Durham University

COMP 3647
Human-AI Interaction Design

Topic 13:
Design Challenges and AI

Prof. Effie L-C Law

Overview

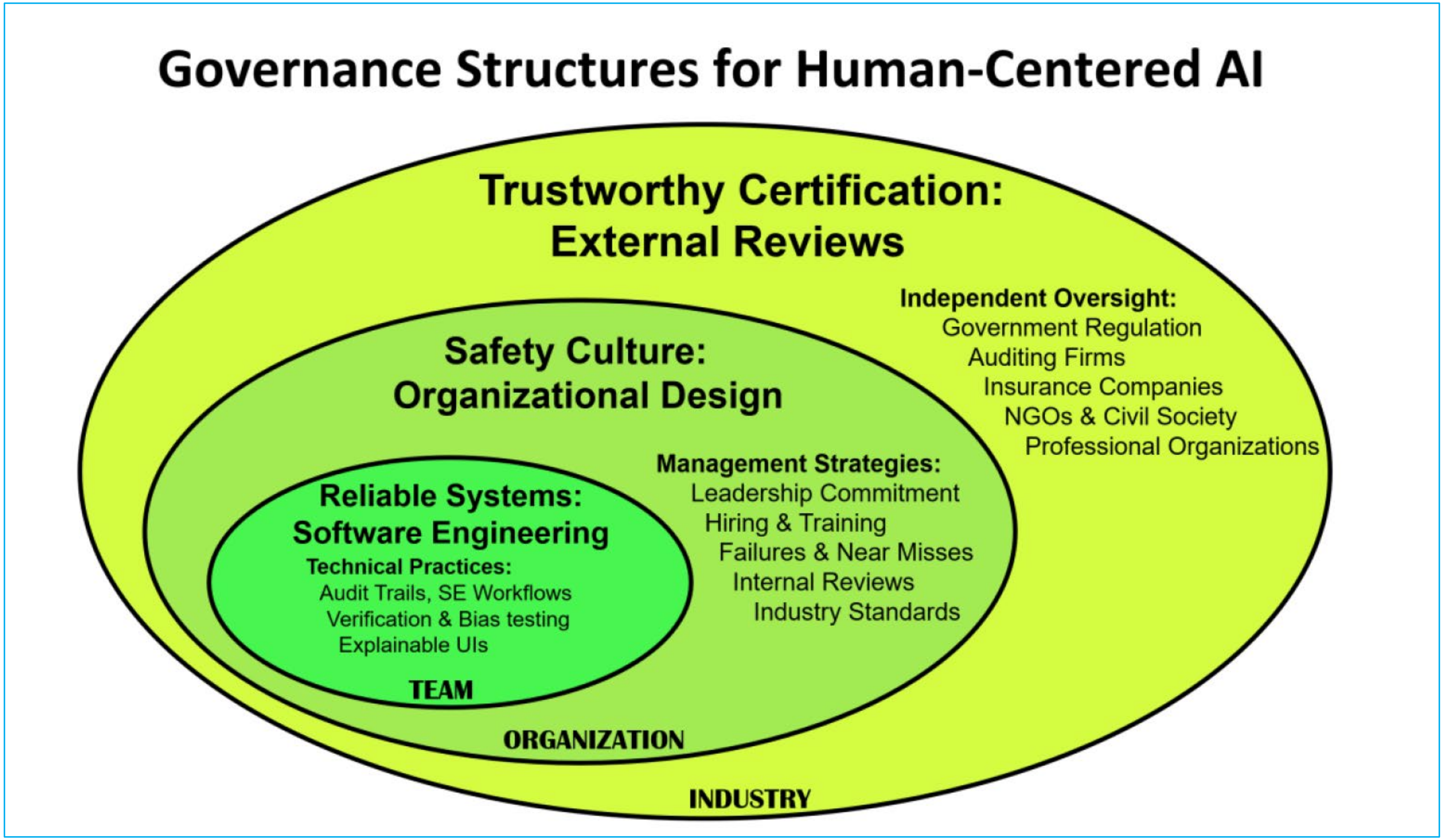
- Human-Centred AI
- AI as a Design Material
- Design Thinking & AI
- Design Methods & AI
- Some Case Studies

Human-Centred AI (Shneiderman, 2021)

- Build devices to **enhance and empower—not replace—humans**
- Autonomous machines should NOT **exceed or replace** any meaningful notion of human intelligence, creativity, and responsibility'
- Human-Centred AI (HAII) is a **vision** of how machines might **augment** humans.
- **Humans in the group; computers in the loop**.

Shneiderman, B. (2021). Human-centered AI. *Issues in Science and Technology, 37(2*), 56-61.

Human-Centred AI (Shneiderman, 2021)



Governance Structures for Human-Centered AI

Trustworthy Certification: External Reviews

Independent Oversight:
Government Regulation
Auditing Firms
Insurance Companies
NGOs & Civil Society
Professional Organizations

Safety Culture: Organizational Design

Management Strategies:
Leadership Commitment
Hiring & Training
Failures & Near Misses
Internal Reviews
Industry Standards

Reliable Systems: Software Engineering

Technical Practices:
Audit Trails, SE Workflows
Verification & Bias testing
Explainable UIs

TEAM

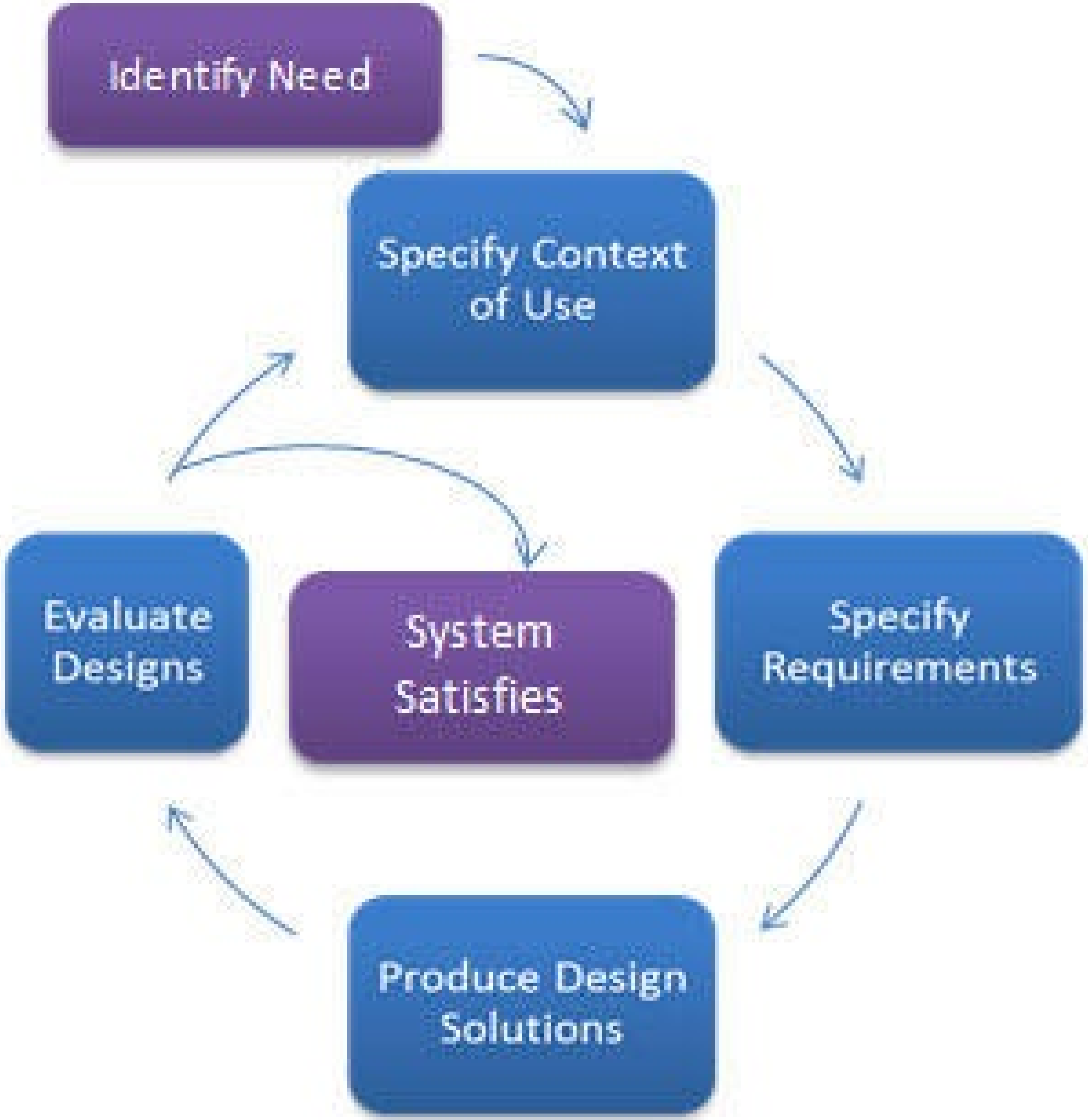
ORGANIZATION

INDUSTRY

User-centred Design

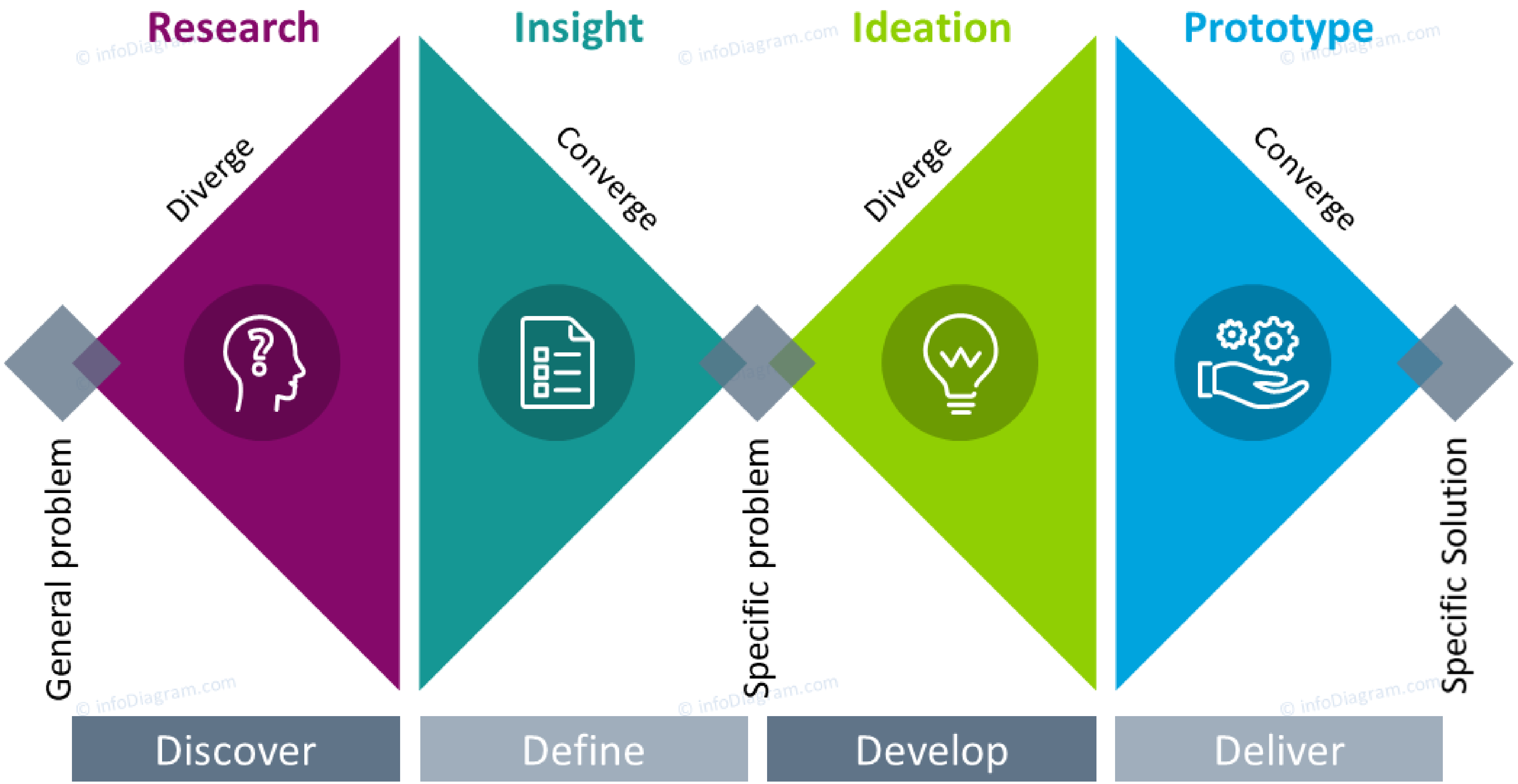
User-/Human-centred Design (UCD/HCD)

