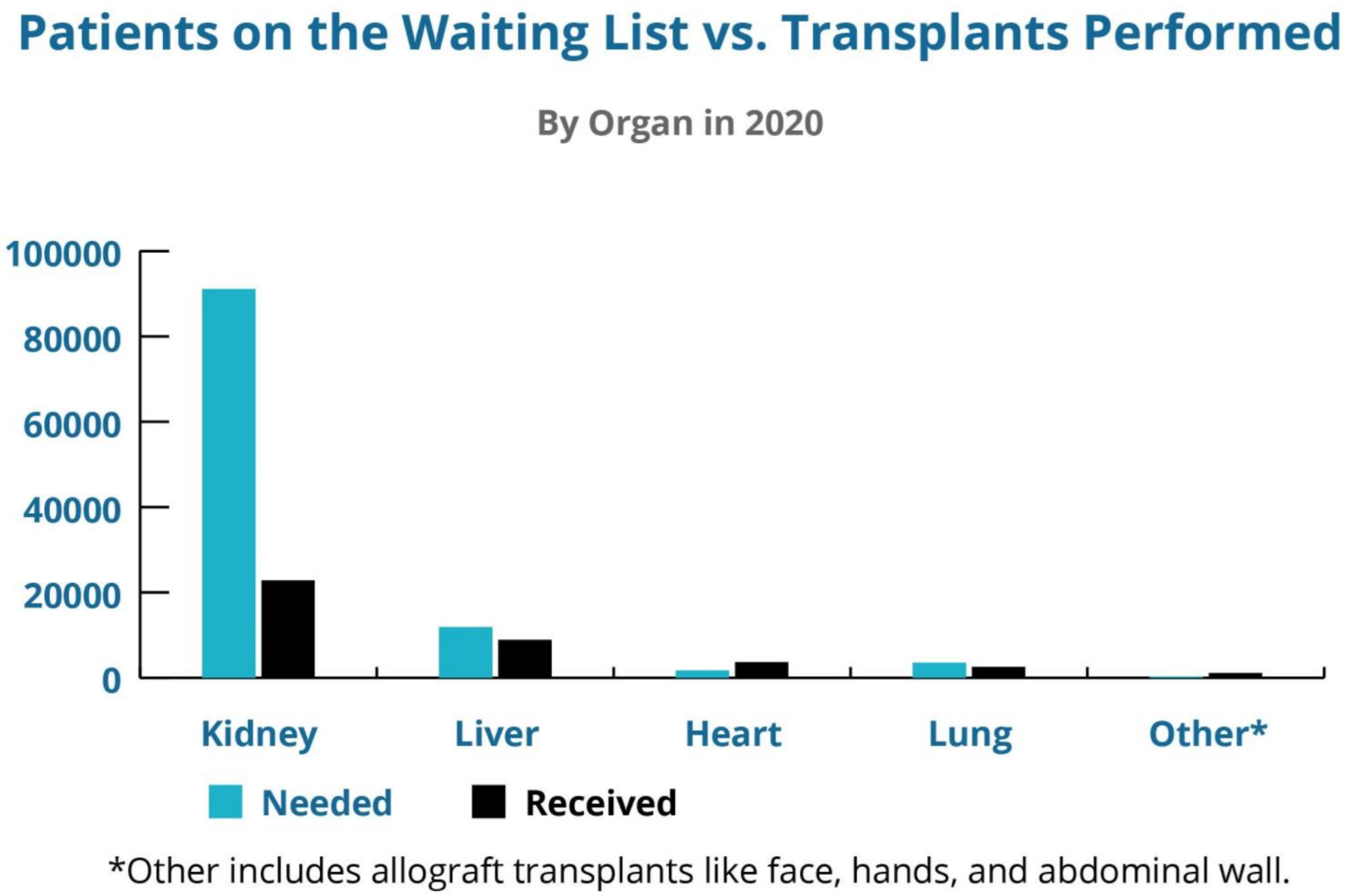
AI in Kidney Acceptance Research Update Webinar

8/1/2023

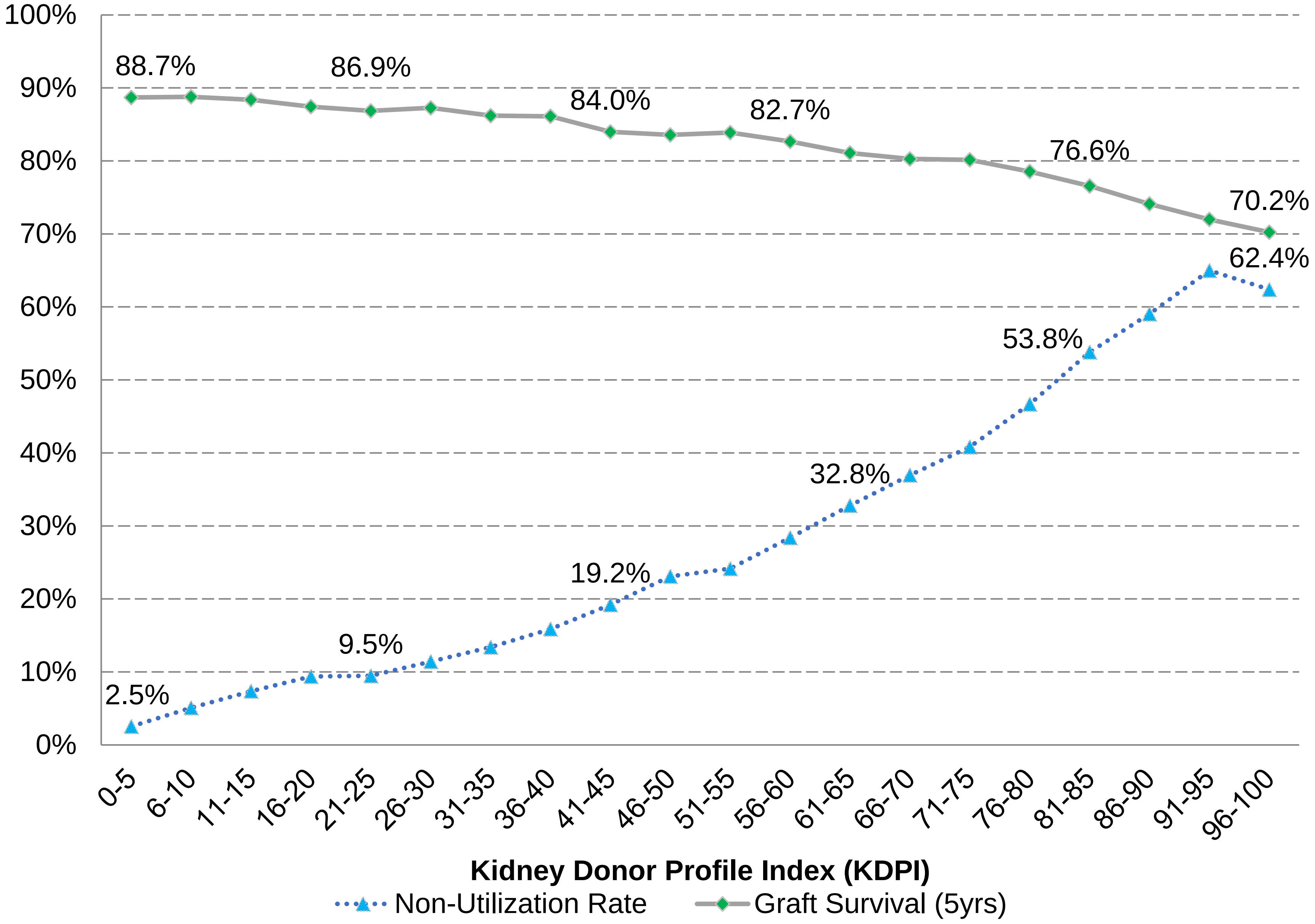
PI: Casey Canfield, Assistant Professor
Engineering Management & Systems Engineering
Missouri University of Science & Technology
canfieldci@mst.edu

