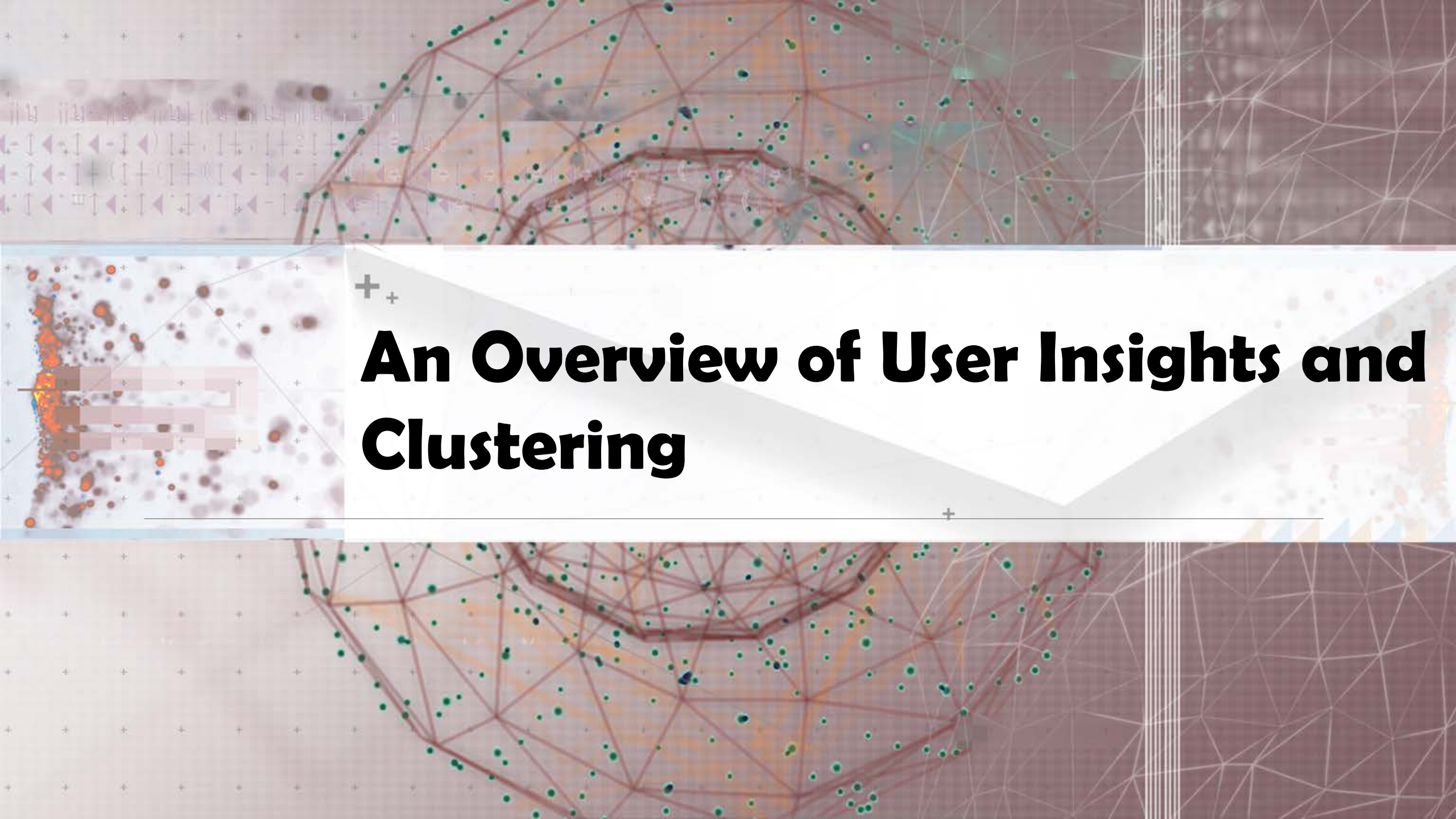

An Overview of User Insights and Clustering

User Insights and Interactions in Clustering

- **Visual insights**: One picture is worth a thousand words
  - Human eyes: High-speed processor linking with a rich knowledge-base
  - A human can provide intuitive insights; HD-eye: visualizing HD clusters
- **Semi-supervised insights**: Passing user's insights or intention to system
  - User-seeding: A user provides a number of labeled examples, approximately representing categories of interest
- **Multi-view and ensemble-based insights**
  - Multi-view clustering: Multiple clusterings represent different perspectives
  - Multiple clustering results can be ensembled to provide a more robust solution
- **Validation-based insights**: Evaluation of the quality of clusters generated
  - May use case studies, specific measures, or pre-existing labels

Recommended Readings

- Major Reference Books on Cluster Analysis
  - Jiawei Han, Micheline Kamber, and Jian Pei. Data Mining: Concepts and Techniques. Morgan Kaufmann, 3rd ed. , 2011 (Chapters 10 & 11)
  - Charu Aggarwal and Chandran K. Reddy (eds.). Data Clustering: Algorithms and Applications. CRC Press, 2014
  - Mohammed J. Zaki and Wagner Meira, Jr.. Data Mining and Analysis: Fundamental Concepts and Algorithms. Cambridge University Press, 2014
- Reference paper for this lecture
  - Charu Aggarwal. An Introduction to Clustering Analysis. *in* Aggarwal and Reddy (eds.). Data Clustering: Algorithms and Applications (Chapter 1). CRC Press, 2014