Power to the People: The Role of Humans in Interactive Machine Learning

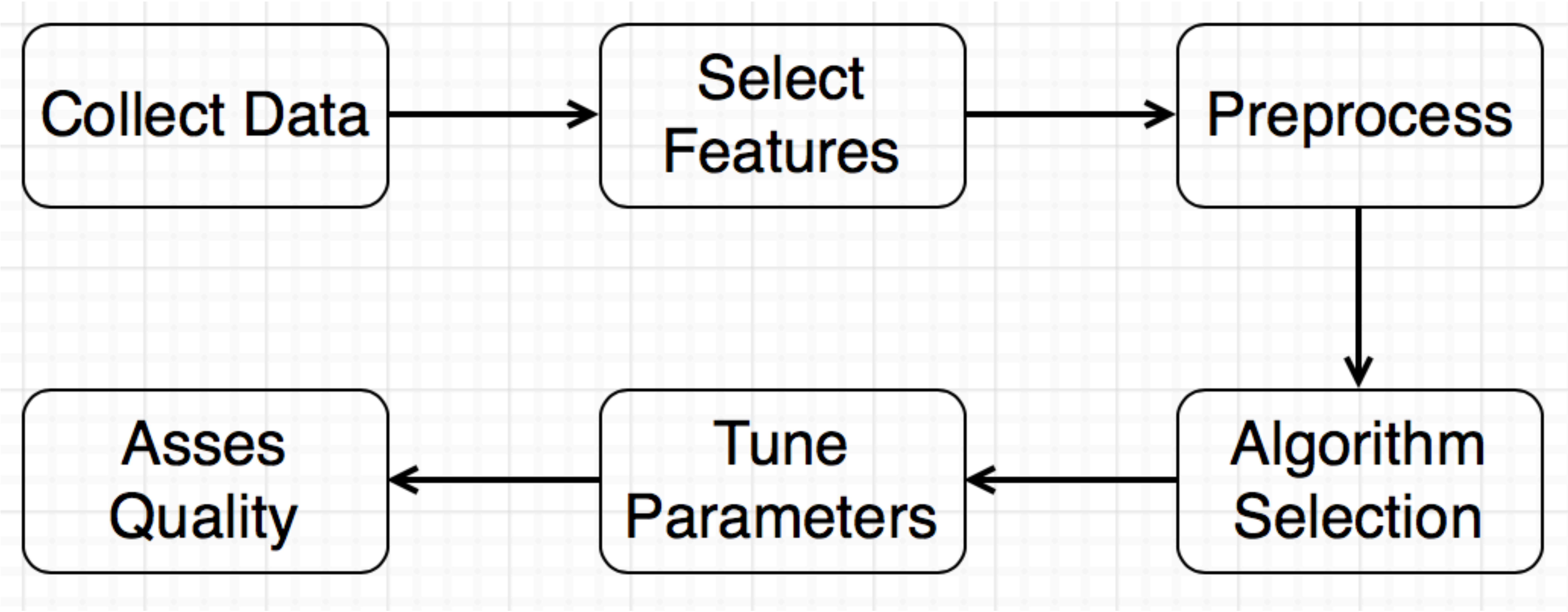
S Amershi, M Cakmak, WB Knox, T Kulesza

Presented by

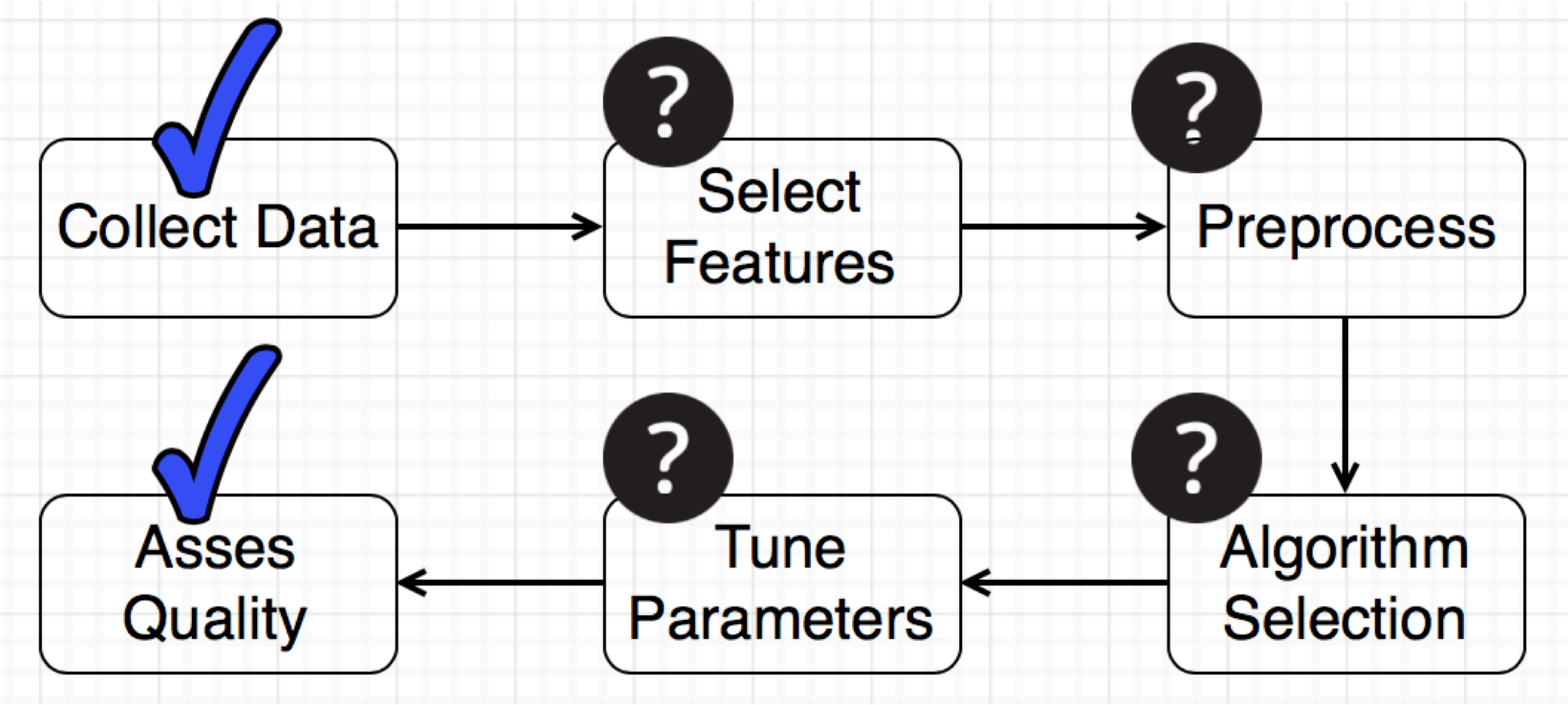
Saif Charaniya

November 23rd 2016

A Typical Approach to ML



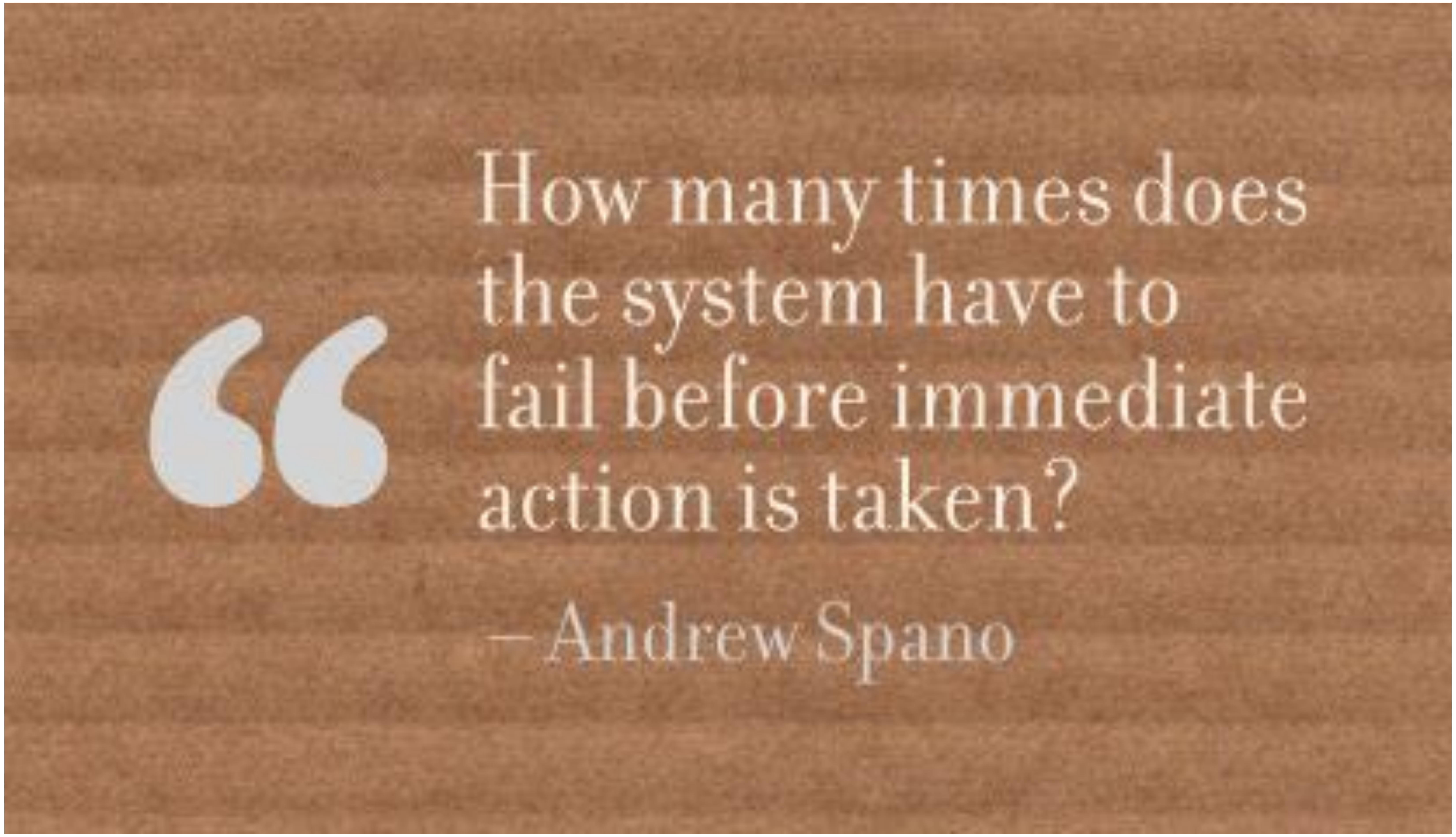
How End Users See the Process



Interactive Machine Learning

- Rapid, Focused, and Incremental!
- Allows users to explores the model space visually and interactively
- Reduces the need of supervision by ML experts
- Intelligent user interfaces and iML have been around for a over a decade (HooK 2000, Cohn 2003)

Rapid Updates



Taken from: http://quotespictures.com/wp-content/uploads/2013/03/how-many-times-does-the-system-have-to-fail-before-immediate-action-is-taken.jpg

Focused

Taken from: http://gym.westernsydney.edu.au/wp-content/uploads/2016/10/SmallSteps-01.png

Incremental

Taken from: https://cdn.shopify.com/s/files/1/0070/7032/files/The_10_Strategy.jpg?754

Discussion

- iML results in a tight coupling which leads to cross influence
- Studying user interaction can challenge assumptions of traditional learning systems
- End user interaction can be expanded into same aspects as of ML experts

Discussion Outline

1. User Interaction with iML
2. Interfaces for iML
3. Challenges in iML

User Interaction with iML

- People vs oracles
- Positive vs negative feedback
- People want to demonstrate how learners *should* behave
- People want to provide more than *just* data Labels
- People value transparency in learning systems
- Transparency can improve label quality

People vs Oracles

- Cakmak (2010): Pairs a robot with a person using three types of interaction
- Found people tend to underestimate the performance of Simon
- Found people want to control how the robots interaction

Simon the robot interacting with a human. Taken from: Cakmak, M., Chao, C., & Thomaz, A. L. 2010. Designing interactions for robot active learners. *Autonomous Mental Development, IEEE Transactions on*, *2*(2), 108-118.

