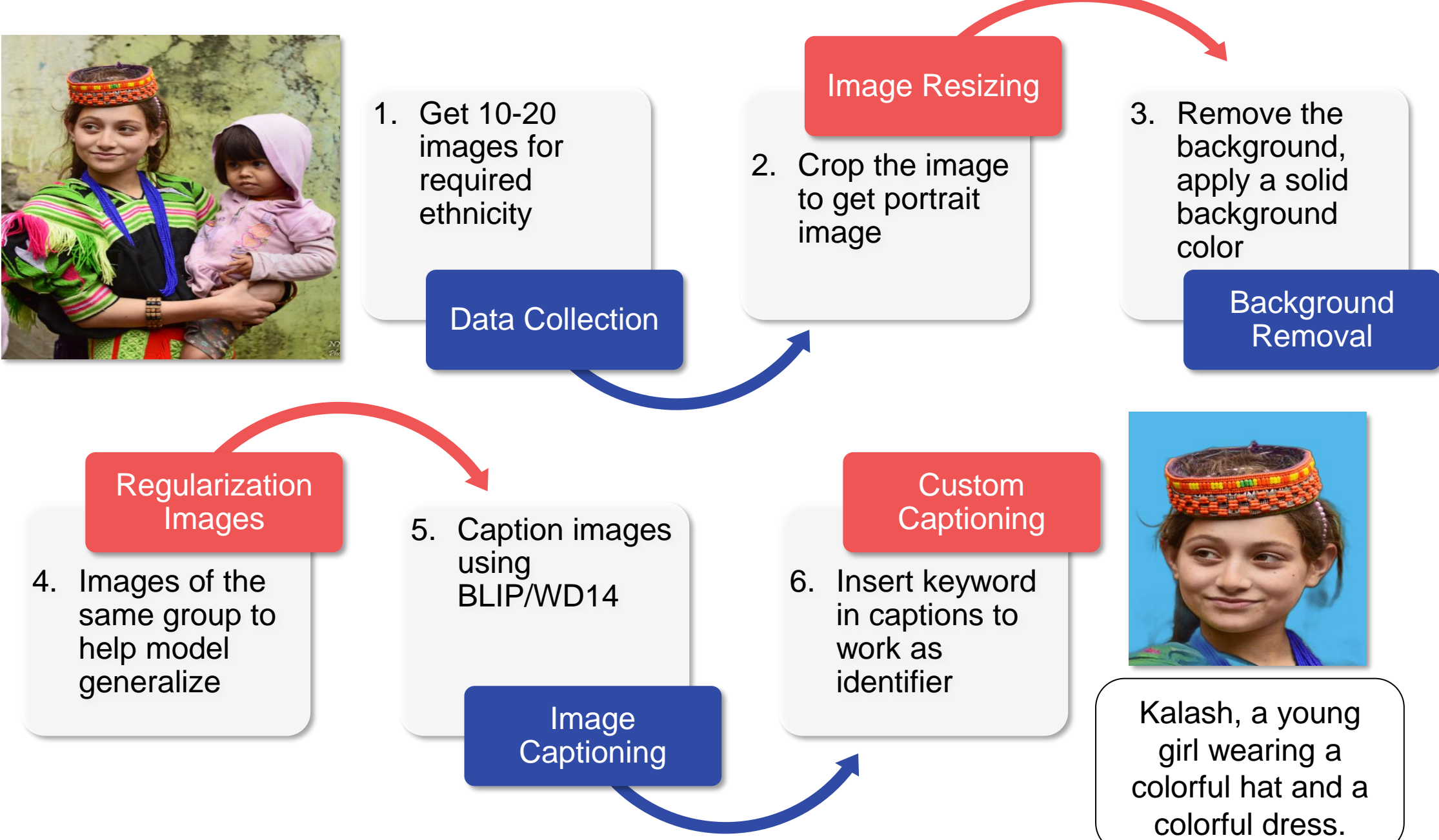
## ANALYTICS PROBLEM

Current text-to-image models transform text into graphics but exhibit inconsistency and biases, favoring Western individuals and celebrities, due to skewed training data. We aim to leverage diverse datasets and technologies like Low-Rank Approximation (LoRA) models, k-means clustering and interfaces like ComfyUI to develop a model that accurately and consistently represents diverse identities in AI-generated content.



A text-to-image model, trained with relevant data, can be utilized to produce tailored content for these communities at scale and efficiently. Such a solution can also create multiple opportunities for these communities.