- Participatory Design (PD)
- Usability
- User Experience (UX)



6 Principles for Human-centred Design
(ISO 9241-210: 2010)

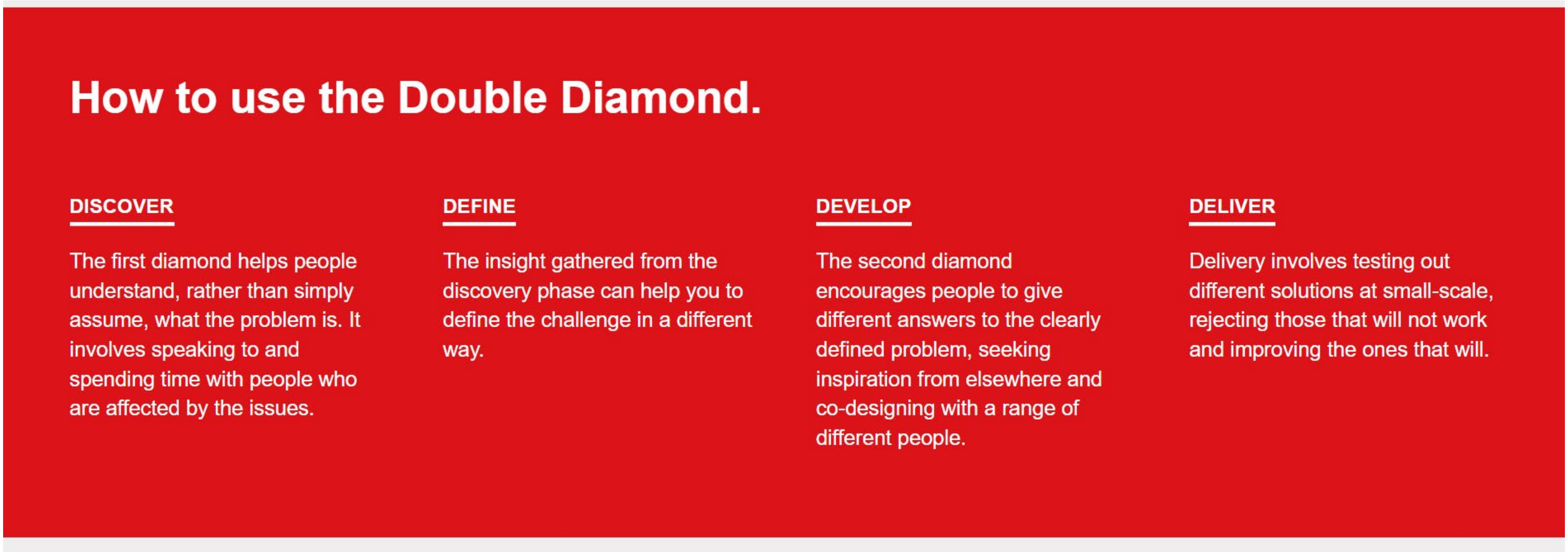
Double Diamond Method (the British Design Council, 2005)



Durham University

https://www.designcouncil.org.uk/our-resources/the-double-diamond/

Double Diamond Process (Richard Eisermann, 2003)



How to use the Double Diamond.

DISCOVER

The first diamond helps people understand, rather than simply assume, what the problem is. It involves speaking to and spending time with people who are affected by the issues.

DEFINE

The insight gathered from the discovery phase can help you to define the challenge in a different way.

DEVELOP

The second diamond encourages people to give different answers to the clearly defined problem, seeking inspiration from elsewhere and co-designing with a range of different people.

DELIVER

Delivery involves testing out different solutions at small-scale, rejecting those that will not work and improving the ones that will.

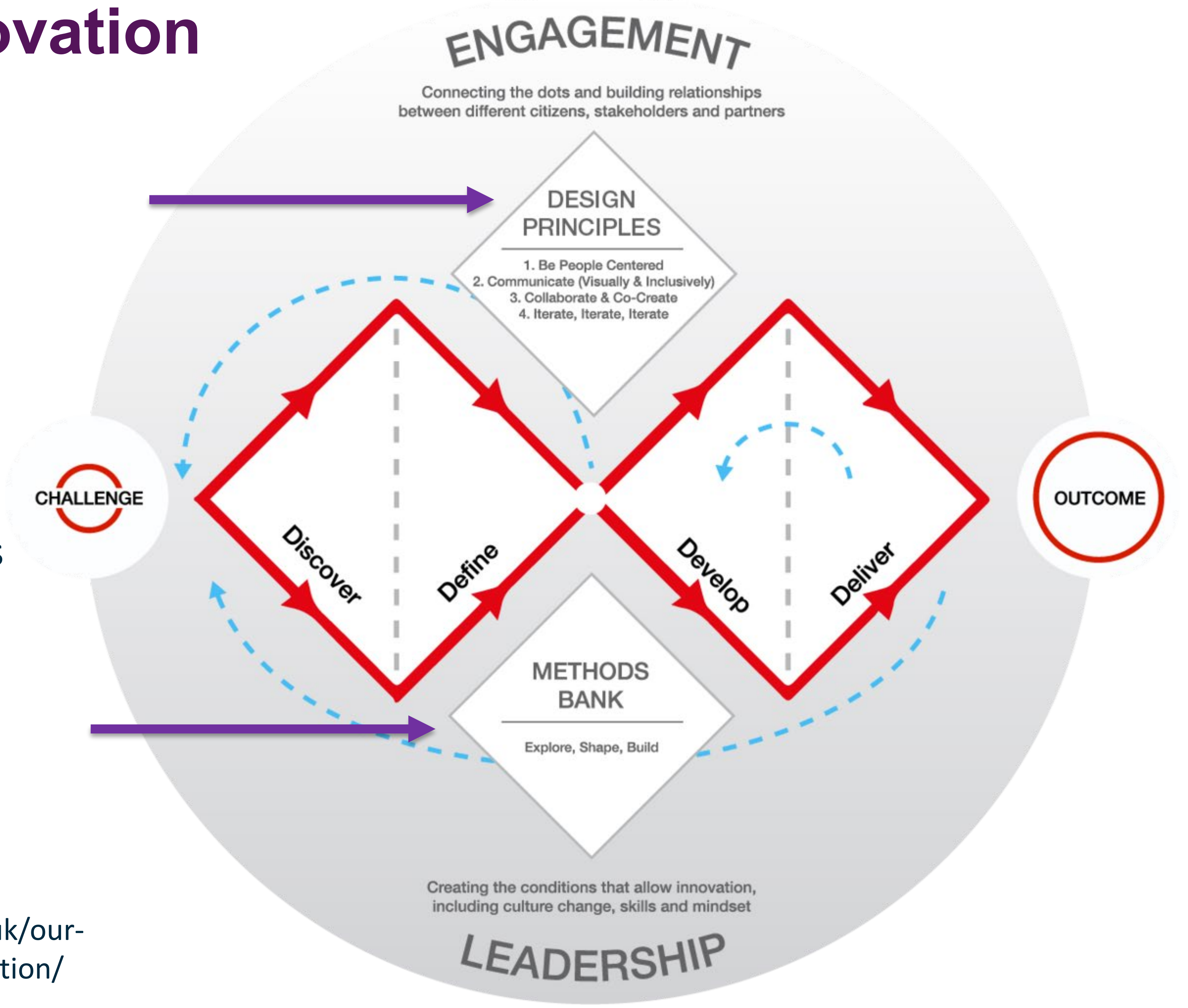
Framework for Innovation

4 Design Principles

- *Put people first*.
- *Communicate visually and inclusively.*
- *Collaborate and co-create*
- *Iterate, iterate, iterate.*

Methods Bank

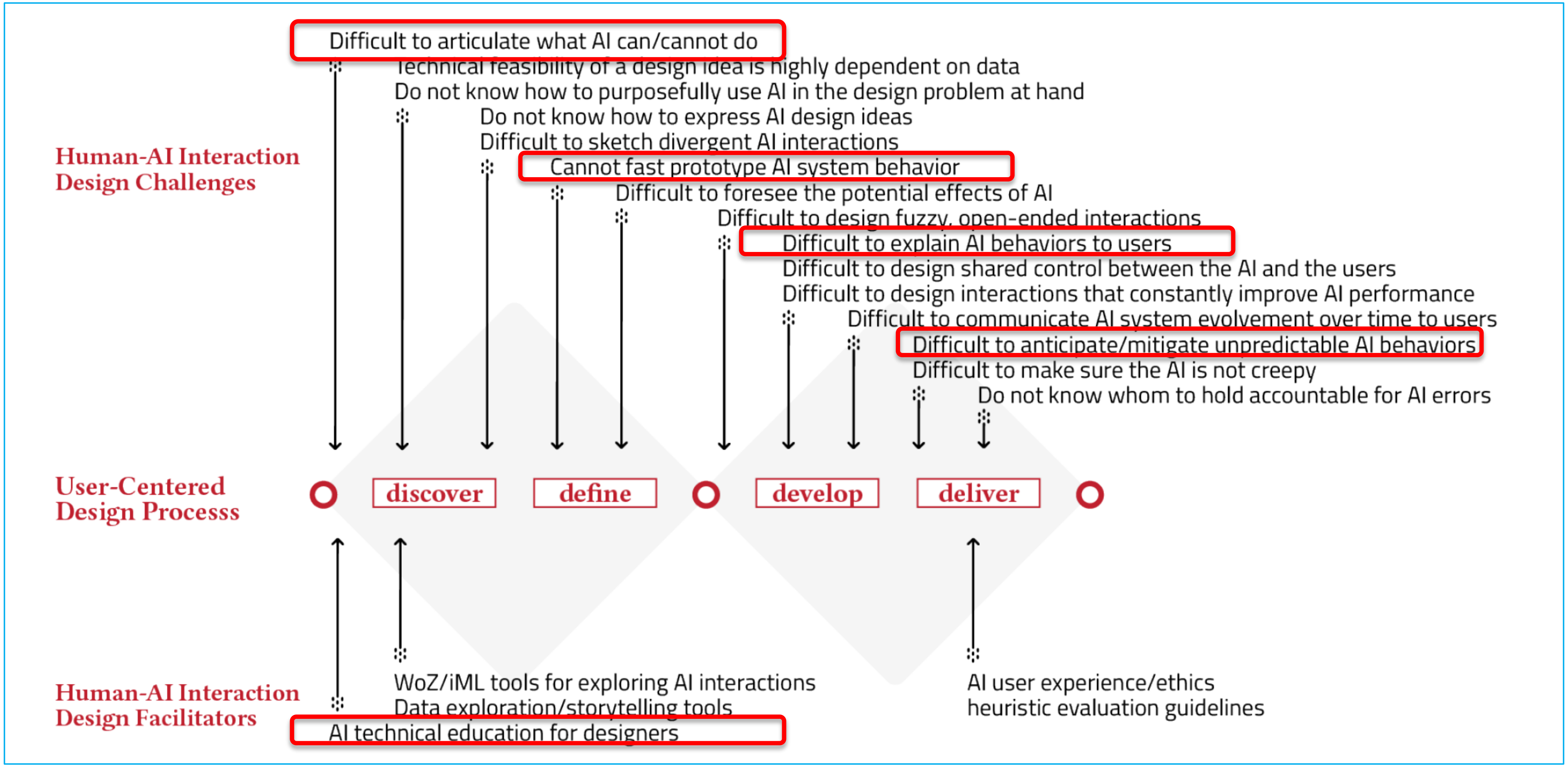
- *Explore*: challenges, needs and opportunities
- *Shape*: prototypes, insights and visions
- *Build:* ideas, plans and expertise