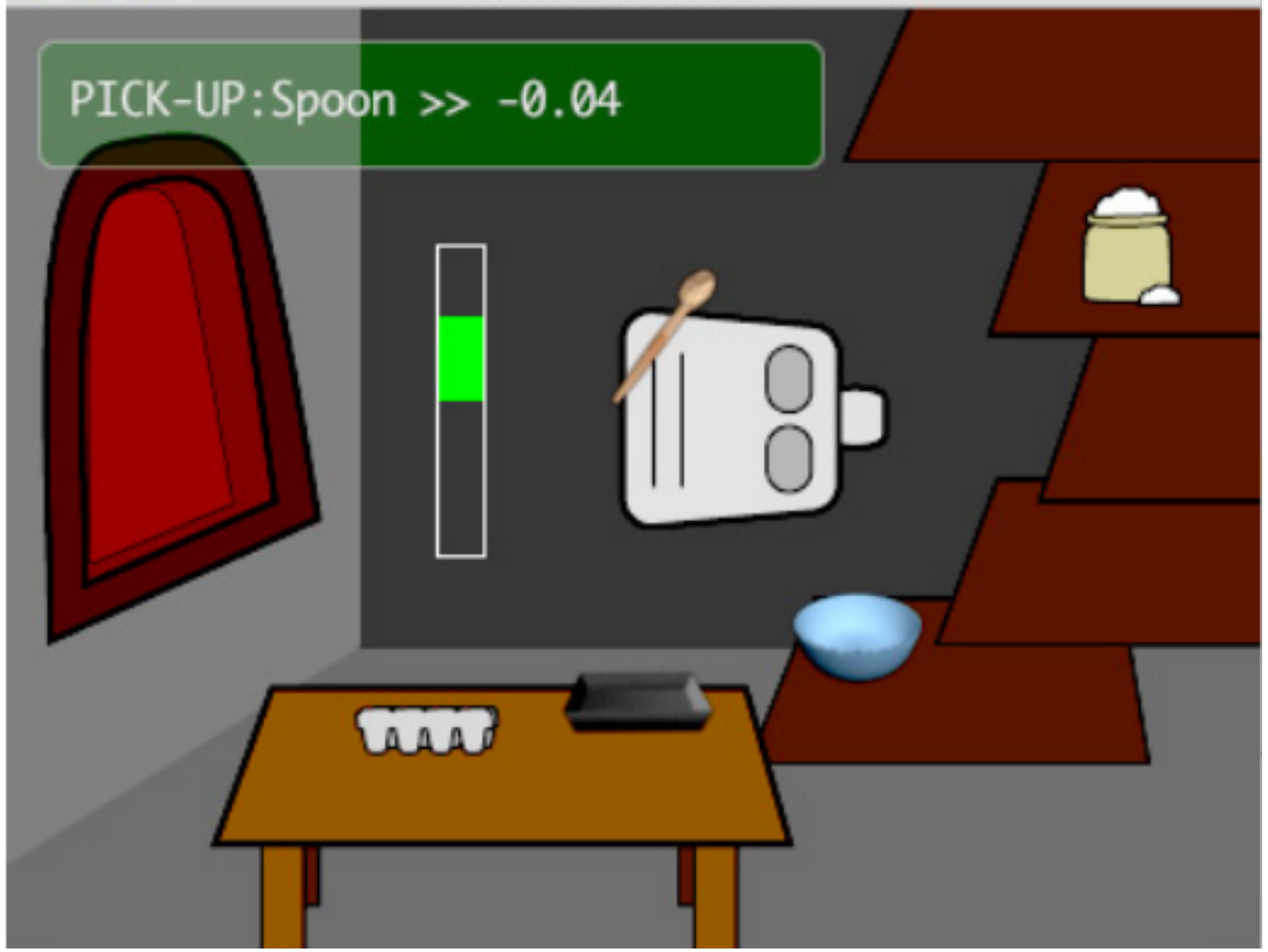
Positive vs Negative Feedback

- Thomaz + Breazeal (2008): Found people tend to give more positive feedback than negative feedback in episodic tasks
- Myopic algorithms don’t work well with this
- Knox + Stone (2013): Created VI Tamer using MDP. First to learn successfully non myopically from human generate reward!

Demonstrating how Learners Should Behave

- Thomaz + Breazeal (2008): Find people often violate rules of interaction with robots

- Human interaction can change the overall goal of learners



Sophie's Kitchen MDP. From: Thomaz, A. L., & Breazeal, C. 2008. Teachable robots: Understanding human teaching behavior to build more effective robot learners. *Artificial Intelligence*, *172*(6), 716-737.

Providing more than just Data Labels

- Stumpf (2007): Designed a text classification system and allowed people to provide feedback based on explanations of the system
- Showed humans want/like to provide feedback
- Can you think of other ways humans can help the ML process?