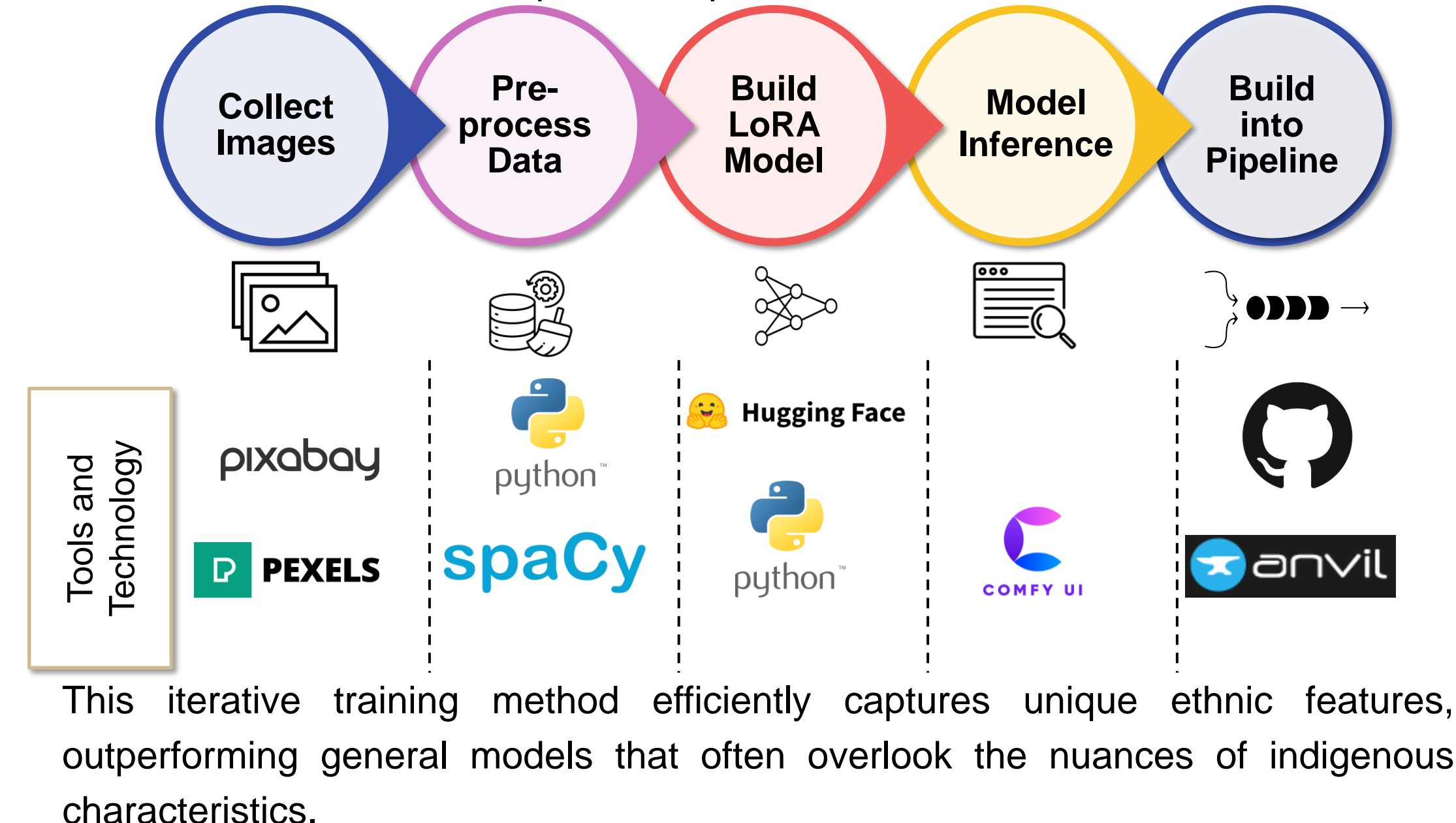
## DATA COLLECTION AND PREPROCESSING



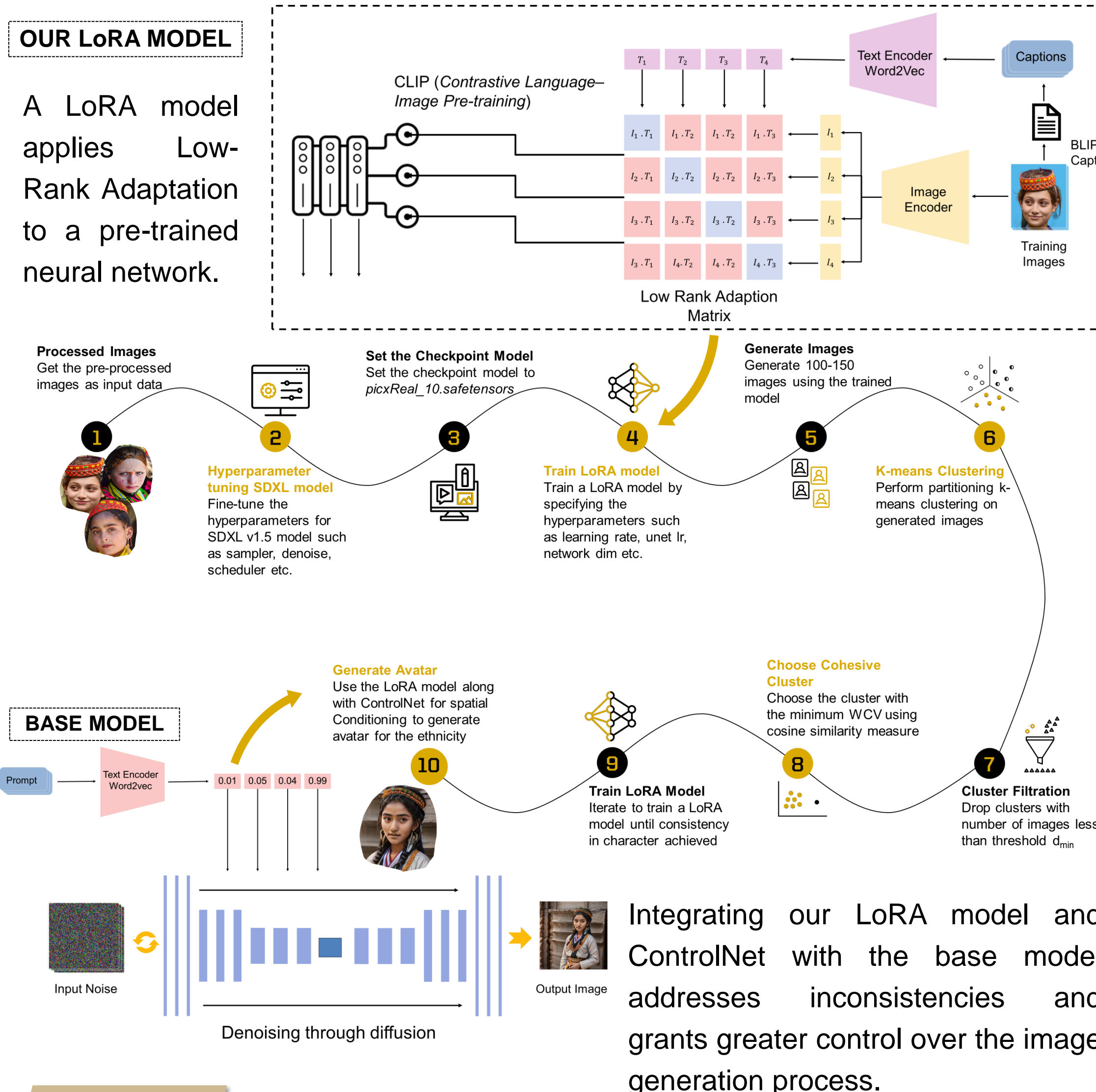
We gather 10 to 20 images per ethnicity from multiple sources, including online and private collections, covering ethnicities such as Kyrgyz, Maasai, Newar, among others. The dual-captioning approach aids in reinforcing the model's association between images and their respective ethnicities during training. Our thorough data curation and preprocessing efforts are aimed at developing a text-to-image model that accurately depicts diverse ethnic characteristics.

## METHODOLOGY

We utilize a base LoRA model trained on a small dataset of 10-15 pre-processed images for generating ethnically specific avatars, employing K-means clustering and ControlNet for consistent posture representation.

