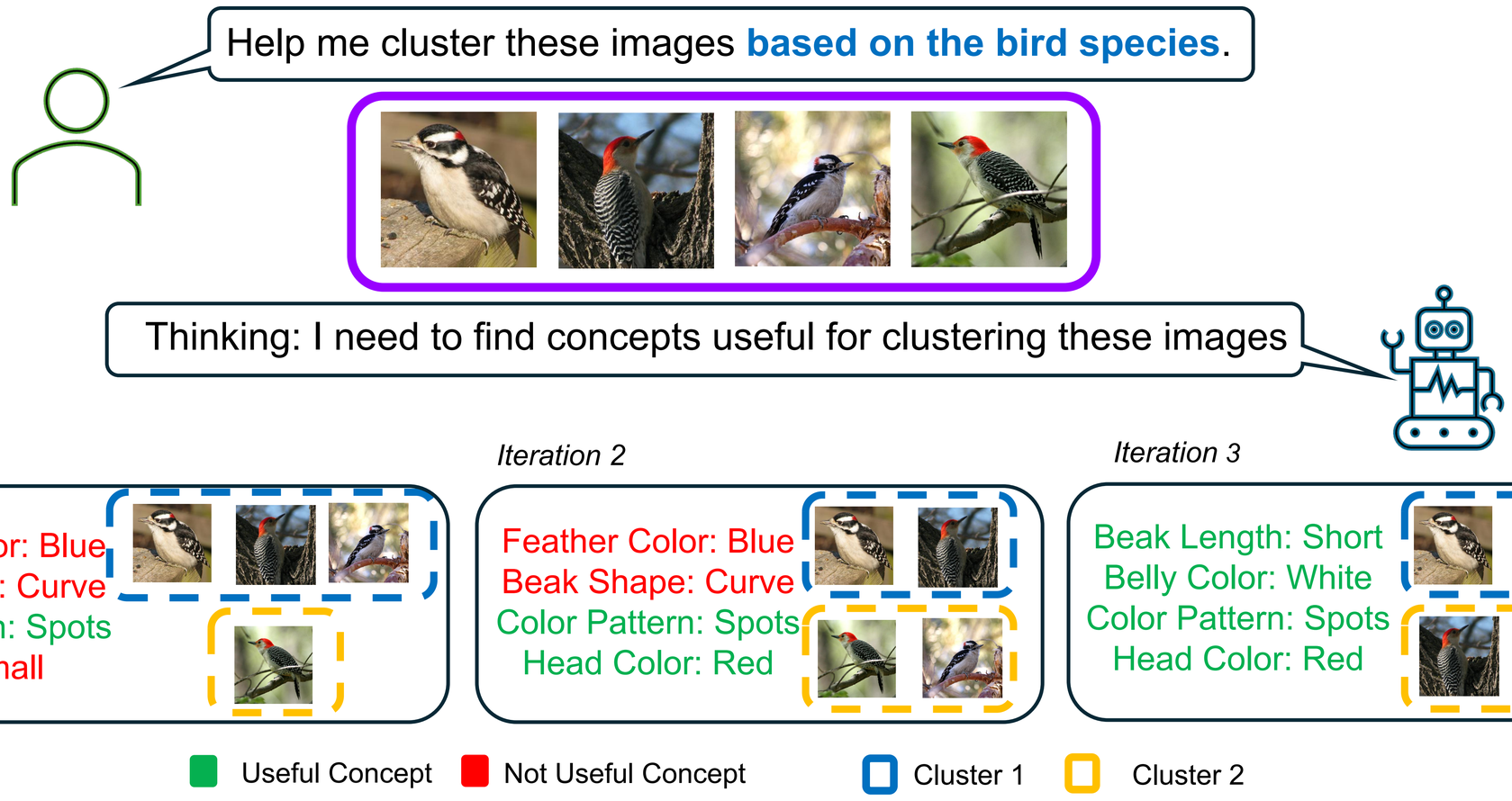
Interpretable Text-Guided Image Clustering via Iterative Search

♦ **Bingchen Zhao** Oisin Mac Aodha
University of Edinburgh

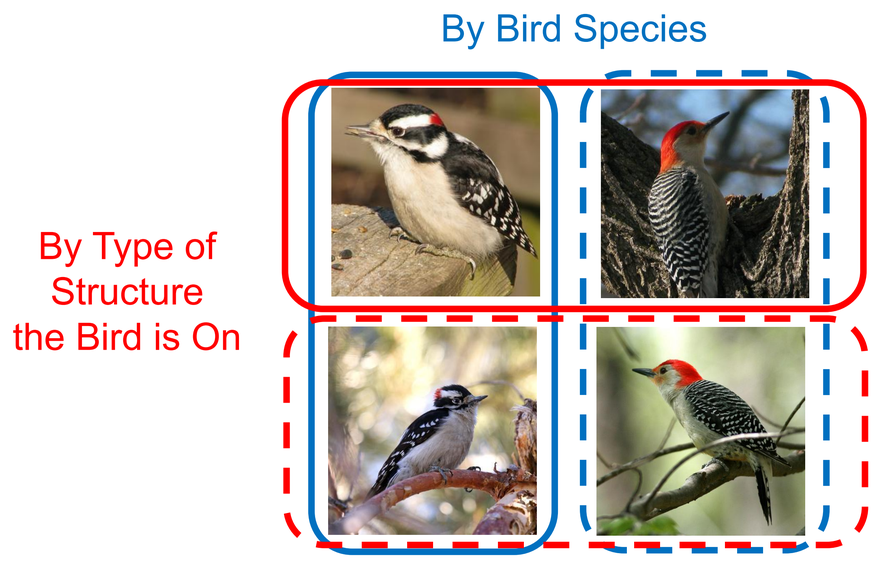
SCAN ME

Question: How to cluster **unlabeled** images to satisfy a user's **natural language query**?