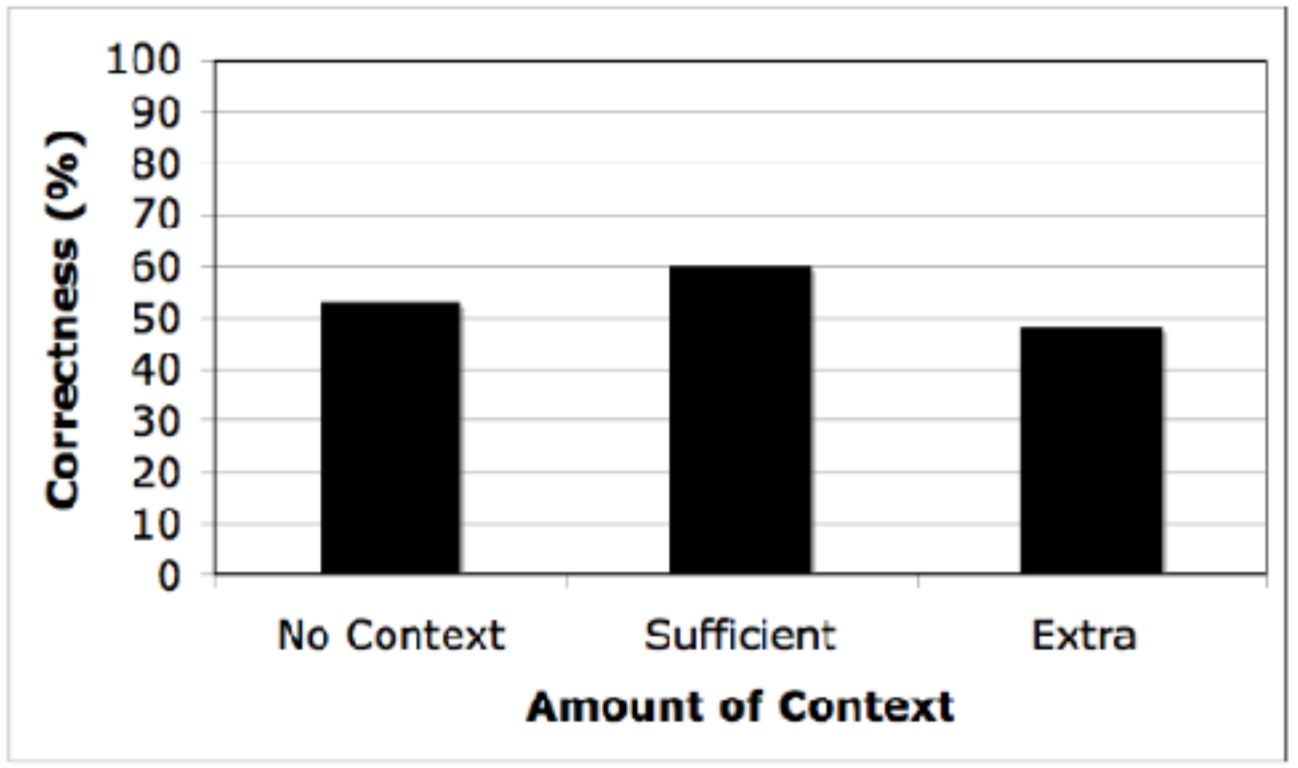
Valuing Transparency

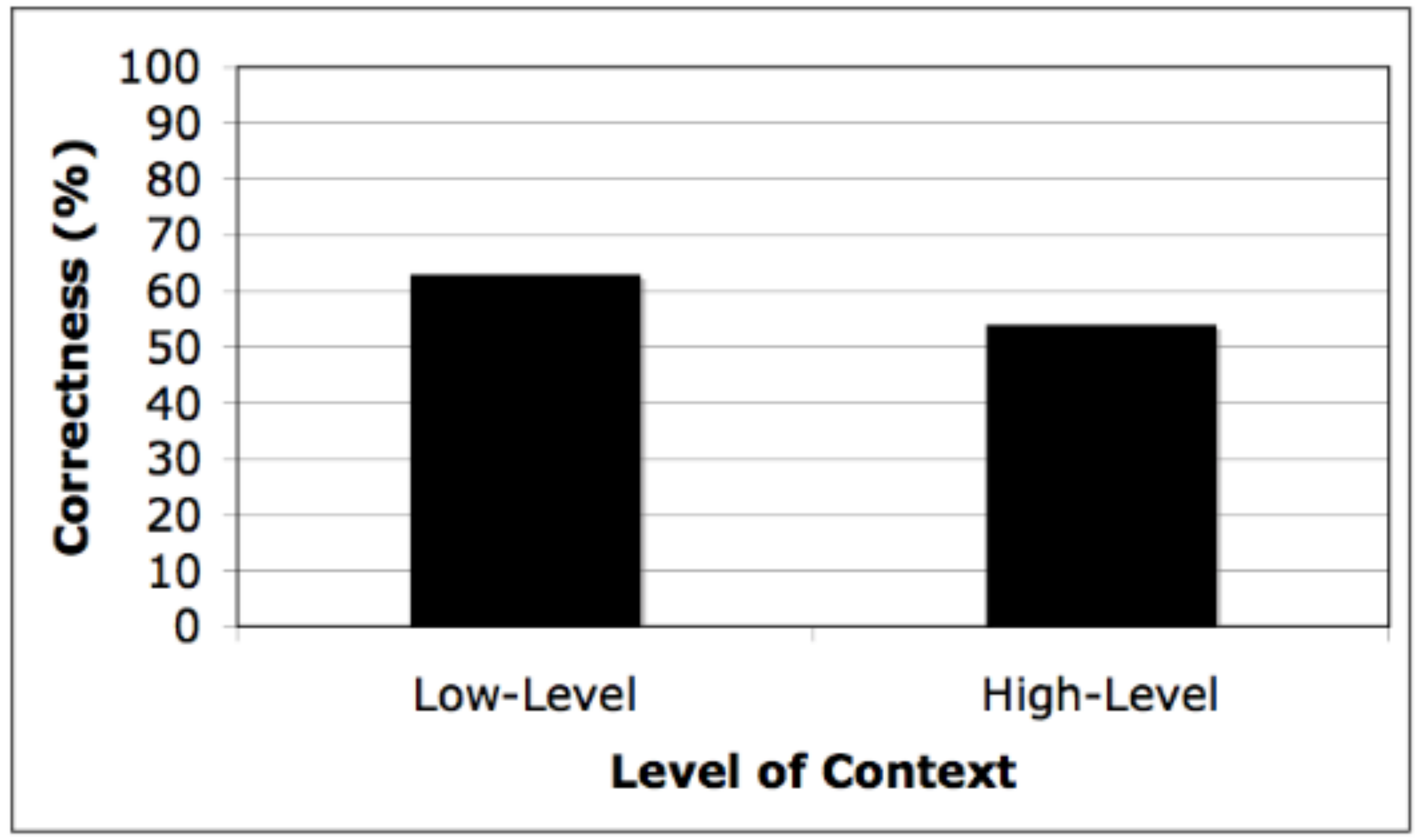
- Kulesza (2012): Explained to a group of user how a music app's recommender works and how user feedback in the app is used
- Found humans that had the explanation provided better feedback and were more satisfied with the app