Demand for kidney transplants exceeds supply.



- Almost 24,000 candidates received a kidney transplant in 2022
- Almost 90,000 candidates are on the waitlist today
- Every year, thousands of donated kidneys are not used

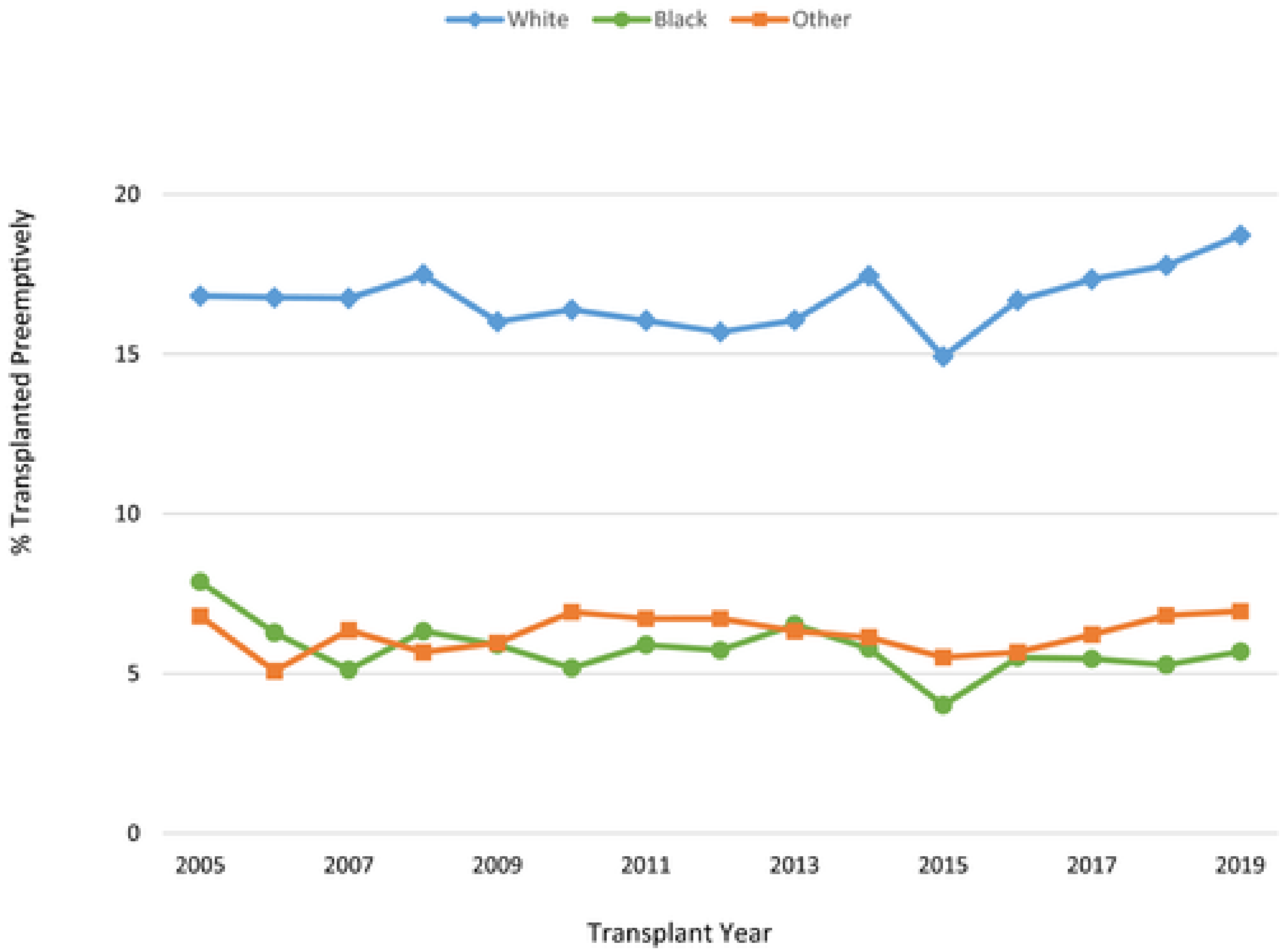
https://www.organdonor.gov/learn/organ-donation-statistics; https://optn.transplant.hrsa.gov/data/



Some unused kidneys are missed opportunities.

Maybe AI can help?

There is unfairness in the existing system, which we need to avoid perpetuating.



Reese, P. P., Mohan, S., King, K. L., Williams, W. W., Potluri, V. S., Harhay, M. N., & Eneanya, N. D. (2021). Racial disparities in preemptive waitlisting and deceased donor kidney transplantation: Ethics and solutions. *American Journal of Transplantation*, *21*(3), 958-967.