ENGAGEMENT
Connecting the dots and building relationships between different citizens, stakeholders and partners
DESIGN PRINCIPLES
1. Be People Centered
2. Communicate (Visually & Inclusively)
3. Collaborate & Co-Create
4. Iterate, Iterate, Iterate
CHALLENGE
Discover
Define
Develop
Deliver
OUTCOME
METHODS BANK
Explore, Shape, Build
Creating the conditions that allow innovation, including culture change, skills and mindset
LEADERSHIP

Durham University

https://www.designcouncil.org.uk/our-resources/framework-for-innovation/

Human-Centred AI (Yang et al. 2020)



Human-AI Interaction Design Challenges

Difficult to articulate what AI can/cannot do

Technical feasibility of a design idea is highly dependent on data

Do not know how to purposefully use AI in the design problem at hand

Do not know how to express AI design ideas

Difficult to sketch divergent AI interactions

Cannot fast prototype AI system behavior

Difficult to foresee the potential effects of AI

Difficult to design fuzzy, open-ended interactions

Difficult to explain AI behaviors to users

Difficult to design shared control between the AI and the users

Difficult to design interactions that constantly improve AI performance

Difficult to communicate AI system evolvement over time to users

Difficult to anticipate/mitigate unpredictable AI behaviors

Difficult to make sure the AI is not creepy

Do not know whom to hold accountable for AI errors

User-Centered Design Processs

Human-AI Interaction Design Facilitators

WoZ/iML tools for exploring AI interactions

Data exploration/storytelling tools

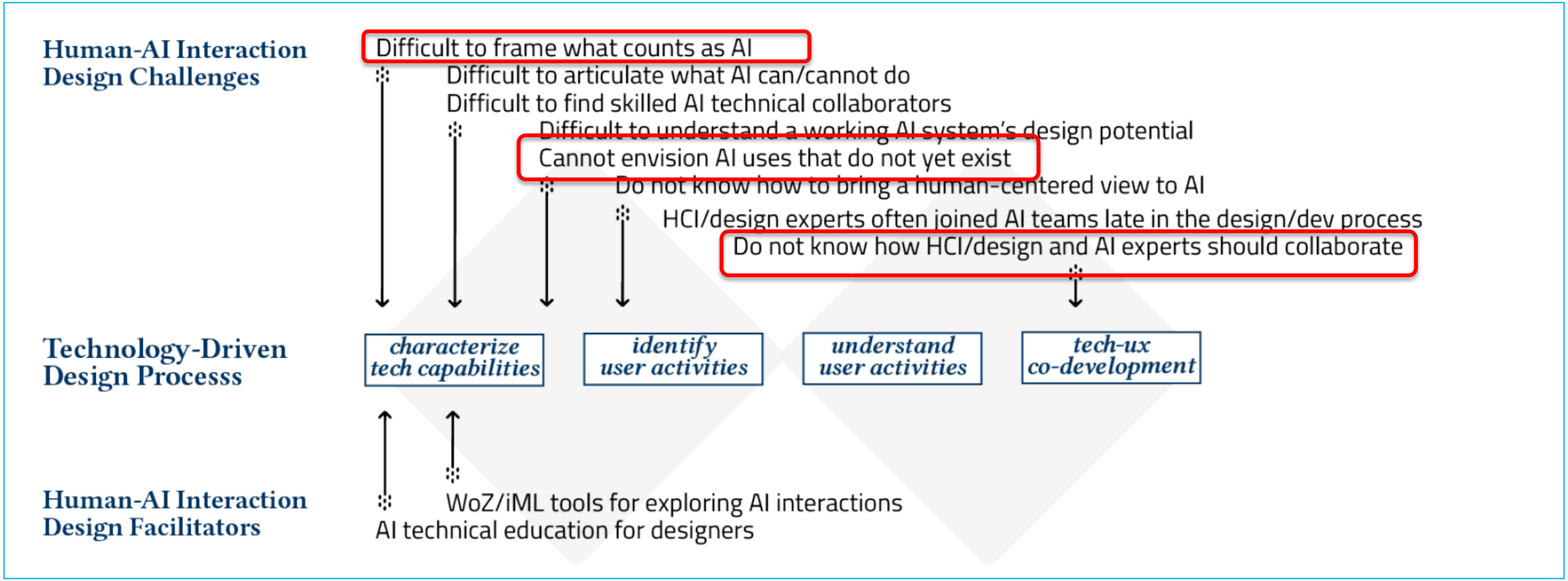
AI technical education for designers

AI user experience/ethics heuristic evaluation guidelines

Durham University

Yang, Q., Steinfeld, A., Rosé, C., & Zimmerman, J. (2020, April). Re-examining whether, why, and how human-AI interaction is uniquely difficult to design. In Proceedings of the 2020 chi conference on human factors in computing systems (pp. 1-13).

Human-Centred AI (Yang et al., 2020)



Human-AI Interaction Design Challenges

Difficult to frame what counts as AI

Difficult to articulate what AI can/cannot do

Difficult to find skilled AI technical collaborators

Difficult to understand a working AI system's design potential

Cannot envision AI uses that do not yet exist

Do not know how to bring a human-centered view to AI

HCI/design experts often joined AI teams late in the design/dev process

Do not know how HCI/design and AI experts should collaborate

Technology-Driven Design Processs

characterize tech capabilities

identify user activities

understand user activities

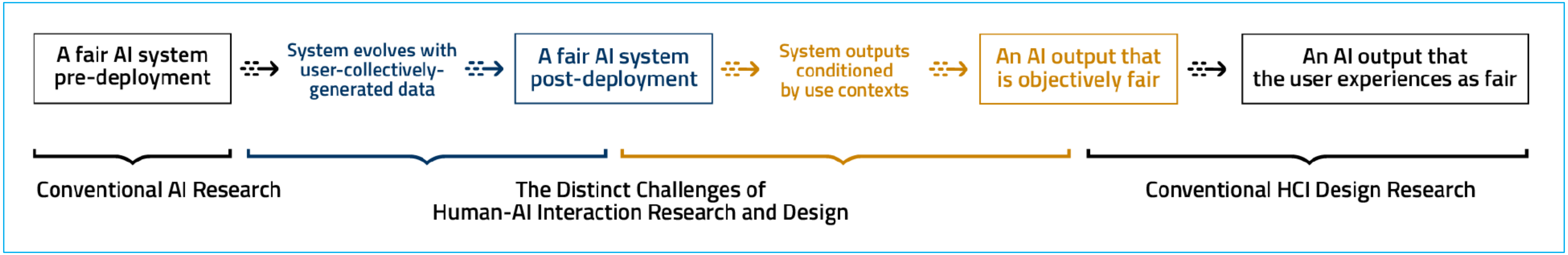
tech-ux co-development

Human-AI Interaction Design Facilitators

WoZ/iML tools for exploring AI interactions

AI technical education for designers

Human-Centred AI (Yang et al. 2020)

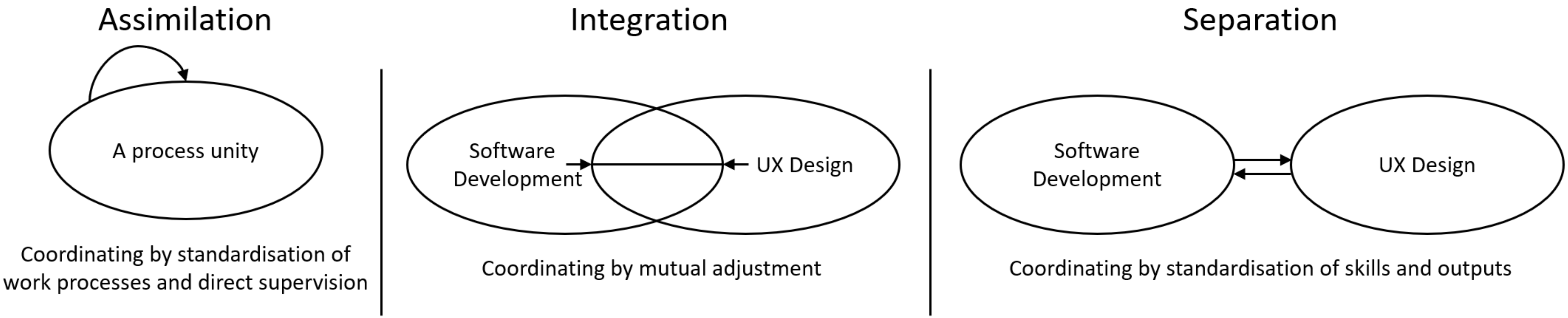

A fair AI system pre-deployment
System evolves with user-collectively-generated data
A fair AI system post-deployment
System outputs conditioned by use contexts
An AI output that is objectively fair
An AI output that the user experiences as fair
Conventional AI Research
The Distinct Challenges of Human-AI Interaction Research and Design
Conventional HCI Design Research

Durham University

Human-Centred AI (Yang et al. 2020)

- Improving designers’ **technical literacy.**
- Facilitating design-oriented data exploration
- Enabling designers to more easily **play with** AI in support of design ideation to gain a felt sense of what AI can do.
- Aiding designers in evaluating AI outputs.
- Creating AI-specific design processes.