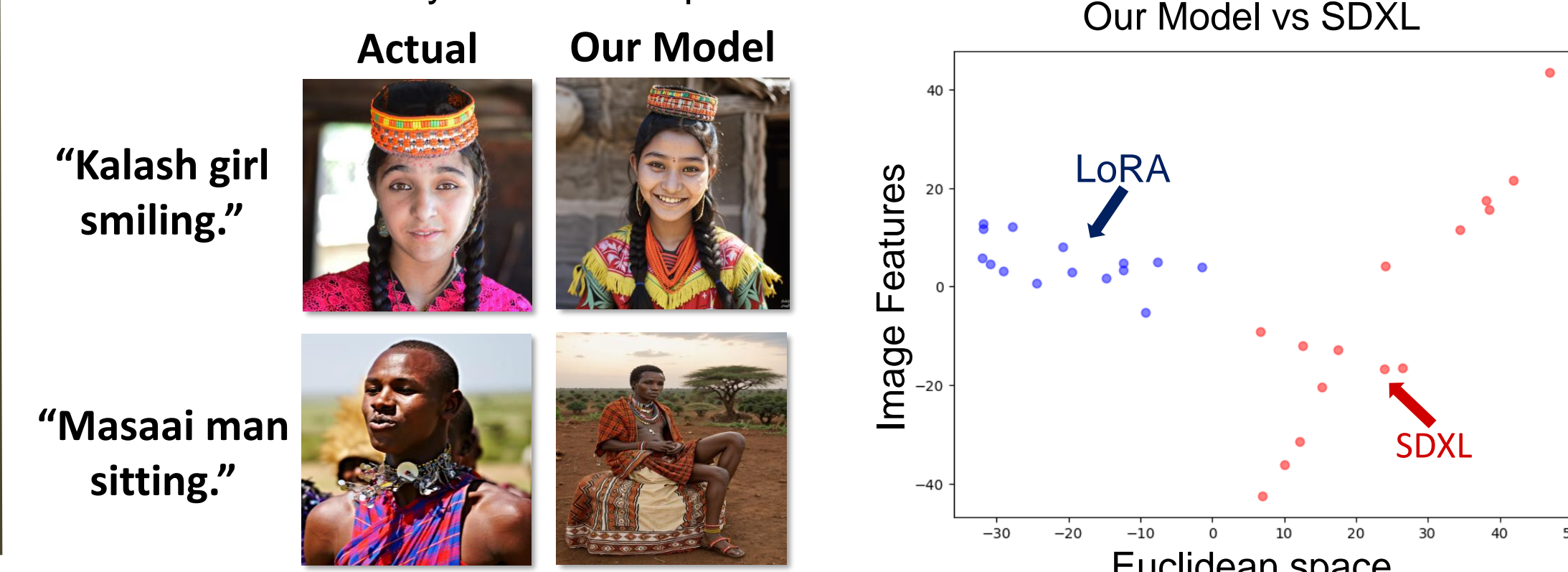


## MODEL BUILDING



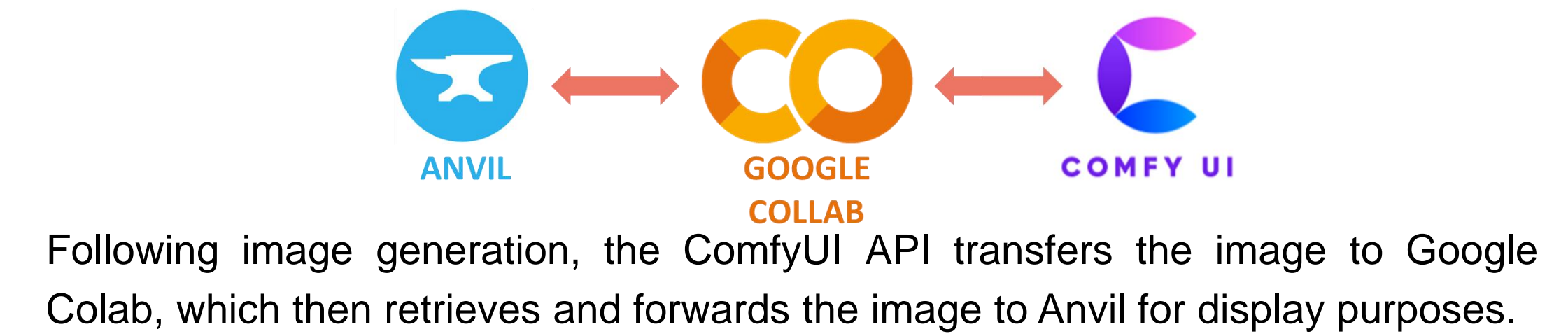
## RESULTS

Our model accurately reflects diverse ethnic facial and style characteristics, and faithfully captures details from textual prompts. The graphical representation indicates a higher clustering of images produced by our LoRA model, signifying enhanced consistency in feature representation.