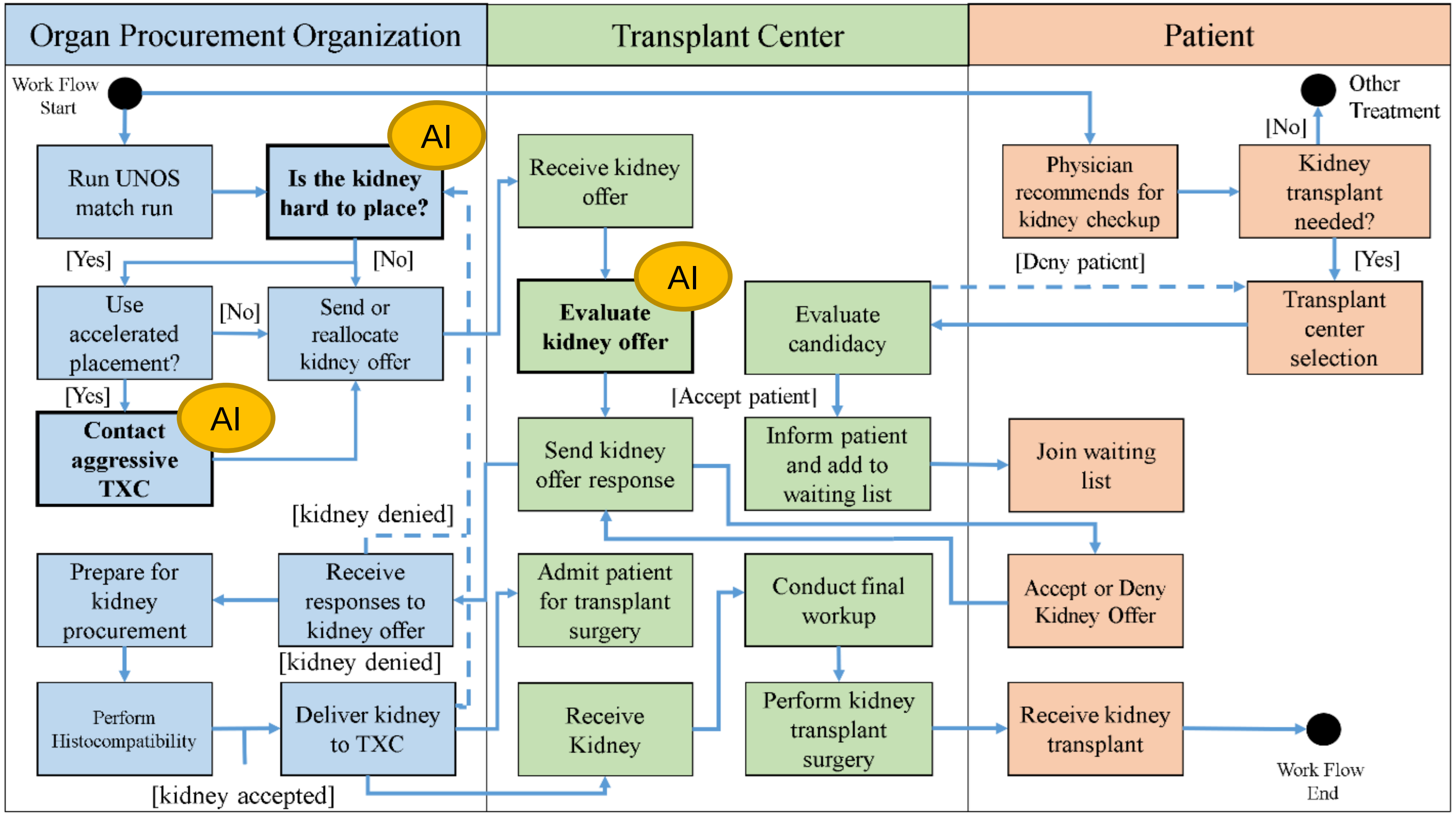
With a planning grant, we created the initial design for an AI system in collaboration with transplant stakeholders.

1. **Interviews**
   - Draft current work-flow
2. **Workshop**
   - Brainstorm system requirements
3. **Workshop**
   - Review system architecture
   - Evaluate mock-up
4. **Interviews**
   - Evaluate mock-up
5. **Workshop**
   - Choose interfaces

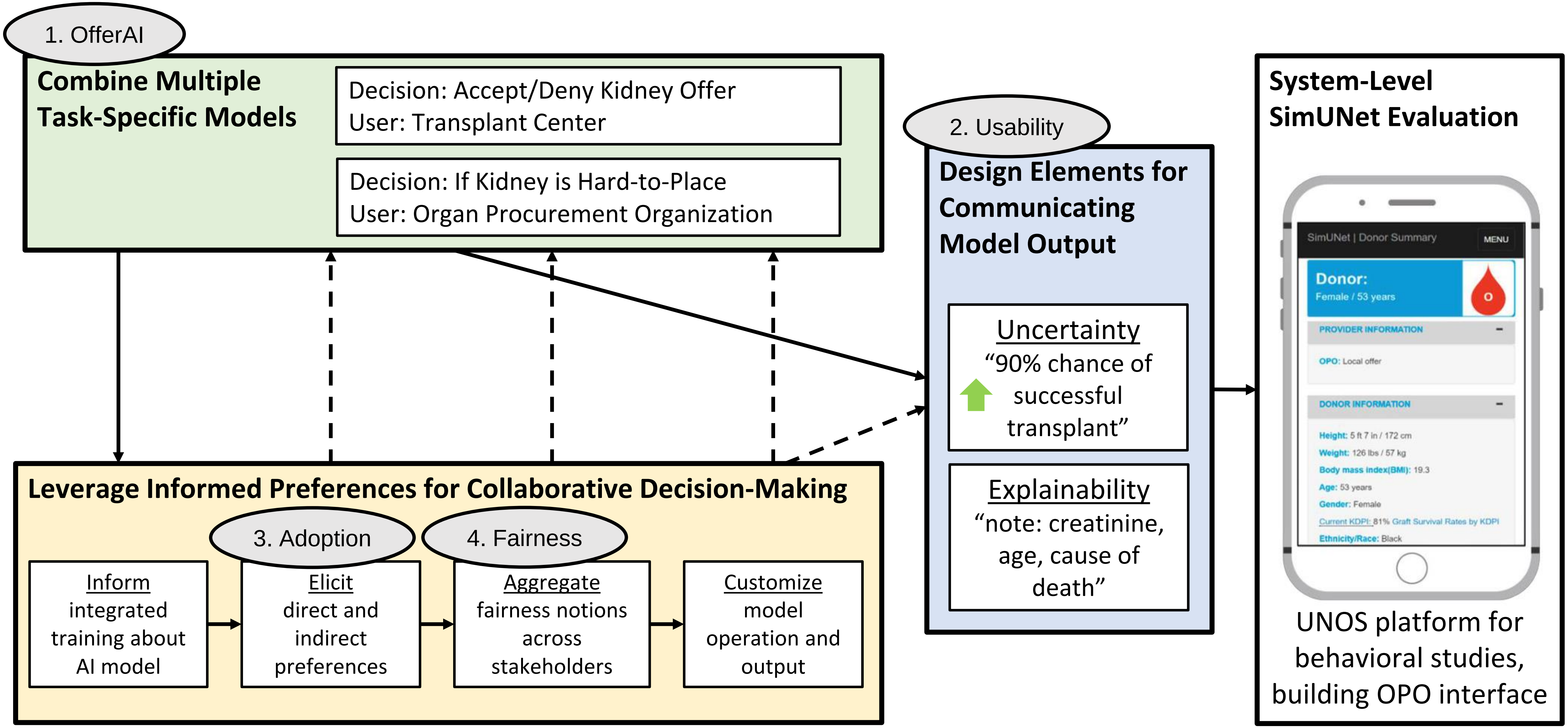
2020-2021. National Science Foundation, FW-HTF-P: Teaming Transplant Professionals and Artificial Intelligence Tools to Reduce Kidney Discard

We are targeting AI support for tension between transplant professionals.