Coordination Mechanisms in AI Development



Assimilation

A process unity

Coordinating by standardisation of work processes and direct supervision

Integration

Software Development

UX Design

Coordinating by mutual adjustment

Separation

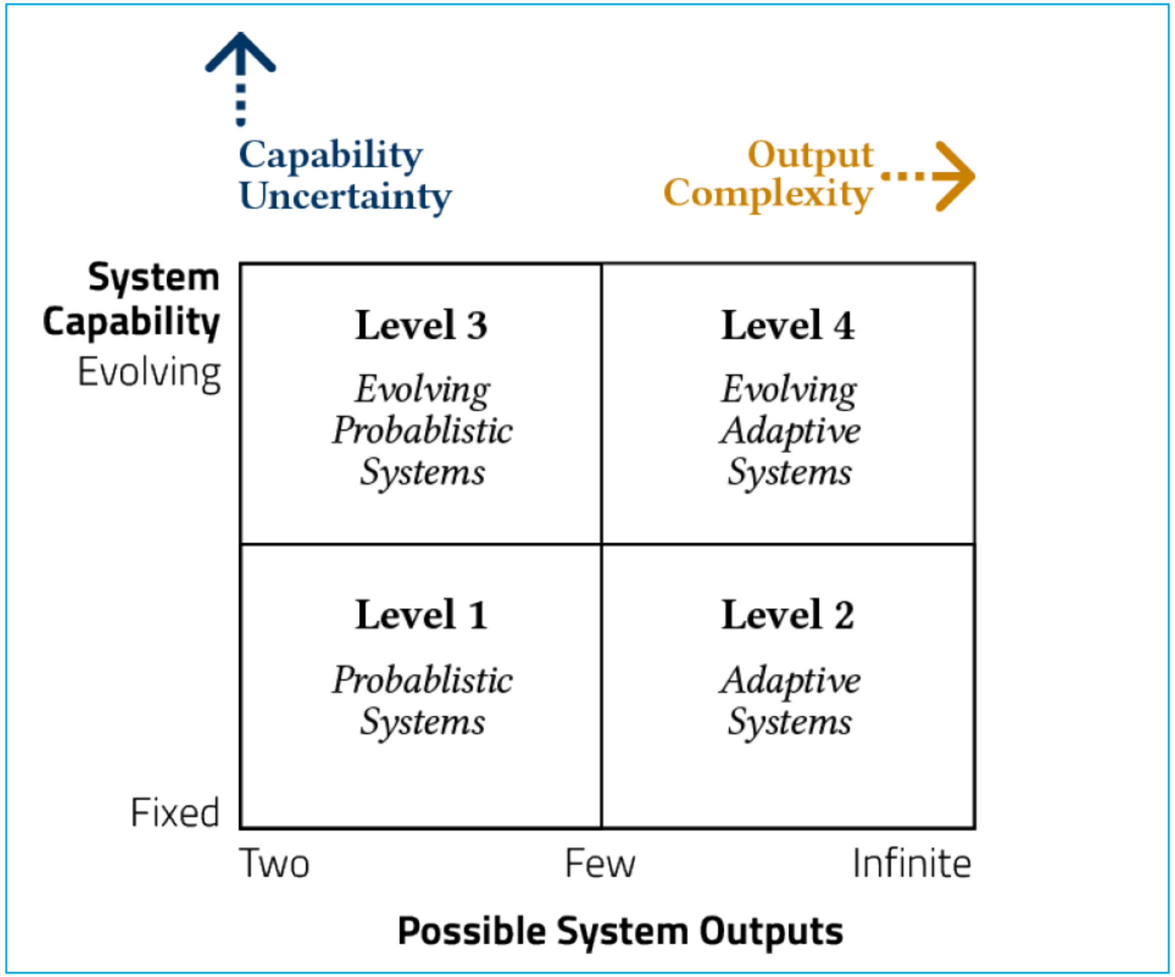
Software Development

UX Design

Coordinating by standardisation of skills and outputs

John Stouby Persson, Anders Bruun, Marta Kristín Larusdóttir, and Peter Axel Nielsen. 2022. Agile Software Development and UX Design: A Case Study of Integration by Mutual Adjustment. Journal of Information Systems and Technology, *152*, 107059.

Human-Centred AI (Yang et al. 2020)



Capability Uncertainty

Output Complexity

System Capability

Evolving

Fixed

Level 3

Evolving Probablistic Systems

Level 4

Evolving Adaptive Systems

Level 1

Probablistic Systems

Level 2

Adaptive Systems

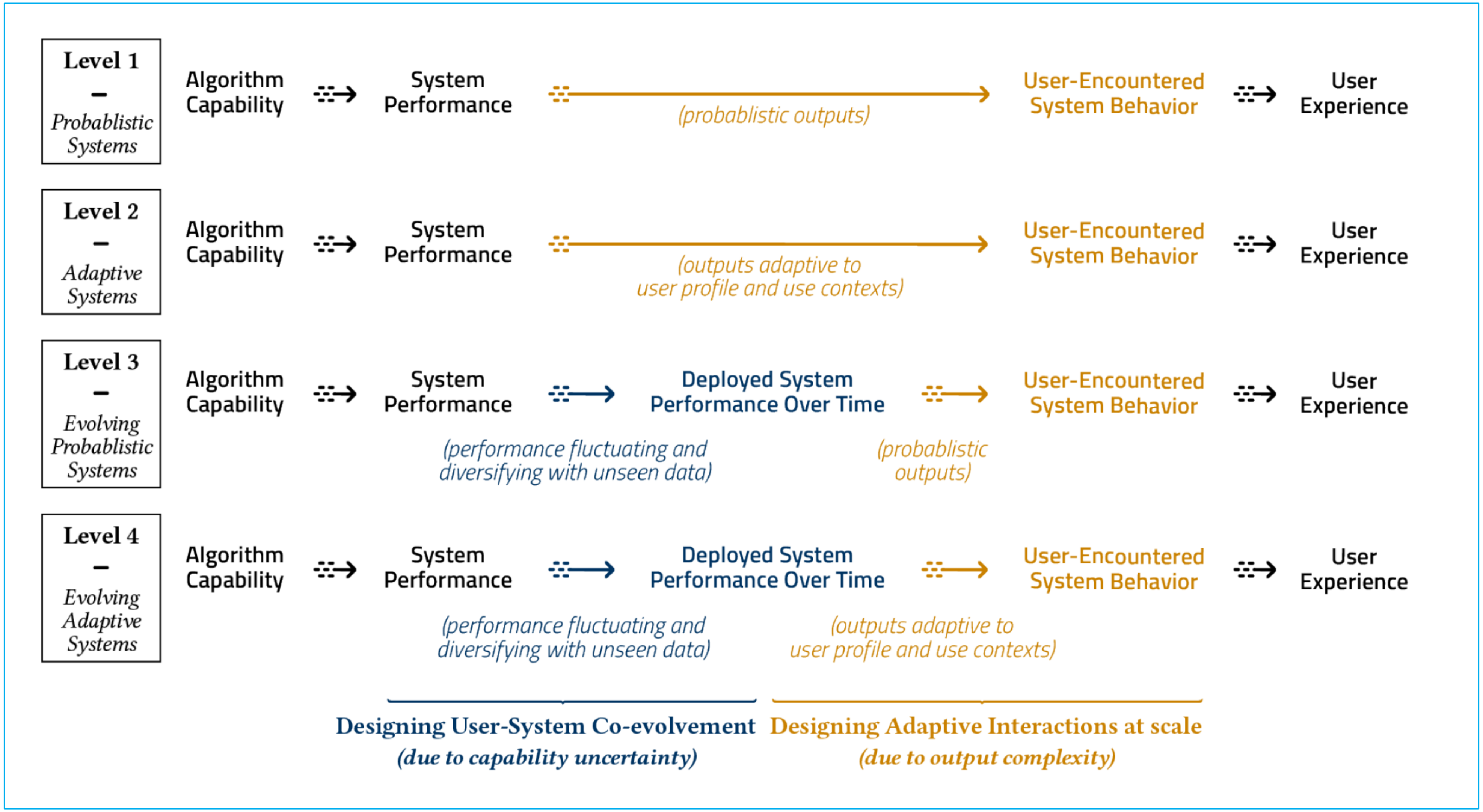
Two

Few

Infinite

Possible System Outputs

Human-Centred AI (Yang et al. 2020)



Level 1 – *Probablistic Systems*

Algorithm Capability → System Performance → User-Encountered System Behavior → User Experience

(probablistic outputs)

Level 2 – *Adaptive Systems*

Algorithm Capability → System Performance → User-Encountered System Behavior → User Experience

(outputs adaptive to user profile and use contexts)

Level 3 – *Evolving Probablistic Systems*

Algorithm Capability → System Performance → Deployed System Performance Over Time → User-Encountered System Behavior → User Experience

(performance fluctuating and diversifying with unseen data)

(probablistic outputs)

Level 4 – *Evolving Adaptive Systems*

Algorithm Capability → System Performance → Deployed System Performance Over Time → User-Encountered System Behavior → User Experience

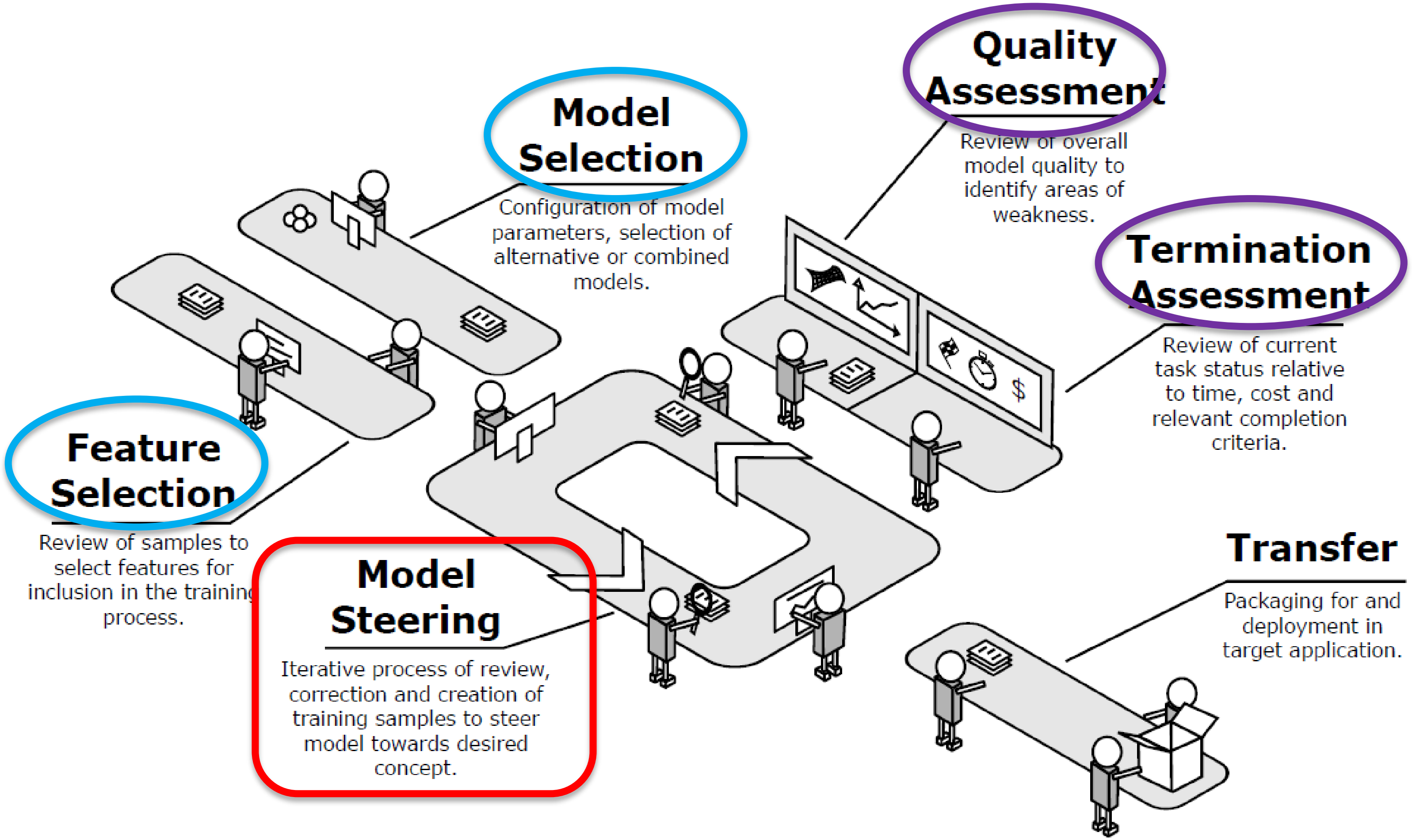
(performance fluctuating and diversifying with unseen data)

(outputs adaptive to user profile and use contexts)

Designing User-System Co-evolvement
(due to capability uncertainty)

Designing Adaptive Interactions at scale
(due to output complexity)

UI Design for iML (Dudly & Kristensson, 2018)



Feature Selection

Review of samples to select features for inclusion in the training process.

Model Selection

Configuration of model parameters, selection of alternative or combined models.

Model Steering

Iterative process of review, correction and creation of training samples to steer model towards desired concept.

Quality Assessment

Review of overall model quality to identify areas of weakness.

Termination Assessment

Review of current task status relative to time, cost and relevant completion criteria.

Transfer

Packaging for and deployment in target application.