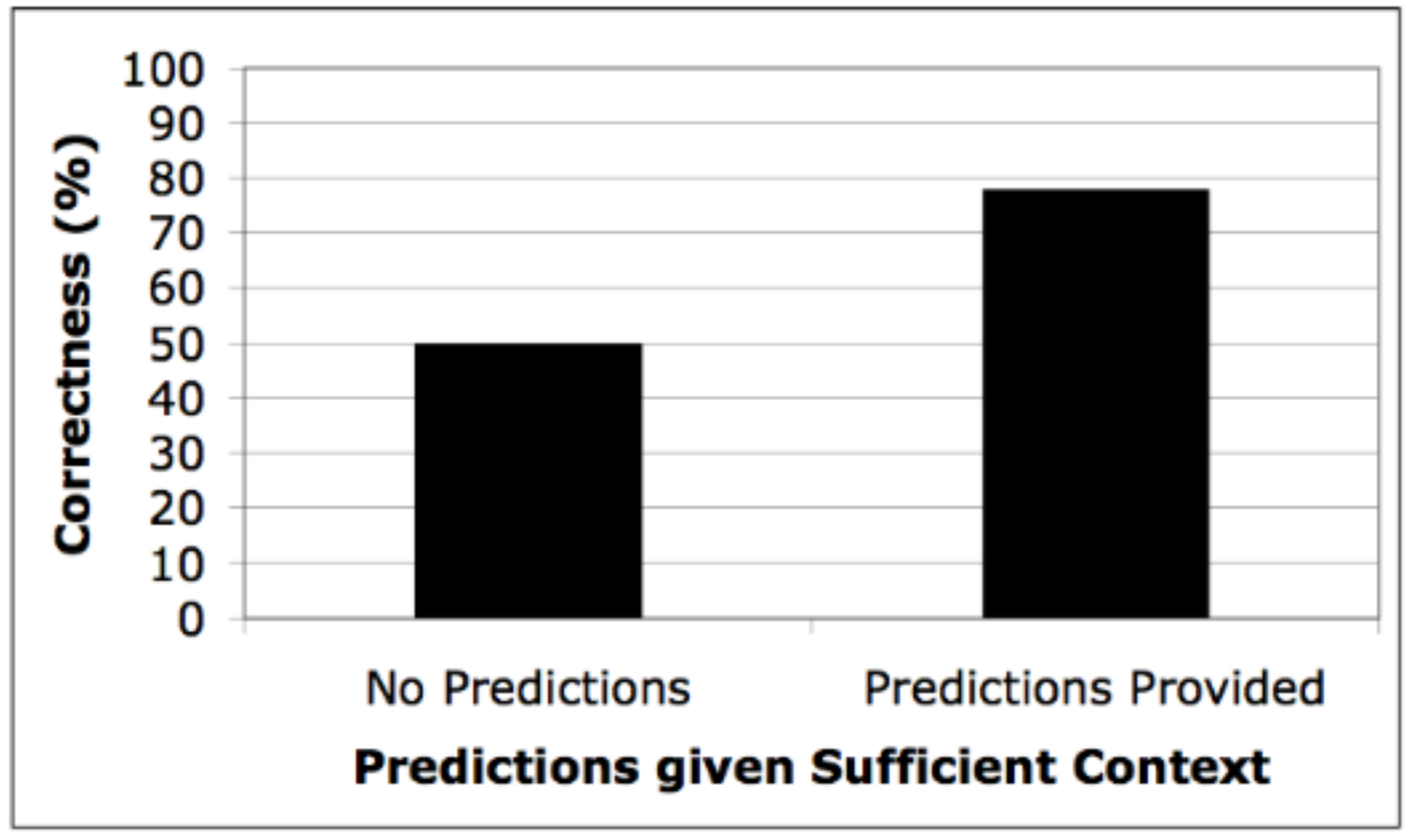
Improving Label Quality

- Rosenthal (2010): Studied how five additional features that may assist label processes

- Found that with **sufficient** context and prediction of answer, humans can provide better labels







Taken from: Rosenthal, S. L., & Dey, A. K. 2010. Towards maximizing the accuracy of human-labeled sensor data. In *Proceedings of the 15th international conference on Intelligent user interfaces* (pp. 259- 268). ACM.

Interfaces with iML

- Supporting assessment of model quality
- Supporting experimentation with model inputs
- Appropriately timing queries
- Enabling users to query the learner
- Enabling users to critique learner output
- Allowing users to specify preferences on errors
- Combining Models

Supporting Assessment of Model Quality

- Forgart (2010): created CueFlik which allows users to view information from both classes

- Amershi (2009): Found the best way is to show users high value examples with model summary helps users train better models

Supporting Experimentation with Model Input

- What if you did not have a backspace button on your keyboard?
- Amershi (2010): Expanded CueFlik to include an undo button with visualizations of user history
- Found users were able to create better models in the same amount of time
- Users are not perfect; users have expectations

Appropriately Timing Queries

- How do you ask a question? How would Simon do it?
- Users preferred teacher triggered queries => more control
- Economics of utility play an interesting role in human utilization

Enabling Users to Query the Learner

- Kulesza (2011): created a text classifier that would display statistics on features to a user, allowing a user to adjust features
- How can iML explain itself effectively for a user to provide feedback?

Enabling Critique of Learner Output

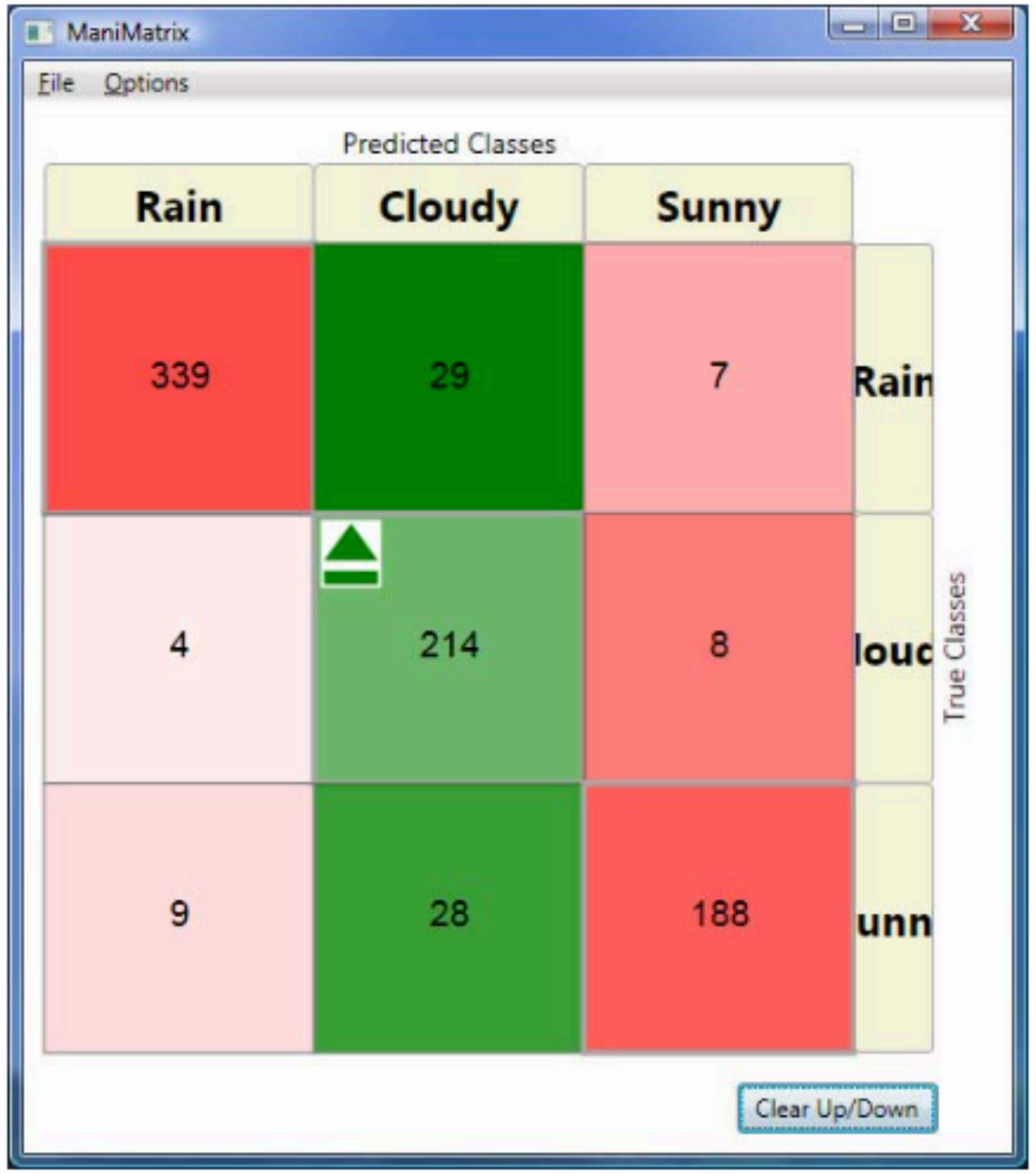
- Vig (2011) : studied this interaction using MovieLens to find similar items using KNN
- 89% of users liked the tool! 79% wanted it to become a permanent feature
- User attitude toward a learner can change when they are given interactive control