- OPOs want transplant centers to deny kidneys faster
- Transplant centers do not want to miss a successful transplant



Threlkeld, R., Ashiku, L., Canfield, C., Shank, D., Schnitzler, M., Lentine, K., Axelrod, D., Battineni, A. C. R., Randall, H., Dagli, C. (2021). Reducing Kidney Discard with Artificial Intelligence Decision Support: The Need for a Transdisciplinary Systems Approach. *Current Transplantation Reports*, 8, 263-271.



2022-2026. National Science Foundation, Collaborative Research: FW-HTF-R: Embedding Preferences in Adaptable Artificial Intelligence Decision Support for Transplant Healthcare to Reduce Kidney Discard

This work involves a large, interdisciplinary team.

MISSOURI S&T


Casey Canfield (PI)
Eng Mgt & Sys Eng


Cihan Dagli (Co-PI)
Eng Mgt & Sys Eng


Daniel Shank (Co-PI)
Psych


Sid Nadendla
Computer Science

Lirim Ashiku
Richard Threlkeld
Hari Subramanian
Rachel Dzieran
Mukund Telukunta
Elham Babaee
Chase Johnson
Gabriella Stickney
Sukruth Rao

SAINT LOUIS UNIVERSITY


Mark Schnitzler (Co-PI)
Health Economics


Krista Lentine (Co-PI)
Nephrology


Henry Randall
Transplant Surgery


Jason Eberl
Health Care Ethics


Michael Miller
Health Care Ethics

UNOS
UNITED NETWORK FOR ORGAN SHARING


Laura Cartwright
Behavioral Science


Stephanie Rose
UNOS Labs


Samantha Noreen
Research Science


Brendon Cummiskey
Behavioral Science

Please add Questions to the Q&A!

- Agenda:
  - Short updates from 4 sub-teams

- We will respond to questions and comments at the end + along the way in the chat
- The recording and slides will be posted on our website (https://sites.mst.edu/aifortransplant/)

OfferAI

Usability

Adoption

Fairness

Sub-Team

Dr. Cihan Dagli
Co-PI
Systems Engineering

Dr. Lirim Ashiku
Research Associate
Systems Engineering

Richard Threlkeld
PhD Candidate
Systems Engineering

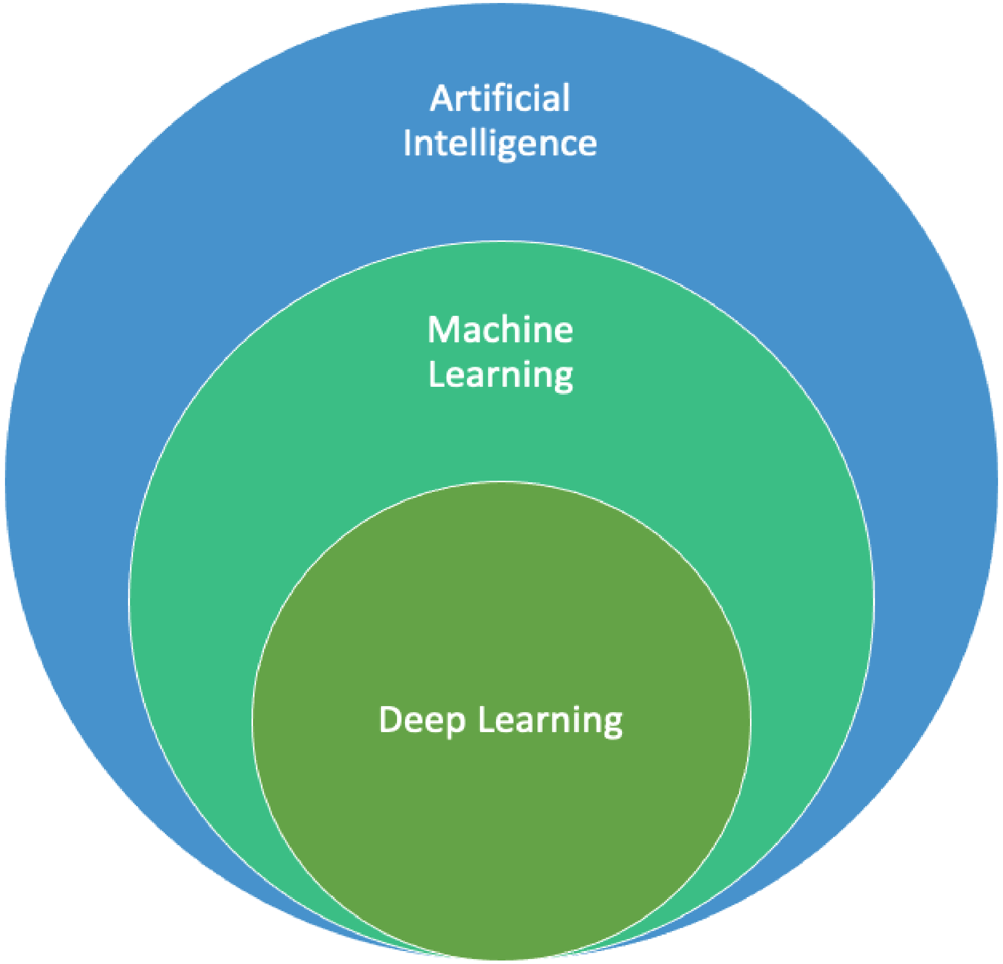
Rachel Dzieran
PhD Student
Systems Engineering