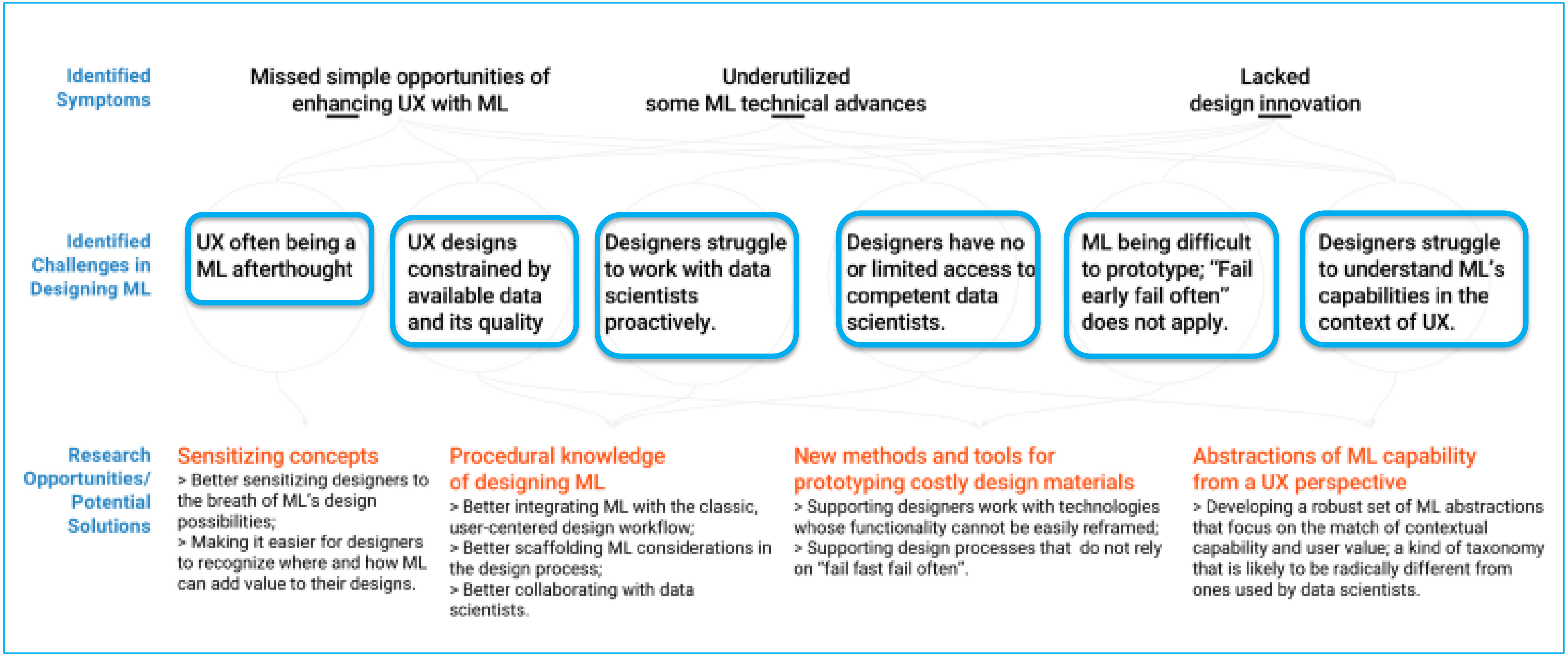
Dudley, J. J., & Kristensson, P. O. (2018). A review of user interface design for interactive machine learning. ACM Transactions on Interactive Intelligent Systems (TiiS), 8(2), 1-37.

Machine Learning (ML) as a Design Material (Dove et al., 2017)

- ML difficult to prototype because such interactions are dynamic, and their outcomes are potentially unpredictable. Potential difficulties in gathering enough user data for successful ML, and in drawing useful insights from these data
- Represent ML's dependency on data in early prototypes - foreground ethical considerations of ML
- The technical complexity is a challenge - the need to better understand and design for that complexity. It can get deep and unfamiliar very quickly, and designers need some level of expertise to function and contribute to the work at hand.

Graham Dove, Kim Halskov, Jodi Forlizzi, and John Zimmerman. 2017. UX Design Innovation: Challenges for Working with Machine Learning as a Design Material. In Proceedings of CHI '17. ACM.

AI as a Design Material (Yang et al., 2018)



Identified Symptoms

Missed simple opportunities of enhancing UX with ML

Underutilized some ML technical advances

Lacked design innovation

Identified Challenges in Designing ML

UX often being a ML afterthought

UX designs constrained by available data and its quality

Designers struggle to work with data scientists proactively.

Designers have no or limited access to competent data scientists.

ML being difficult to prototype; "Fail early fail often" does not apply.

Designers struggle to understand ML's capabilities in the context of UX.

Research Opportunities/ Potential Solutions

Sensitizing concepts

> Better sensitizing designers to the breath of ML's design possibilities;
> Making it easier for designers to recognize where and how ML can add value to their designs.

Procedural knowledge of designing ML

> Better integrating ML with the classic, user-centered design workflow;
> Better scaffolding ML considerations in the design process;
> Better collaborating with data scientists.

New methods and tools for prototyping costly design materials

> Supporting designers work with technologies whose functionality cannot be easily reframed;
> Supporting design processes that do not rely on "fail fast fail often".

Abstractions of ML capability from a UX perspective

> Developing a robust set of ML abstractions that focus on the match of contextual capability and user value; a kind of taxonomy that is likely to be radically different from ones used by data scientists.

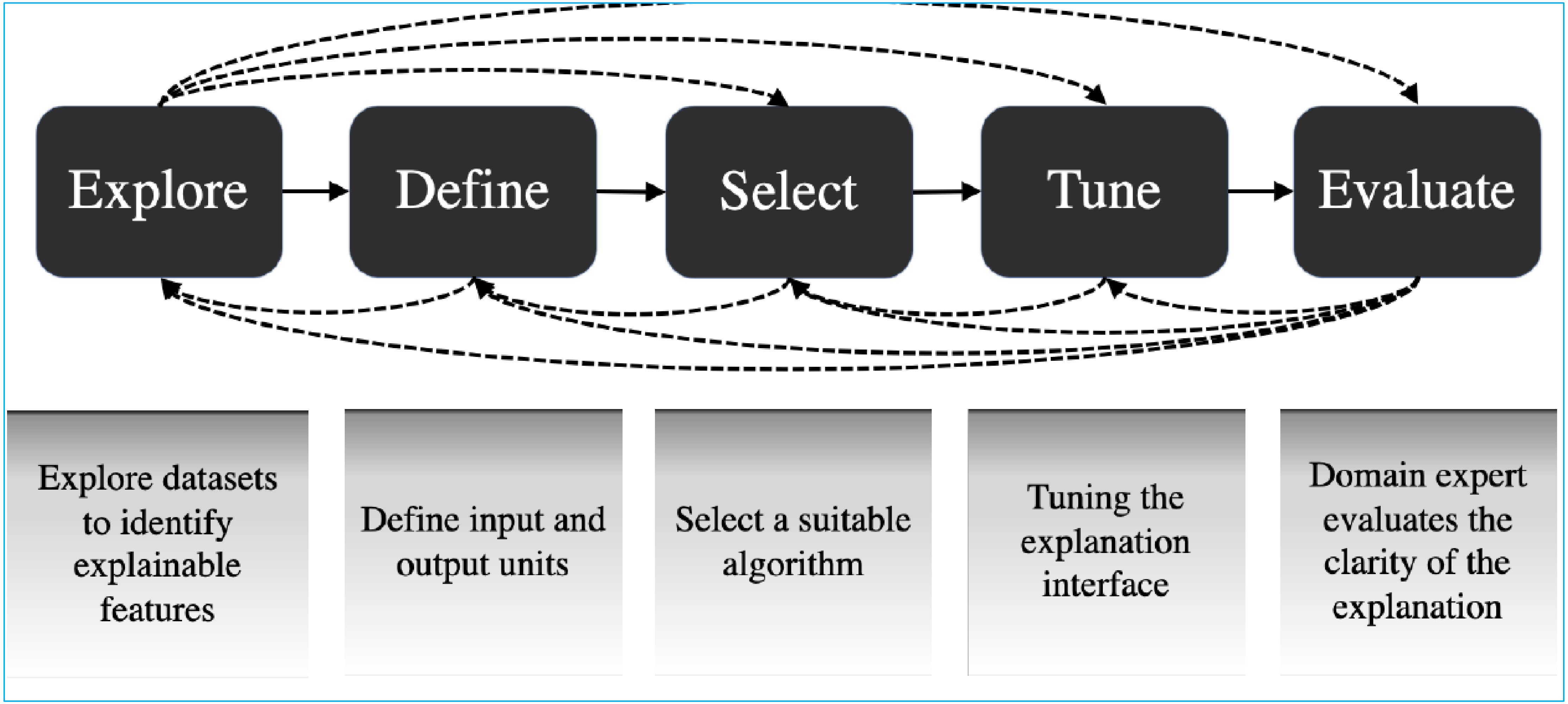
Yang, Q., Suh, J., Chen, N. C., & Ramos, G. (2018, June). Grounding interactive machine learning tool design in how non-experts actually build models. In Proceedings of the 2018 designing interactive systems conference (pp. 573-584).

AI as a Design Material (Yildirim et al. 2022)

- *"… we'll automate the low value tasks for people to move to better jobs"*, but no one designs what those jobs are. *"… you're not going to reach a 100% automation as the data itself is changing over time,"* and there is a role for the human in the loop to deal with the hard cases, but also to directly train the algorithm
- To understand the technicalness of what's going on with an AI solution ... what data is there that we can use as a design material.
- Collaboration between designers and data scientists: communication with tools (white board, flow diagrams …) A shared understanding or mental model of what [the AI system] does

Yildirim, N., Kass, A., Tung, T., Upton, C., Costello, D., Giusti, R., ... & Zimmerman, J. (2022, April). How Experienced Designers of Enterprise Applications Engage AI as a Design Material. In CHI Conference on Human Factors in Computing Systems (pp. 1-13).

Design Thinking & AI (Kurti et al. 2021)