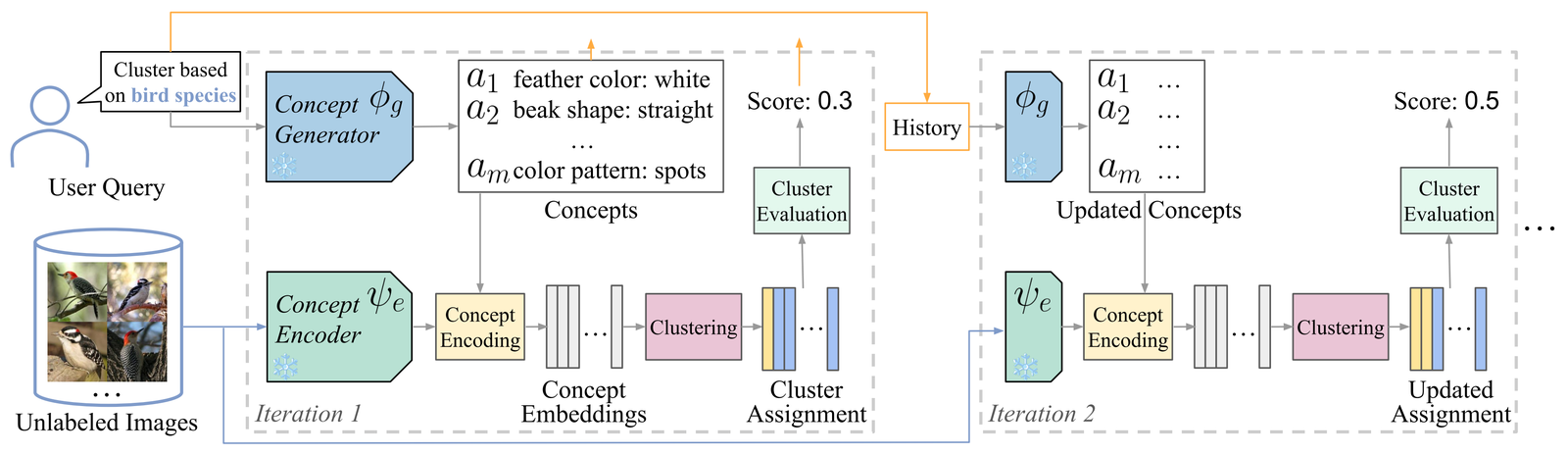
The Challenge

Unsupervised clustering is ill-posed



Our Approach

Iteratively refine visual concepts w. feedback from an unsupervised score, improving alignment to user query



Related Work

Multi-Criteria Clustering: Unsupervised Clustering for Multiple Clustering Results.

Cannot take user inputs.

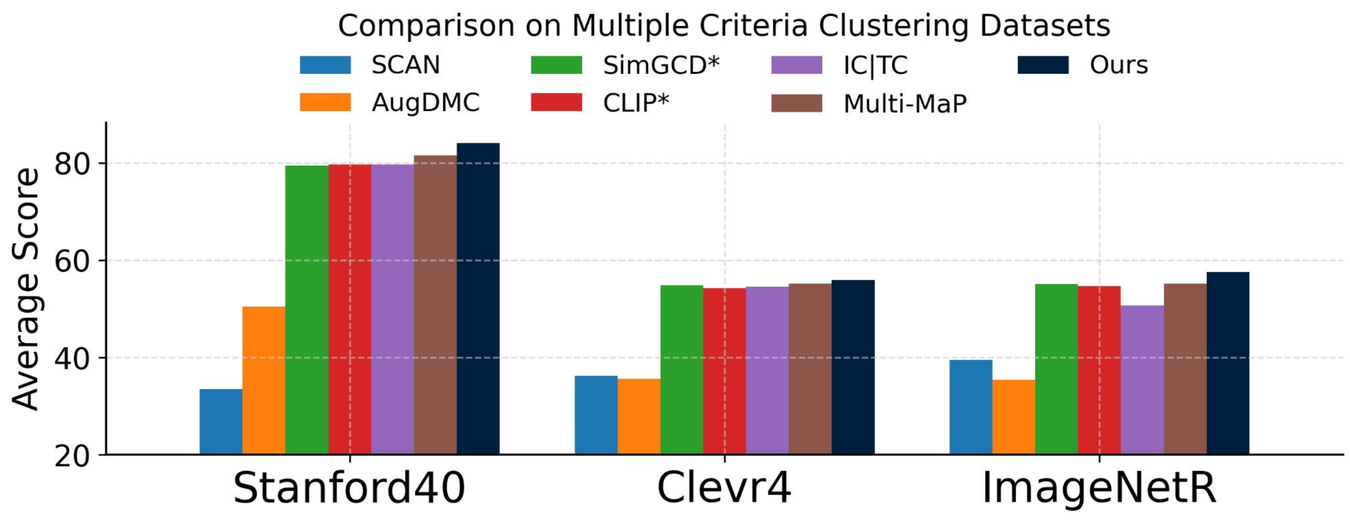
Category Discovery: Determining the Clustering Criteria by Labeled Data.

User inputs need to be complex (labeled data).

Text-guided Clustering: Determining the Clustering Criteria by Natural Language Input.

Often zero-shot or one-shot.

Evaluation



* indicates the use of supervision.