Human-AI Teaming

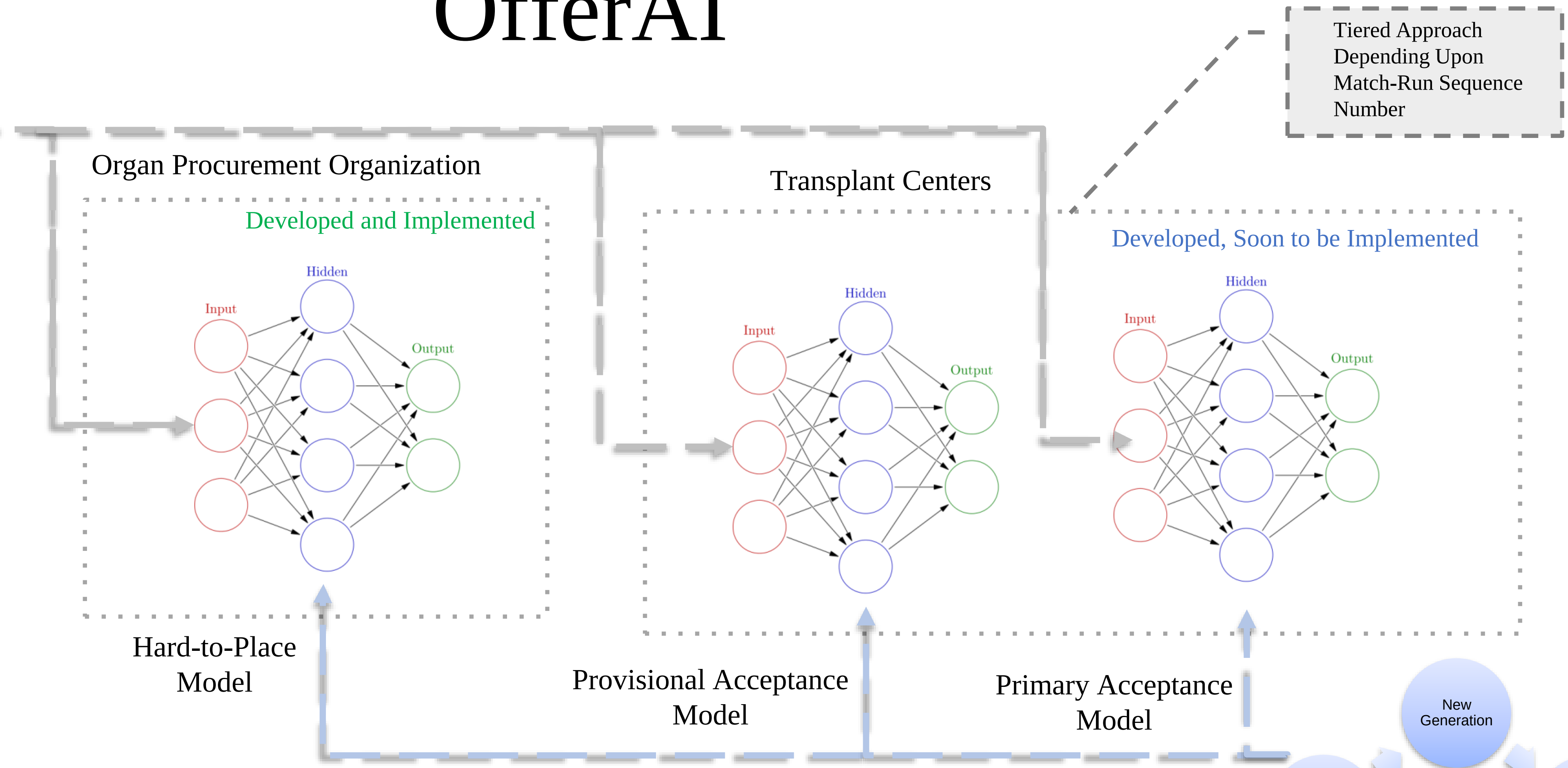
What is Artificial Intelligence (AI)?

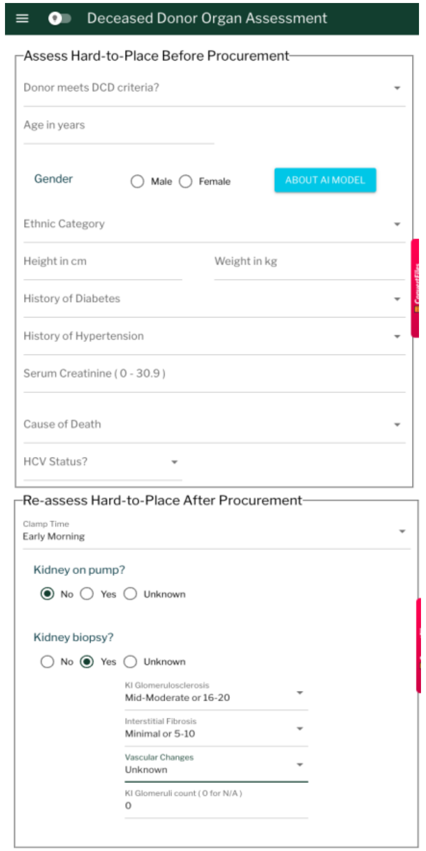
- AI, Machine Learning, and Deep Learning are often used interchangeably to describe algorithms

- Machine Learning
  - can learn and operate on large datasets
  - require shorter training times
- Deep Learning
  - captures non-linear or complex correlations for very larger datasets, including images
  - takes longer to train
  - generally yields better performance



OfferAI




Deceased Donor Organ Assessment

Assess Hard-to-Place Before Procurement

Donor meets DCD criteria?

Age in years

Gender ○ Male ○ Female ABOUT AI MODEL

Ethnic Category

Height in cm

Weight in kg

History of Diabetes

History of Hypertension

Serum Creatinine (0 - 30.9)

Cause of Death

HCV Status?

Re-assess Hard-to-Place After Procurement

Clamp Time
Early Morning

Kidney on pump?
◉ No ○ Yes ○ Unknown

Kidney biopsy?
○ No ◉ Yes ○ Unknown

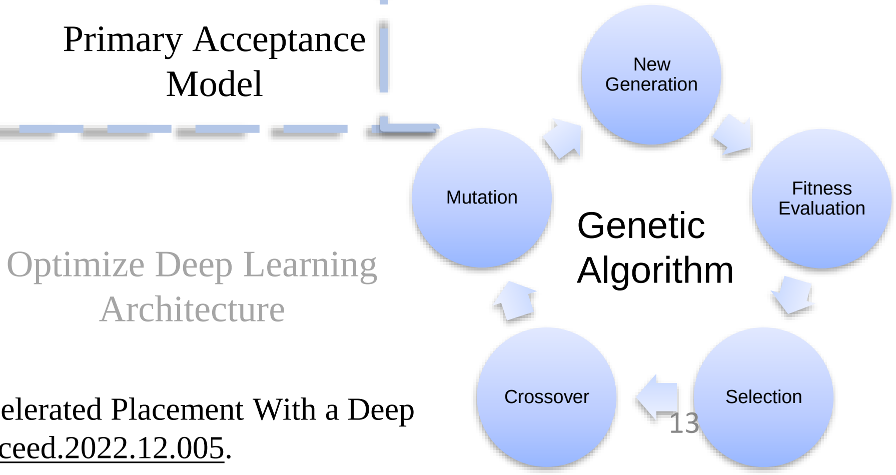
KI Glomerulosclerosis
Mid-Moderate or 16-20

Interstitial Fibrosis
Minimal or 5-10

Vascular Changes
Unknown

KI Glomeruli count (0 for N/A)
0

ddoa.mst.hekademeia.org/#/

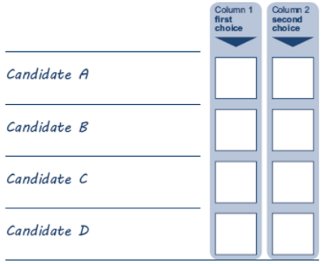


Ashiku L, Dagli C. (2023). Identify Hard-to-Place Kidneys for Early Engagement in Accelerated Placement With a Deep Learning Optimization Approach. *Transplant Proc*. 55(1), 38-48. doi: 10.1016/j.transproceed.2022.12.005.