Explore → Define → Select → Tune → Evaluate

Explore datasets to identify explainable features

Define input and output units

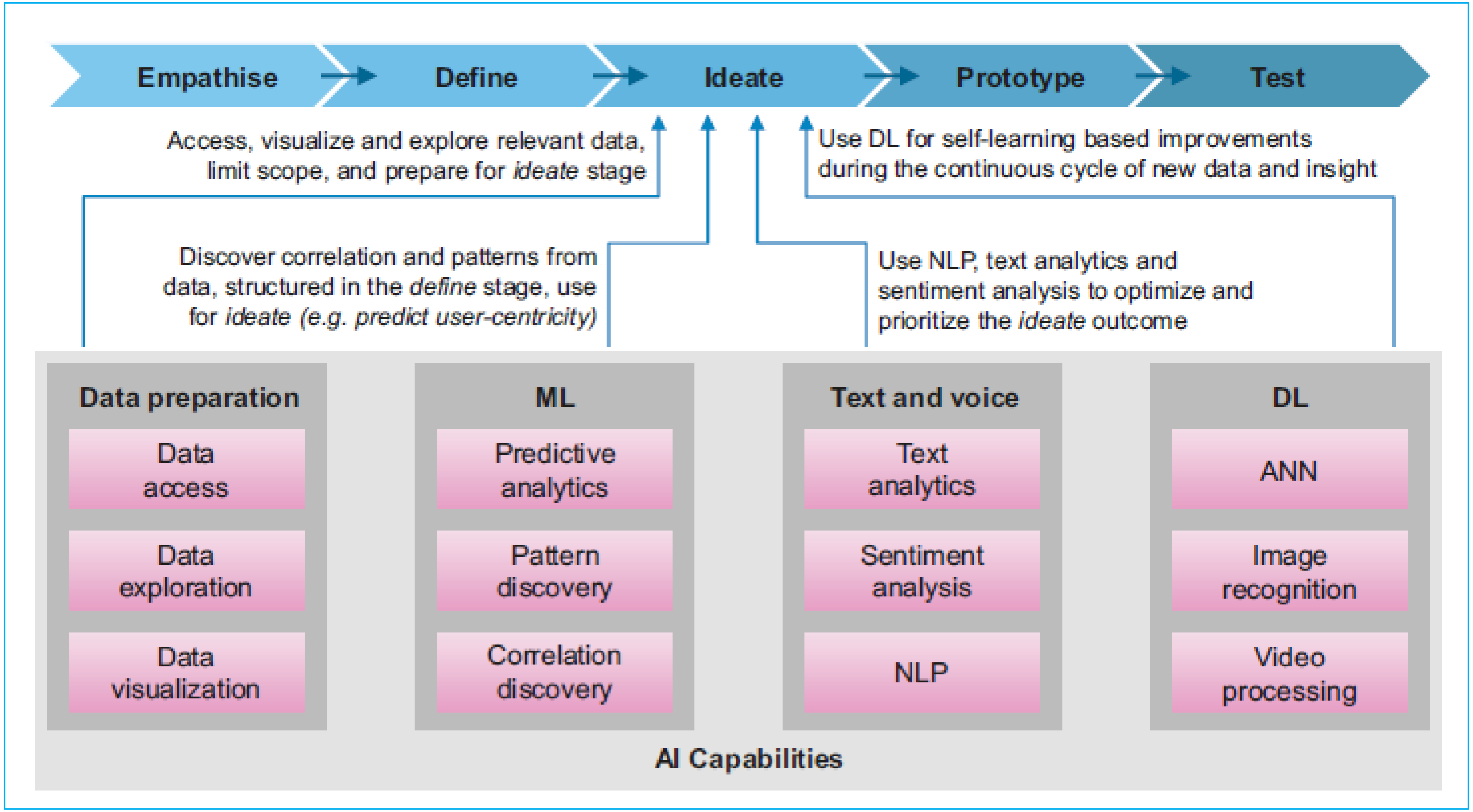
Select a suitable algorithm

Tuning the explanation interface

Domain expert evaluates the clarity of the explanation

Kurti, A., Dalipi, F., Ferati, M., & Kastrati, Z. (2021). Increasing the understandability and explainability of machine learning and artificial intelligence solutions: a design thinking approach. In *Proceedings of the 4th International Conference on Human Interaction and Emerging Technologies: Future Applications (IHIET–AI 2021), April 28-30, 2021, Strasbourg, France 4* (pp. 37-42). Springer International Publishing.

Design Thinking & AI (Hechler et al., 2020)



Empathise
Define
Ideate
Prototype
Test
Access, visualize and explore relevant data, limit scope, and prepare for *ideate* stage
Use DL for self-learning based improvements during the continuous cycle of new data and insight
Discover correlation and patterns from data, structured in the *define* stage, use for *ideate (e.g. predict user-centricity)*
Use NLP, text analytics and sentiment analysis to optimize and prioritize the *ideate* outcome
Data preparation
Data access
Data exploration
Data visualization
ML
Predictive analytics
Pattern discovery
Correlation discovery
Text and voice
Text analytics
Sentiment analysis
NLP
DL
ANN
Image recognition
Video processing
AI Capabilities

Hechler, E., Oberhofer, M., & Schaeck, T. (2020). Design Thinking and DevOps in the AI Context. In Deploying AI in the Enterprise (pp. 141-161). Apress, Berkeley, CA.

Design Methods & AI (van Allen, 2018)

Tools must address the prototyping and design of:

- Personality and character for autonomous behavior
- Multimodal, non-visual interactions
- User training/pruning/tending/learning
- Point of view and biases (considering diversity in people and machines)
- Mixed social interactions—M2H and M2M
- Intentions/goals/rules
- Ethics/civic responsibility
- Indication of expertise and affordance.

van Allen, P. (2018). Prototyping ways of prototyping AI. Interactions, 25(6), 46-51.

Design Methods & AI (van Allen, 2018)

And the tools must enable the designer to use methods such as:

- Fast, iterative, experimental prototyping of interactions and autonomous behaviours
- Wizard of Oz (WoZ) design experiments and testing with people
- Comparing different algorithms, datasets, and training methods
- Iterative testing of embodied working prototypes
- Finding a minimum viable product (MVP) and minimum viable data (MVD)
- Creating new AI technology requirements.

Design Methods & AI – Model Cards (Mitchell et al., 2019)

Model cards are used to disclose information about a trained ML model, including how the model was built, assumptions for its development, how users of different demographics experience different model behaviour… – **Standardised Documentation**

Responsible democratisation of ML/AL; Increasing transparency of ML/AI

- *Model Details: Basic information about the model*
- *Intended Use: Use cases envisioned*
- *Factors: Demographics, environmental conditions, technical attributes, etc.*
- *Metrics: Measuring real-world impact*
- *Evaluation Data: Datasets for quantitative analysis*
- *Training Data*
- *Quantitative Analyses*
- *Ethical Considerations*
- *Caveats and Recommendations*

Mitchell, M., Wu, S., Zaldivar, A., Barnes, P., Vasserman, L., Hutchinson, B., ... & Gebru, T. (2019, January). Model cards for model reporting. In *Proceedings of the Conference on Fairness, Accountability, and Transparency (FAT 2019)* (pp. 220-229).

Design Methods & AI – Model Cards, Example

Model Details

- Developed by researchers at Google and the University of Toronto, 2018, v1.
- Convolutional Neural Net.
- Pretrained for face recognition then fine-tuned with cross-entropy loss for binary smiling classification

Intended Use

- Intended to be used for fun applications, such as creating cartoon smiles on real images; augmentative applications, such as providing details for people who are blind; or assisting applications such as automatically finding smiling photos.
- Particularly intended for younger audiences.
- Not suitable for emotion detection or determining affect; smiles were annotated based on physical appearance, and not underlying emotions.

Factors

- Based on known problems with computer vision face technology, potential relevant factors include groups for gender, age, race, and Fitzpatrick skin type; hardware factors of camera type and lens type; and environmental factors of lighting and humidity.
- Evaluation factors are gender and age group, as annotated in the publicly available dataset CelebA . Further possible factors not currently available in a public smiling dataset. Gender and age determined by third-party annotators based on visual presentation, following a set of examples of male/female gender and young/old age.

Design Methods & AI – Model Cards, Example

Metrics

- Evaluation metrics include False Positive Rate and False Negative Rate to measure disproportionate model performance errors across subgroups. False Discovery Rate and False Omission Rate, which measure the fraction of negative (not smiling) and positive (smiling) predictions that are incorrectly predicted to be positive and negative, respectively, are also reported. [48]
- Together, these four metrics provide values for different errors that can be calculated from the confusion matrix for binary classification systems.
- These also correspond to metrics in recent definitions of “fairness” in machine learning, where parity across subgroups for different metrics correspond to different fairness criteria.
- 95% confidence intervals calculated with bootstrap resampling.
- All metrics reported at the .5 decision threshold, where all error types (FPR, FNR, FDR, FOR) are within the same range (0.04 - 0.14).

Evaluation Data

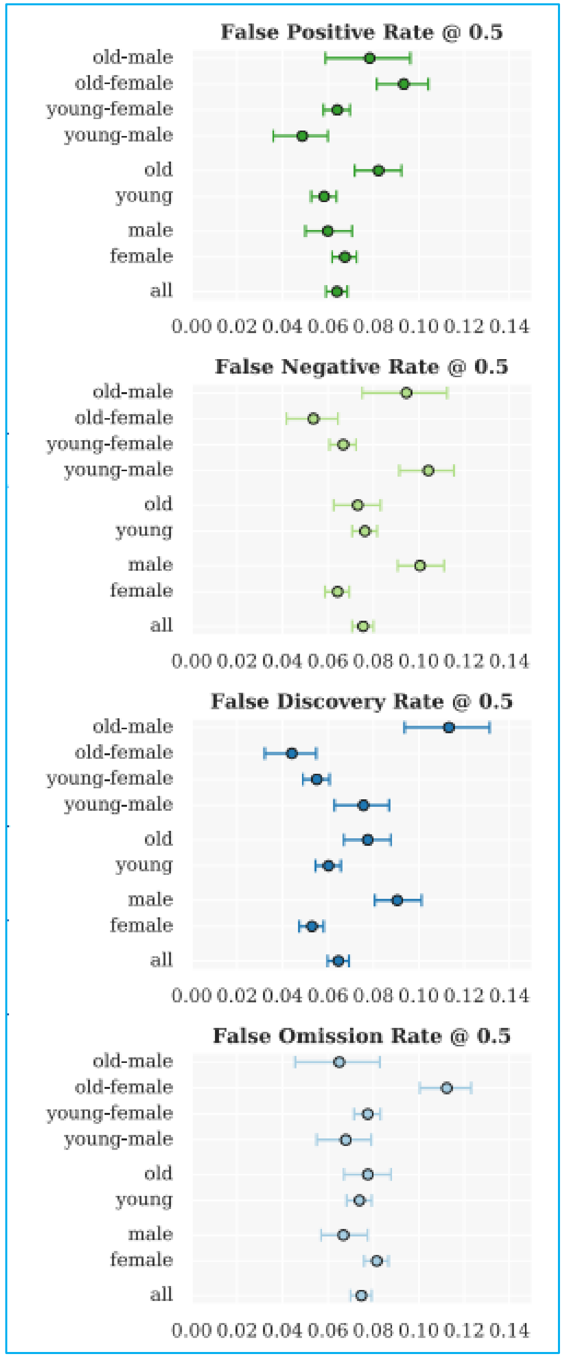
- CelebA, test data split. - Chosen as a basic proof-of-concept

Training Data

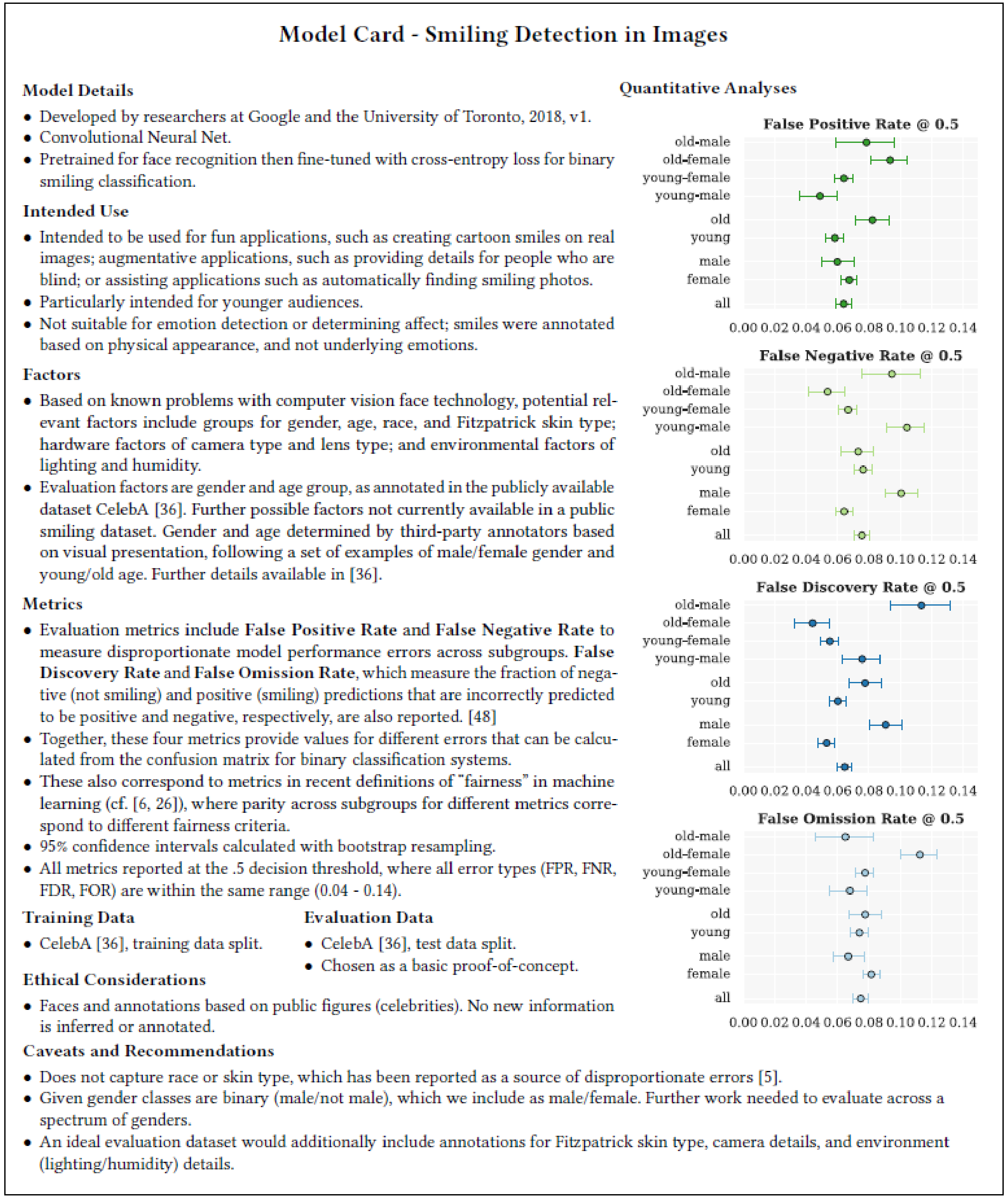
- CelebA, training data split

Design Methods & AI – Model Cards, Example

- **Quantitative Analyses**
- **Ethical Considerations**
  - Faces and annotations based on public figures (celebrities). No new information is inferred or annotated
- **Caveats and Recommendations**
  - Does not capture race or skin type, which has been reported as a source of disproportionate errors.
  - Given gender classes are binary (male/not male), further work needed to evaluate across a spectrum of genders.
  - An ideal evaluation dataset would additionally include annotations for Fitzpatrick skin type, camera details, and environment (lighting/humidity) details