Allowing Users to Specify Preferences

- Precision, accuracy, recall. Which is more important?

- Kapoor (2010) created ManiMatrix allowing users to interactively adjust decision boundaries using the confusion matrix

- Allows non experts to change iML performance based on their needs



Taken from: Kapoor, A., Lee, B., Tan, D., & Horvitz, E. 2010. Interactive optimization for steering machine classification. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 1343-1352). ACM.

Combining Models

- How many ensemble methods are there?
- Talbot (2009): created EnsembleMatrix allowing users to interactively engage with multiple methods
- Allows visualization of building a model, evaluating, and exploring effects
- Combining human intuition with ML allowing users to create better classifiers faster!

Challenges in iML

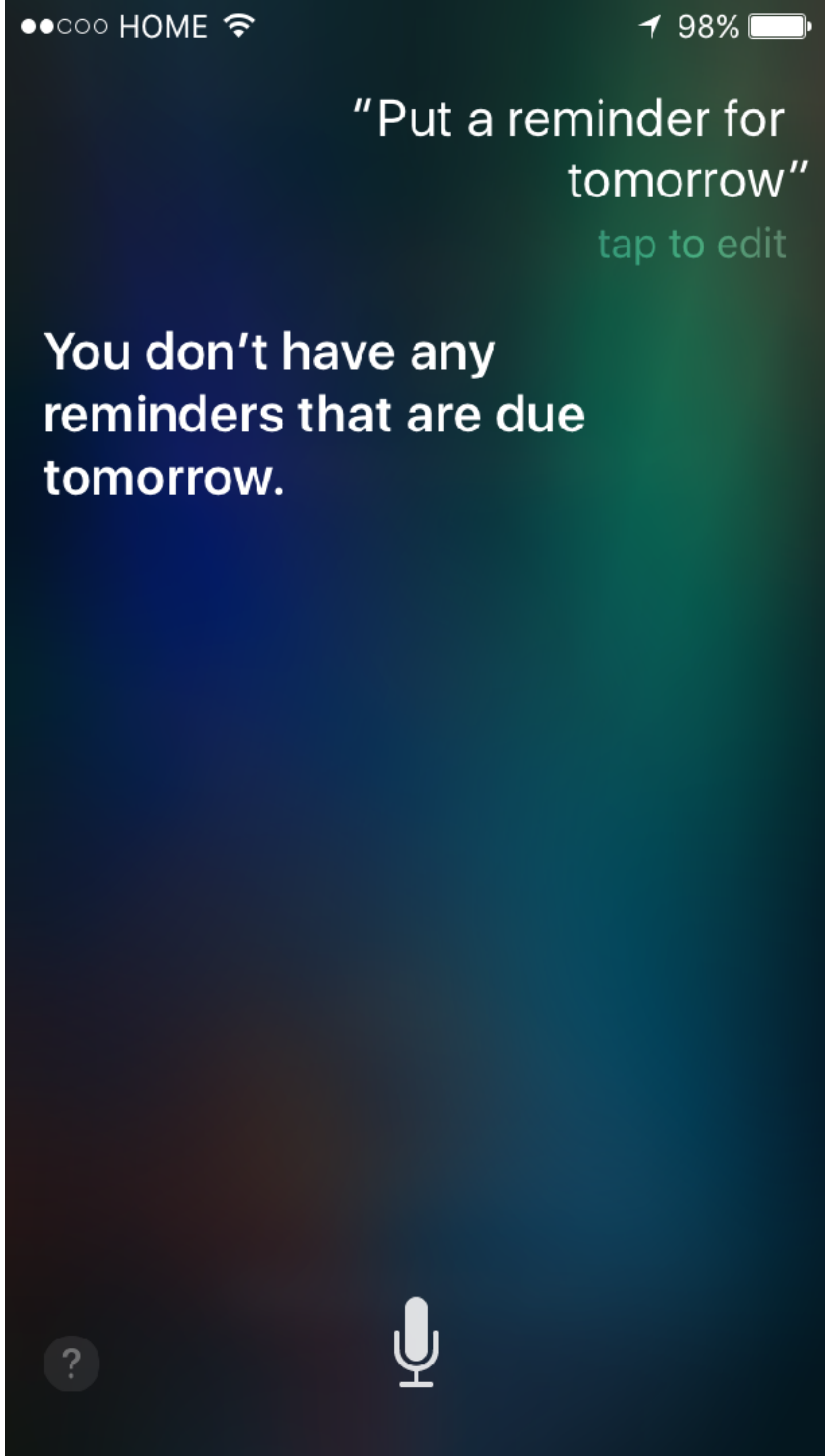
- Common languages across diverse fields
- principles and guidelines for *how* to design human interaction with ML
- Techniques and standards for evaluating iML systems
- Leveraging the *masses*
- Algorithmic problems in iML
- Collaboration across the fields of HCI and ML