OPOAI

Organ Procurement and Transplantation Network (OPTN) Policies

Match-Run List



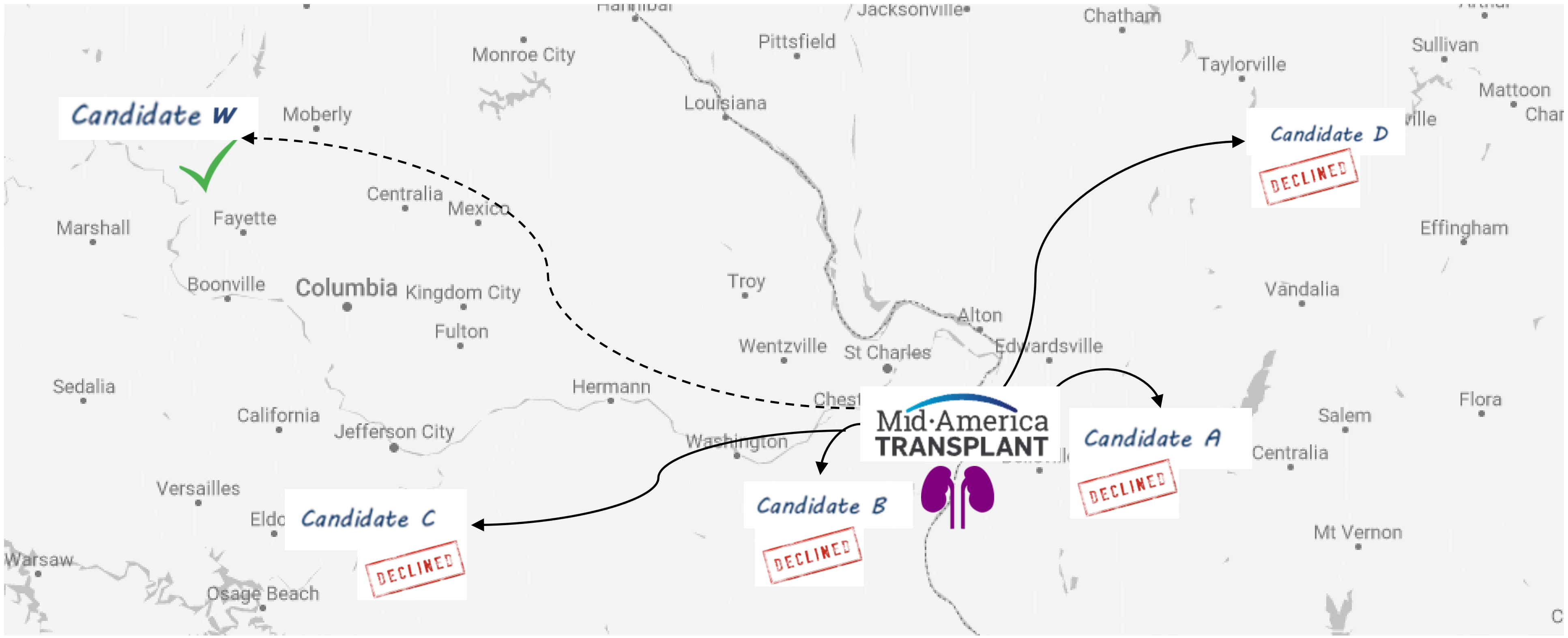
→ Sequence

- - → Out of Sequence

Support decisions over time in simulation before implementation in real life to see the effect on utility and equity.

In Development

Supported by:



- Is the kidney considered high-risk for non-utilization or hard to place?
- Has it been declined by many transplant centers?
- What is the likelihood of it being accepted for transplant?

2022-2024. Mid-America Transplant. AI-Enabled Digital Support to Increase Placement of Hard to Place Deceased Donor Kidneys. PI: Cihan Dagli

Threlkeld, R., Ashiku, L., Dagli, C. (2023). A Use Case for Developing Meta Architectures with Artificial Intelligence and Agent Based Simulation in the Kidney Transplant Complex System of Systems. *18th Annual System of Systems Engineering Conference (SoSe)*, Lille, France, pp. 1-6, doi: 10.1109/SoSE59841.2023.10178576.

Human-AI Teaming Experiments

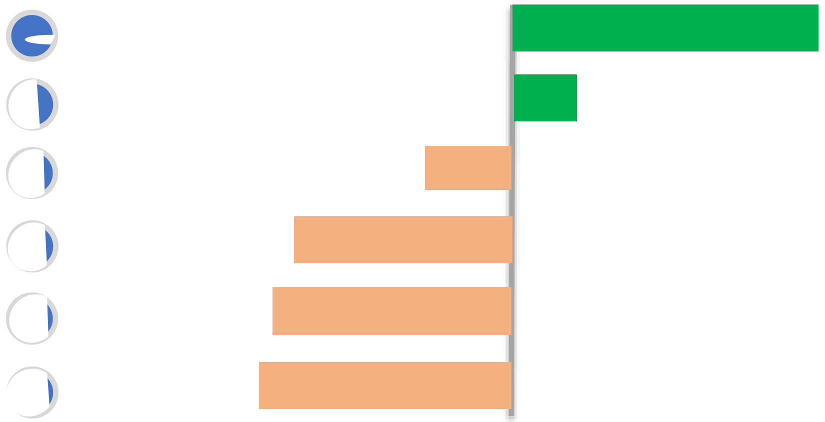
In design and will be conducted in 2024

- Application and measurability of OfferAI models
- Evaluating the inherent bias
- Human-AI teaming experiments conducted via a digital platform based on
  - Algorithms
  - Technology
  - Individual practices & processes