Model Card: Example 1
Smiling Detection in Images



Model Card - Smiling Detection in Images

Model Details

- Developed by researchers at Google and the University of Toronto, 2018, v1.
- Convolutional Neural Net.
- Pretrained for face recognition then fine-tuned with cross-entropy loss for binary smiling classification.

Intended Use

- Intended to be used for fun applications, such as creating cartoon smiles on real images; augmentative applications, such as providing details for people who are blind; or assisting applications such as automatically finding smiling photos.
- Particularly intended for younger audiences.
- Not suitable for emotion detection or determining affect; smiles were annotated based on physical appearance, and not underlying emotions.

Factors

- Based on known problems with computer vision face technology, potential relevant factors include groups for gender, age, race, and Fitzpatrick skin type; hardware factors of camera type and lens type; and environmental factors of lighting and humidity.
- Evaluation factors are gender and age group, as annotated in the publicly available dataset CelebA [36]. Further possible factors not currently available in a public smiling dataset. Gender and age determined by third-party annotators based on visual presentation, following a set of examples of male/female gender and young/old age. Further details available in [36].

Metrics

- Evaluation metrics include **False Positive Rate** and **False Negative Rate** to measure disproportionate model performance errors across subgroups. **False Discovery Rate** and **False Omission Rate**, which measure the fraction of negative (not smiling) and positive (smiling) predictions that are incorrectly predicted to be positive and negative, respectively, are also reported. [48]
- Together, these four metrics provide values for different errors that can be calculated from the confusion matrix for binary classification systems.
- These also correspond to metrics in recent definitions of "fairness" in machine learning (cf. [6, 26]), where parity across subgroups for different metrics correspond to different fairness criteria.
- 95% confidence intervals calculated with bootstrap resampling.
- All metrics reported at the .5 decision threshold, where all error types (FPR, FNR, FDR, FOR) are within the same range (0.04 - 0.14).

Training Data

- CelebA [36], training data split.

Evaluation Data

- CelebA [36], test data split.
- Chosen as a basic proof-of-concept.

Ethical Considerations

- Faces and annotations based on public figures (celebrities). No new information is inferred or annotated.

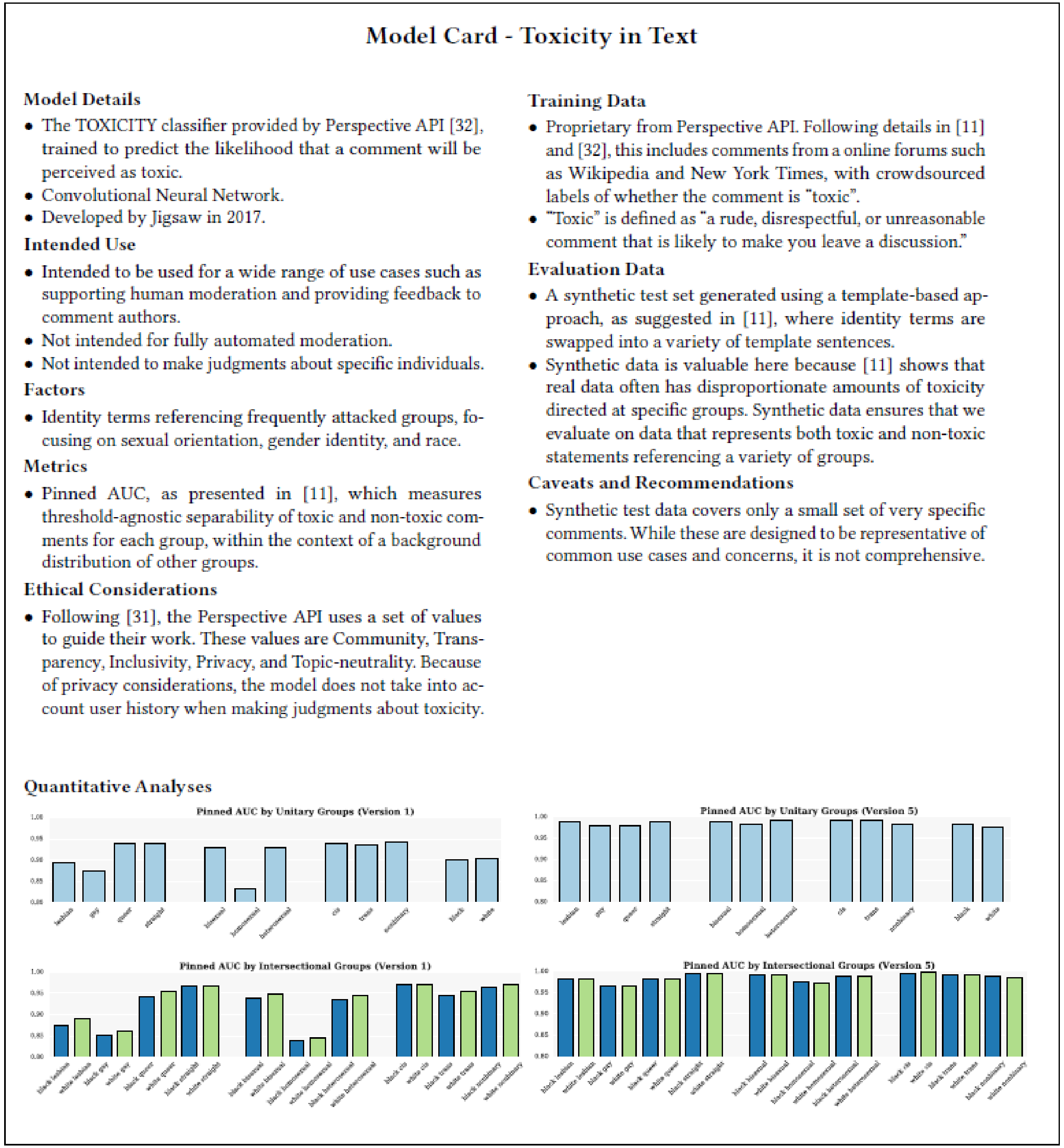
Quantitative Analyses

False Positive Rate @ 0.5
False Negative Rate @ 0.5
False Discovery Rate @ 0.5
False Omission Rate @ 0.5
old-male
old-female
young-female
young-male
old
young
male
female
all
0.00 0.02 0.04 0.06 0.08 0.10 0.12 0.14

Caveats and Recommendations

- Does not capture race or skin type, which has been reported as a source of disproportionate errors [5].
- Given gender classes are binary (male/not male), which we include as male/female. Further work needed to evaluate across a spectrum of genders.
- An ideal evaluation dataset would additionally include annotations for Fitzpatrick skin type, camera details, and environment (lighting/humidity) details.

Model Card: Example 2
Toxicity in Text



Model Card - Toxicity in Text

Model Details

- The TOXICITY classifier provided by Perspective API [32], trained to predict the likelihood that a comment will be perceived as toxic.
- Convolutional Neural Network.
- Developed by Jigsaw in 2017.

Intended Use

- Intended to be used for a wide range of use cases such as supporting human moderation and providing feedback to comment authors.
- Not intended for fully automated moderation.
- Not intended to make judgments about specific individuals.

Factors

- Identity terms referencing frequently attacked groups, focusing on sexual orientation, gender identity, and race.

Metrics

- Pinned AUC, as presented in [11], which measures threshold-agnostic separability of toxic and non-toxic comments for each group, within the context of a background distribution of other groups.

Ethical Considerations

- Following [31], the Perspective API uses a set of values to guide their work. These values are Community, Transparency, Inclusivity, Privacy, and Topic-neutrality. Because of privacy considerations, the model does not take into account user history when making judgments about toxicity.

Training Data

- Proprietary from Perspective API. Following details in [11] and [32], this includes comments from a online forums such as Wikipedia and New York Times, with crowdsourced labels of whether the comment is "toxic".
- "Toxic" is defined as "a rude, disrespectful, or unreasonable comment that is likely to make you leave a discussion."

Evaluation Data

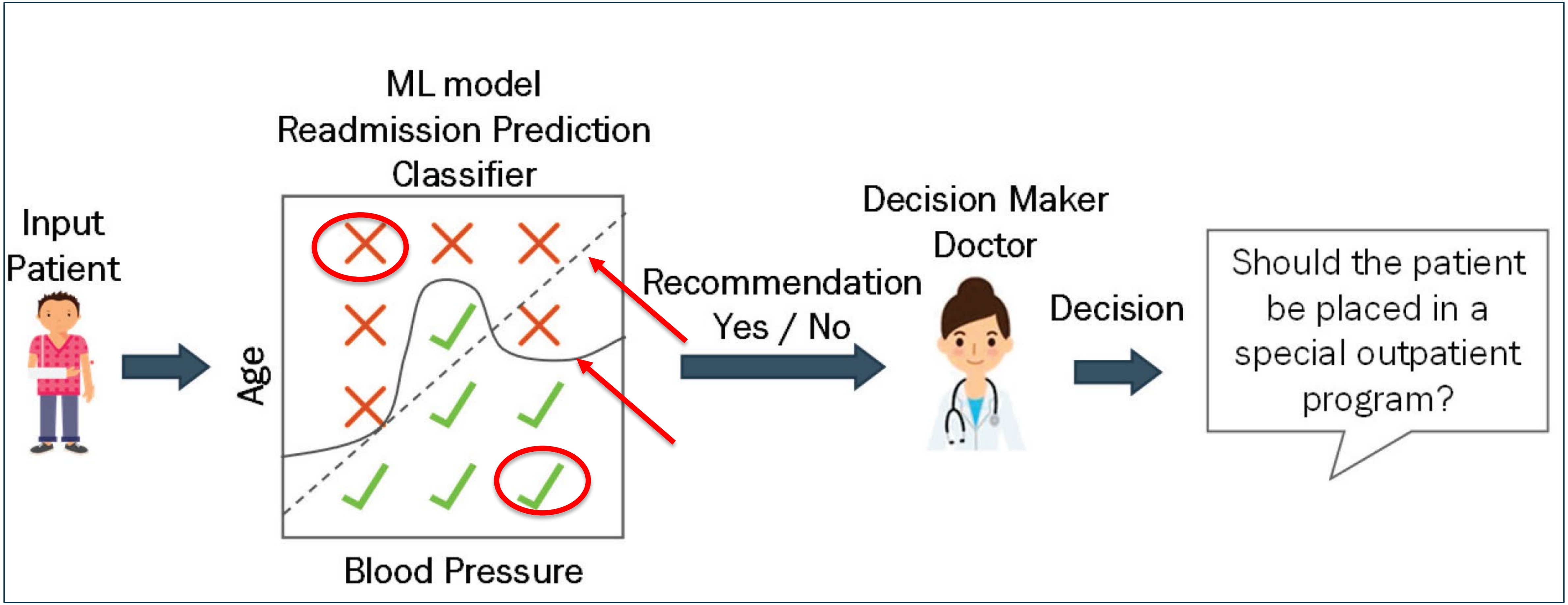
- A synthetic test set generated using a template-based approach, as suggested in [11], where identity terms are swapped into a variety of template sentences.
- Synthetic data is valuable here because [11] shows that real data often has disproportionate amounts of toxicity directed at specific groups. Synthetic data ensures that we evaluate on data that represents both toxic and non-toxic statements referencing a variety of groups.