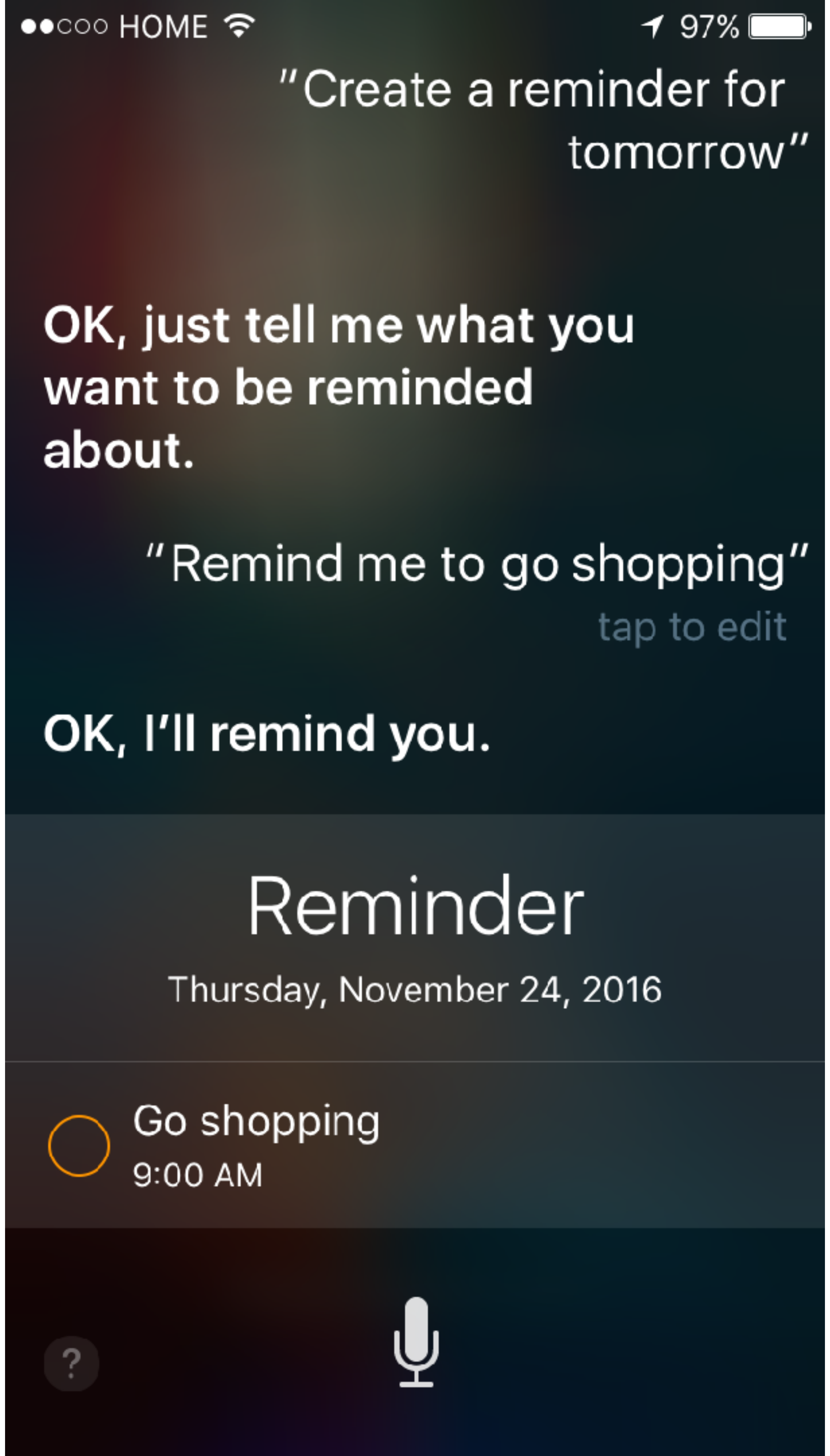
Common Languages

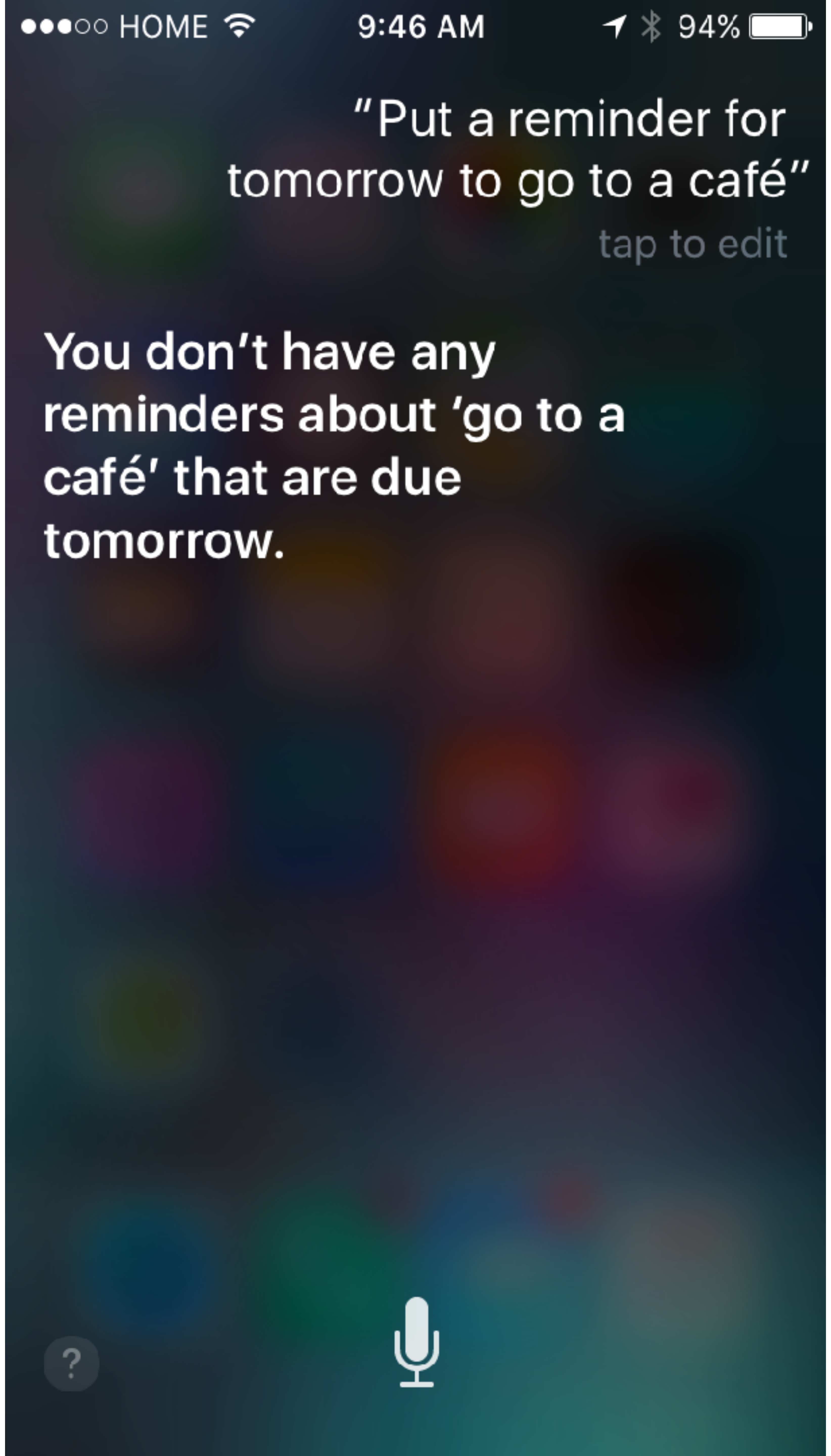
- What is iML? Relevance feedback? Socially guided ML? PBD?
- Impeding awareness and possibility of duplicate work
- Researchers need to look across diverse fields
- Porter (2013): breaks iML into 3 dimensions:
    - Task Decomposition
    - Training Vocabulary
    - Training Dialog