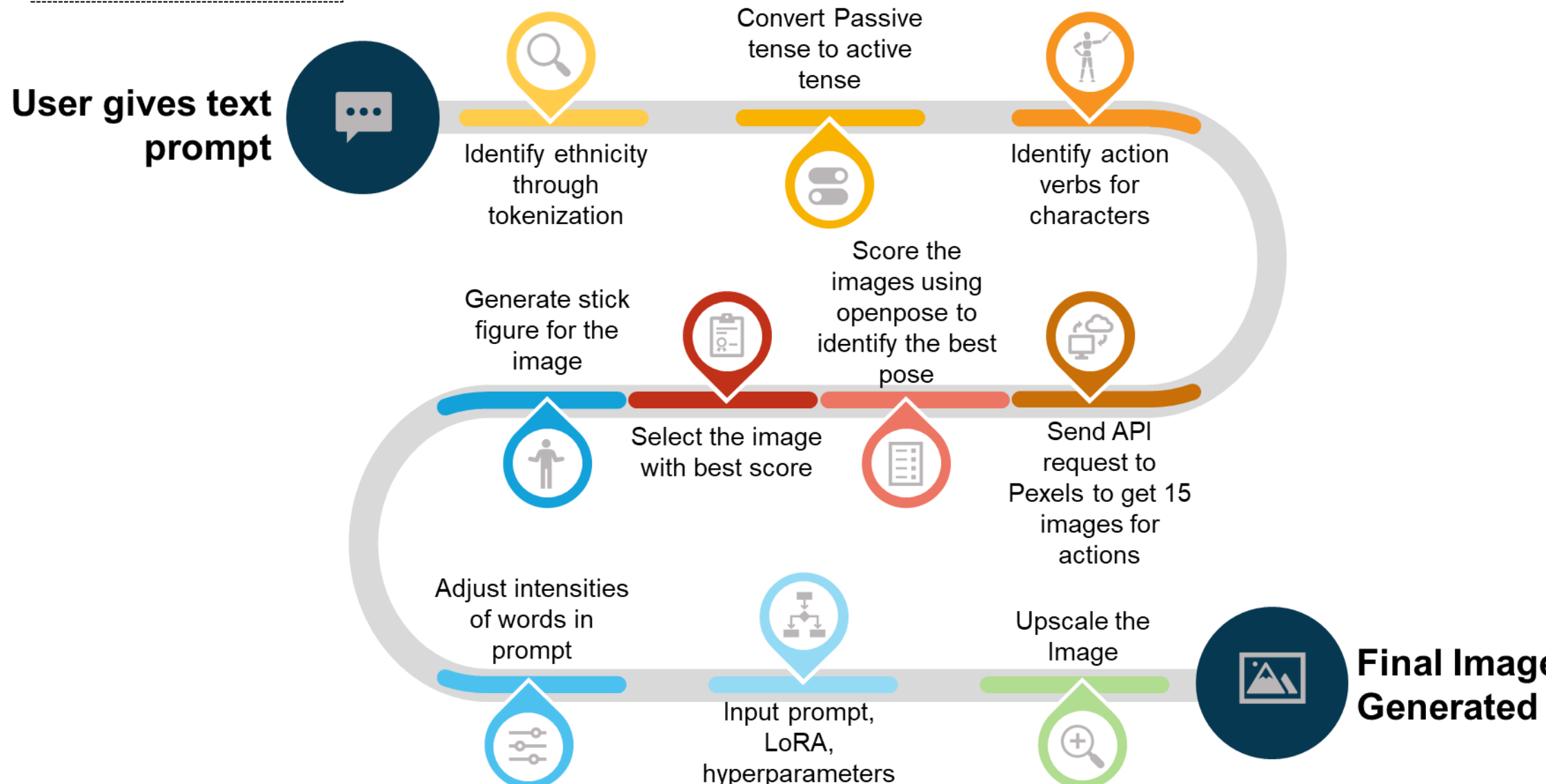


## DEPLOYMENT AND LIFECYCLE MANAGEMENT

**DEPLOYMENT PLATFORMS-** The user submits a prompt through Anvil, which is then securely transmitted to Google Colab via a REST API integration. Google Colab processes and formats the prompt before dispatching it, with the appropriate parameters, to the ComfyUI API. Finally, the ComfyUI API employs LoRA technology to generate the requested images.



## Image Delivery Pipeline



To facilitate a seamless transition and handover of technology, we have developed a comprehensive readme text file accompanying our Google Colab codes, detailing the entire process flow. Additionally, output images are archived in Google Drive for future accessibility and utilization.

## ACKNOWLEDGEMENTS

We thank our industry partner for their trust, support and encouragement while approaching this business problem. We also thank Professor Matthew Lanham for constant guidance on this project.