Caveats and Recommendations

- Synthetic test data covers only a small set of very specific comments. While these are designed to be representative of common use cases and concerns, it is not comprehensive.

Quantitative Analyses

Case Study: *Beyond Accuracy: When does AI err?* (Bansai et al, 2019)



Bansal, G., Nushi, B., Kamar, E., Lasecki, W. S., Weld, D. S., & Horvitz, E. (2019, October). Beyond accuracy: The role of mental models in human-AI team performance. In Proceedings of the AAAI Conference on Human Computation and Crowdsourcing (Vol. 7, pp. 2-11).

Case Study: *Beyond Accuracy*

- **AI Error Boundaries – affecting human's ability to form an accurate mental model**
  - Parsimony
    - Simple to represent and can be modelled with a small number of features 9and conjunctive expressions) and cleanly as well as reliably distinguishes success from errors without uncertainty.
  - Stochasticity
    - having a random probability distribution that may be analysed statistically but may not be predicted precisely.
  - Task Dimensionality: the number of features defining each instance

- **Insights**
  - Do people create mental models of the error boundary? How do mental models evolve with interaction?
  - Do more parsimonious error boundaries facilitate mental model creation?
  - Do less stochastic error boundaries lead to better mental models?

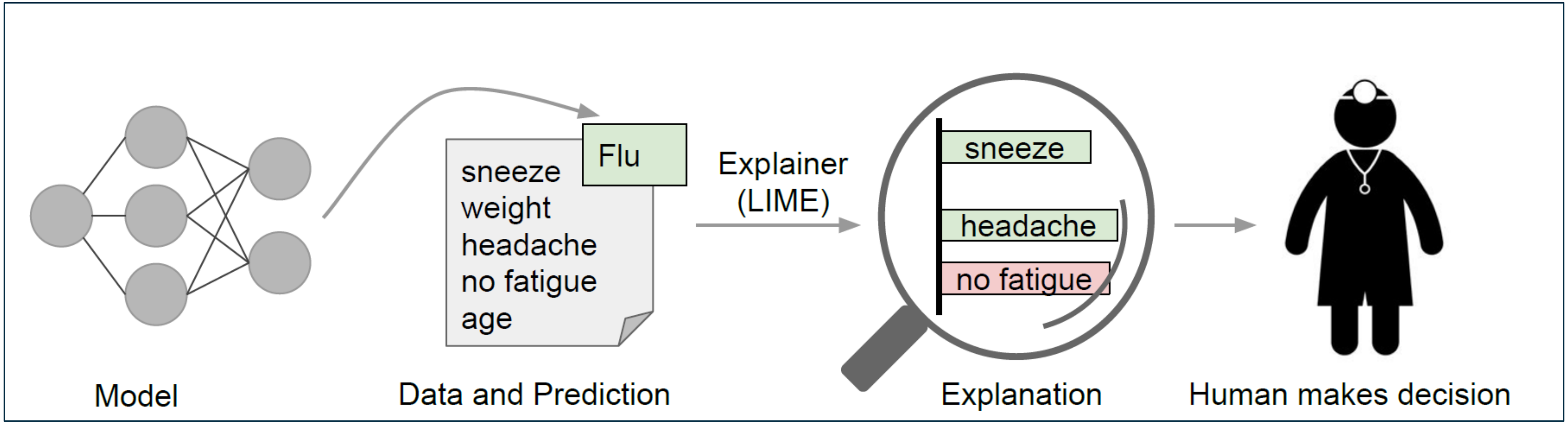
Case Study: *Beyond Accuracy*

The user studies presented suggest the following considerations when developing ML models to be used in AI advised human decision making:

- Build AI systems with parsimonious error boundaries.
- Minimize the stochasticity of system errors.
- Reduce task dimensionality when possible either by eliminating features that are irrelevant for both machine and human reasoning or most importantly by analysing the trade-off between the marginal gain of machine performance per added feature and the marginal loss of the accuracy of human mental models per added feature.
- Deploy models whose error boundaries are *backward compatible*, *i.e.* by regularizing to minimize the introduction of new errors on instances where the user has learned to trust the system.

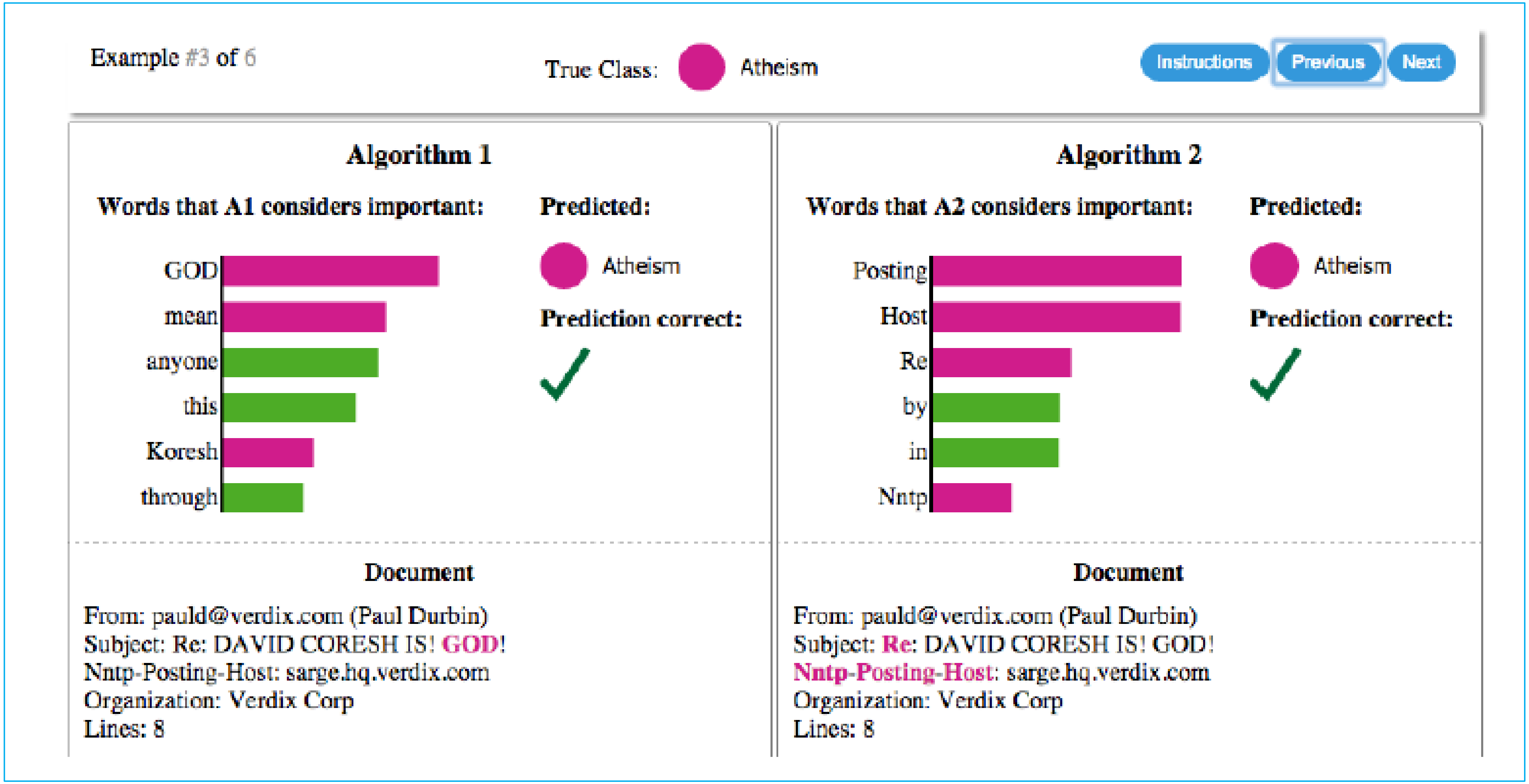
Case Study: *"Why Should I Trust You?"* (Ribeiro et al. 2016)

LIME - Local Interpretable Model-agnostic Explanations. The overall goal of **LIME** is to identify an interpretable model over the interpretable representation that is locally faithful to the classifier.



Ribeiro, M. T., Singh, S., & Guestrin, C. (2016, August). " Why should i trust you?" Explaining the predictions of any classifier. In Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining (pp. 1135-1144).

Case Study: *“Why Should I Trust You?”*



Example #3 of 6

True Class: ● Atheism

Algorithm 1

Words that A1 considers important:

Predicted:

● Atheism

Prediction correct:

✓

Document

From: pauld@verdix.com (Paul Durbin)
Subject: Re: DAVID CORESH IS! **GOD**!
Nntp-Posting-Host: sarge.hq.verdix.com
Organization: Verdix Corp
Lines: 8

Algorithm 2

Words that A2 considers important:

Predicted:

● Atheism

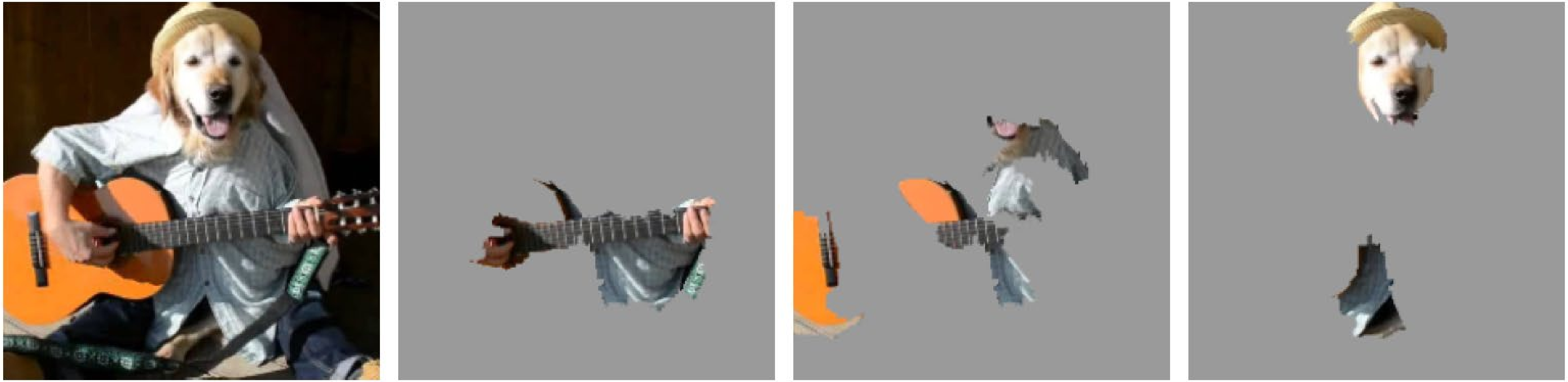
Prediction correct:

✓

Document

From: pauld@verdix.com (Paul Durbin)
Subject: **Re**: DAVID CORESH IS! GOD!
Nntp-Posting-Host: sarge.hq.verdix.com
Organization: Verdix Corp
Lines: 8

Case Study: *“Why Should I Trust You?”*



(a) Original Image (b) Explaining *Electric guitar* (c) Explaining *Acoustic guitar* (d) Explaining *Labrador*

Questions?