Principles and Guidlines

- iML systems do not always follow the principles of *understandability* and *actionability*
- Proposed guidelines include: safety and trust, managing expectations of users, and helping to avoid user frustration
- Experience in developing iML systems …
- Extracting and evaluating dimensions from research …


●●○○○ HOME
98%
"Put a reminder for tomorrow"
tap to edit
You don't have any reminders that are due tomorrow.
?


●●○○○ HOME
97%
"Create a reminder for tomorrow"
OK, just tell me what you want to be reminded about.
"Remind me to go shopping"
tap to edit
OK, I'll remind you.
Reminder
Thursday, November 24, 2016
Go shopping
9:00 AM
?


●●●○○ HOME
9:46 AM
94%
"Put a reminder for tomorrow to go to a café"
tap to edit
You don't have any reminders about 'go to a café' that are due tomorrow.
?

Evaluating iML Systems

- What does it mean for an iML system to fail? or succeed?
- How can we gauge effectiveness?
- Avoid generalizations of specific interaction techniques and instead generalize situations and contexts

Leveraging the *Masses*

- How can iML scale up from one user?
- Create systems that can integrate more users (Crowdsourcing ??)
- Are iML systems reusable? combinable? generalizable?
- iML needs Coordination!

Algorithmic Problems in iML

- What's more important: speed or accuracy?
- Do current iML algorithms allow natural interaction with users?

Collaboration in HCI with ML

- HCI: Human Computer Interaction
- HCI can help in evaluating iML systems with potential users
- Leveraging both solutions!

Examples of iML

- https://www.youtube.com/watch?v=JL-M-1utrlY
- https://vimeo.com/76664145

Conclusion

- You should have idea of what iML is and why it's awesome!
- You should understand the need of exploring user interaction with ML
- You should agree that there are many ways in which iML can harness human power and combine it with ML power
- iML will lead to more capable ML models and more capable end users

Suggested Resources

- https://www.youtube.com/user/SimonTheSocialRobot/videos
- https://www.youtube.com/watch?v=-2ggKevM-_8

References

- Amershi, S., Fogarty, J., Kapoor, A. and Tan, D. 2009. Overview-Based Example Selection in Mixed-Initiative Concept Learning. In Proceedings of the ACM Symposium on User Interface Software and Technology, 2009 (UIST 2009), pp. 247-256.

- Amershi, S., Fogarty, J., Kapoor, A. and Tan, D. 2010. Examining Multiple Potential Models in End-User Interactive Concept Learning. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, 2010 (CHI 2010), pp. 1357-1360.

- Cakmak, M., Chao, C., & Thomaz, A. L. 2010. Designing interactions for robot active learners. Autonomous Mental Development, IEEE Transactions on, 2(2), 108-118.

- Cakmak, M., & Thomaz, A. L. 2010. Optimality of human teachers for robot learners. In Development and Learning (ICDL), 2010 IEEE 9th International Conference on (pp. 64-69). IEEE.

References

- Fogarty, J., Tan, D., Kapoor, A., & Winder, S. 2008. CueFlik: interactive concept learning in image search. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (pp. 29-38). ACM.

- Kapoor, A., Lee, B., Tan, D., & Horvitz, E. 2010. Interactive optimization for steering machine classification. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (pp. 1343-1352). ACM.

- Knox, W. B., & Stone, P. 2013. Learning non-myopically from human-generated reward. In Proceedings of the 2013 International Conference on Intelligent User Interfaces (pp. 191- 202). ACM.

- Kulesza, T., Stumpf, S., Wong, W. K., Burnett, M. M., Perona, S., Ko, A., & Oberst, I. 2011. Why-oriented end-user debugging of naive Bayes text classification. ACM Transactions on Interactive Intelligent Systems (TiiS), 1(1), 2.

References

- Kulesza, T., Stumpf, S., Burnett, M., & Kwan, I. 2012. Tell me more?: the effects of mental model soundness on personalizing an intelligent agent. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (pp. 1-10). ACM.

- Porter, R., Theiler, J., & Hush, D. 2013. Interactive Machine Learning in Data Exploitation. Technical Report. Los Alamos National Lab.

- Rashid, A. M., Ling, K., Tassone, R. D., Resnick, P., Kraut, R., & Riedl, J. 2006. Motivating participation by displaying the value of contribution. In Proceedings of the SIGCHI conference on Human Factors in computing systems (pp. 955-958). ACM.

- Rosenthal, S. L., & Dey, A. K. 2010. Towards maximizing the accuracy of human-labeled sensor data. In Proceedings of the 15th international conference on Intelligent user interfaces (pp. 259- 268). ACM.

References

- Stumpf, S., Rajaram, V., Li, L., Burnett, M., Dietterich, T., Sullivan, E., Drummond, R., & Herlocker, J. 2007. Toward harnessing user feedback for machine learning. In Proceedings of the 12th international conference on Intelligent user interfaces (pp. 82-91). ACM.

- Talbot, J., Lee, B., Kapoor, A., & Tan, D. S. 2009. EnsembleMatrix: interactive visualization to support machine learning with multiple classifiers. In Proceedings of the 27th international conference on Human factors in computing systems (pp. 1283-1292). ACM.

- Thomaz, A. L., & Breazeal, C. 2008. Teachable robots: Understanding human teaching behavior to build more effective robot learners. Artificial Intelligence, 172(6), 716-737.

- Vig, J., Sen, S., & Riedl, J. 2011. Navigating the tag genome. In Proceedings of the 16th international conference on Intelligent user interfaces (pp. 93-102). ACM.

Thank You!