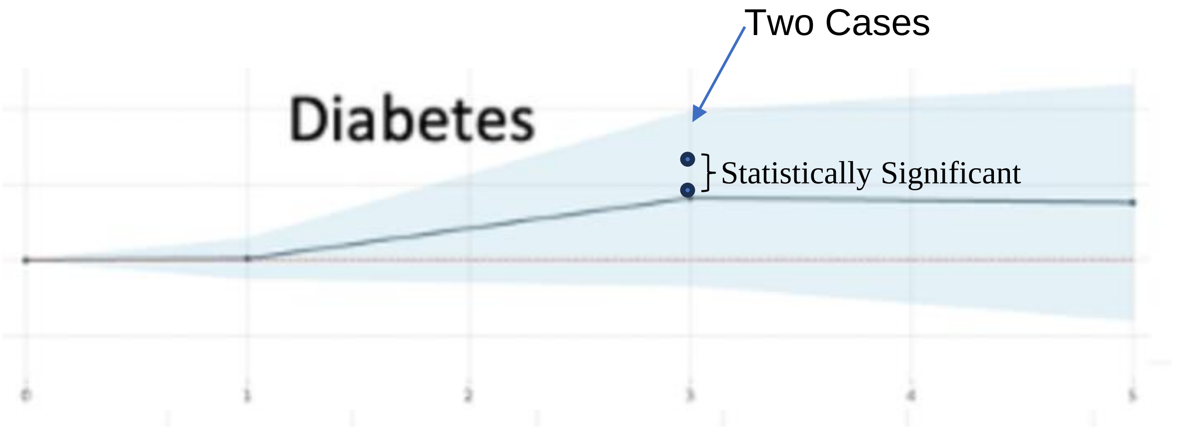
Percentage of Diabetes Donors

0	No
1	0-5
2	6-10
3	> 10
4	Duration Unknown
5	Unknown

Representation



10%	20%	70%
Algorithms	Technology	People & Processes



Duranton S., How humans and AI can work together to create better businesses, ted.com, https://www.ted.com/talks/sylvain_duranton_how_humans_and_ai_can_work_together_to_create_better_businesses/

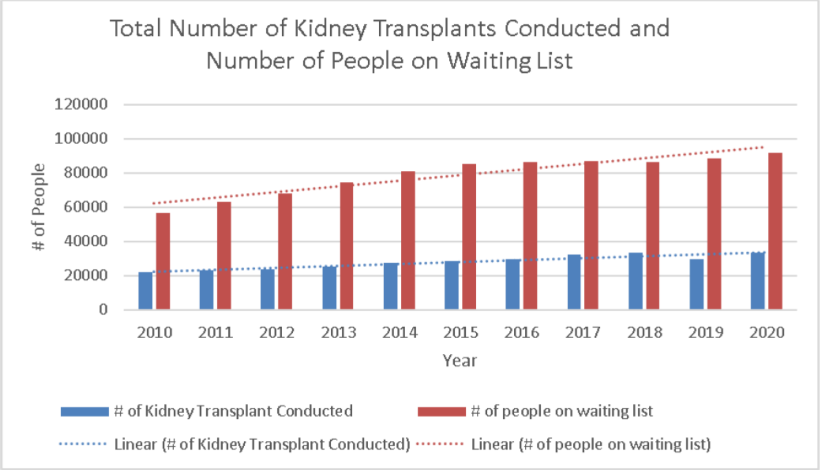
Next steps

- Collaborating with stakeholders to gather feedback to improve our AI and simulation software

- Integrating acceptance model to the online platform

- Conducting platform testing with stakeholders, decision-makers, and users

- Studying human-AI collaboration to understand the decision-making rules and preferences of individuals



Utility Equity

OfferAI

Usability

Adoption

Fairness

Sub-Team

Casey Canfield (PI)
Eng Mgt & Sys Eng

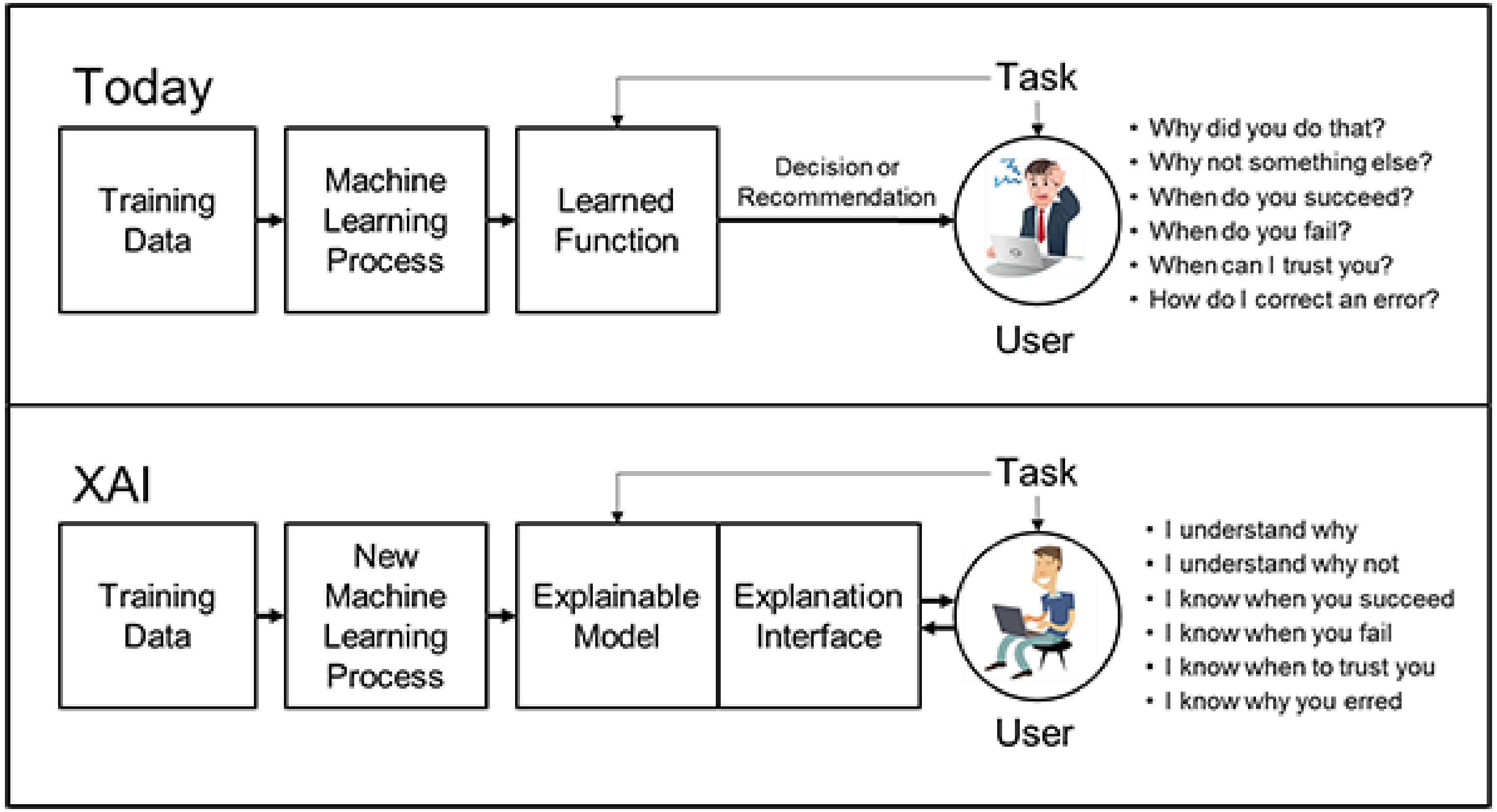
Daniel Shank (Co-PI)
Psych

Hari Subramanian
PhD Student
Engineering Management

An effective AI system should be explainable

Explainable AI (XAI) helps users:

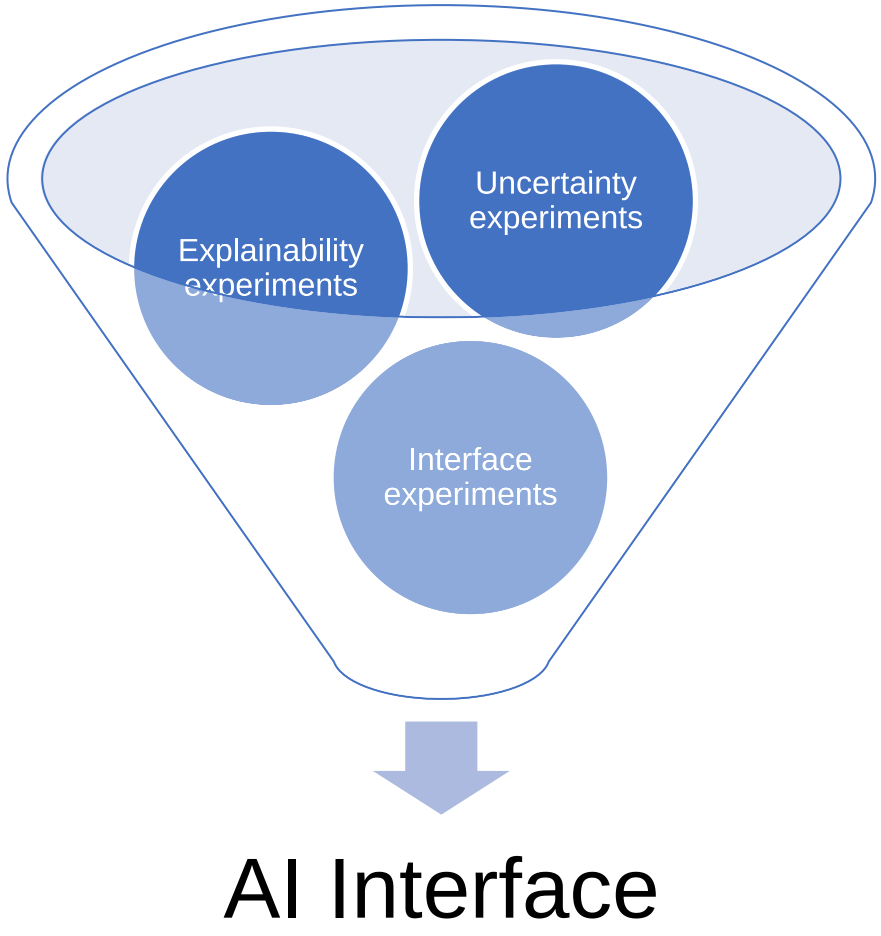
- Understand system’s process and logic
- Appropriately trust the system
- Effectively manage performance

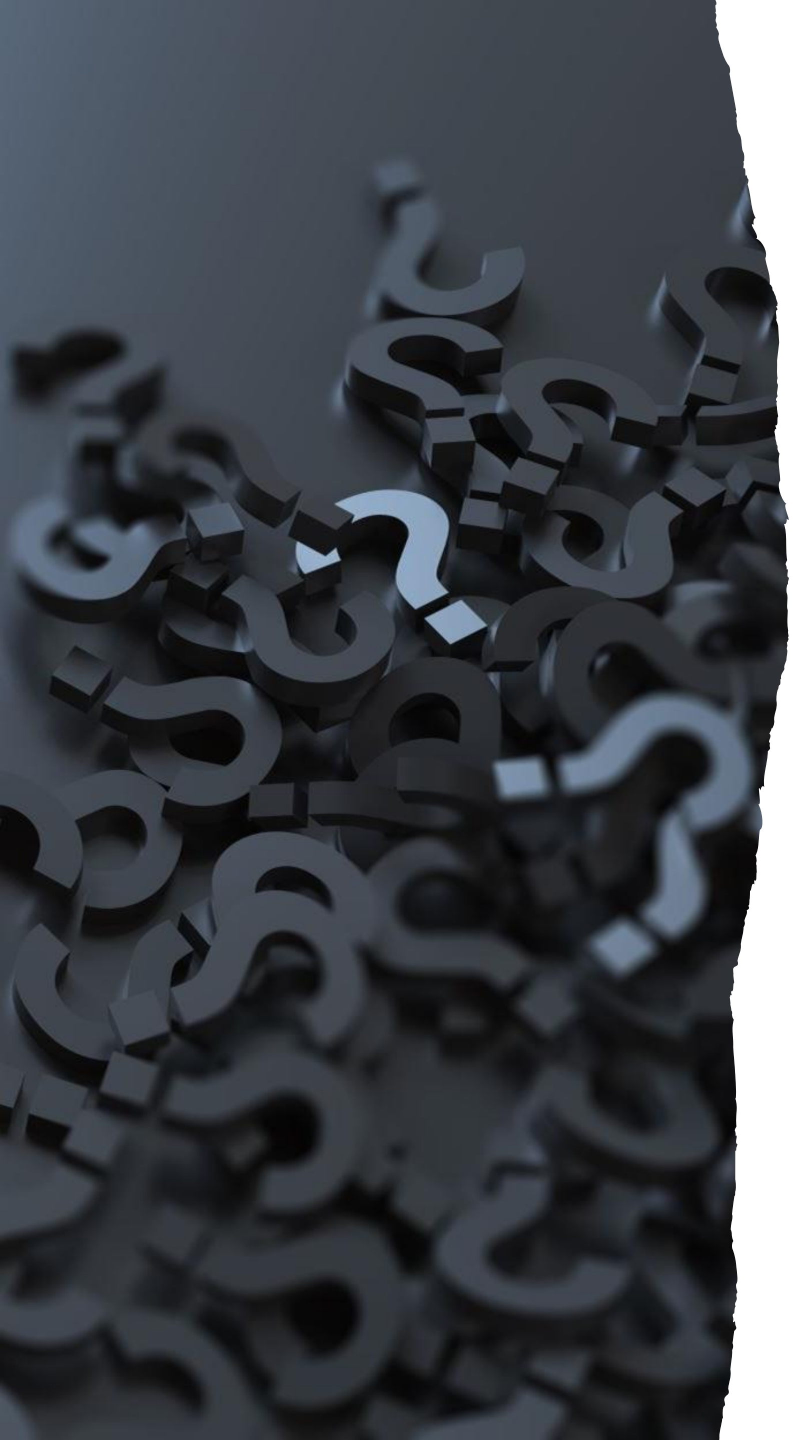


https://www.darpa.mil/program/explainable-artificial-intelligence

Goals of Usable Communication

- Examine user reactions to:
  - AI explainability and transparency: what does the AI know?
  - Uncertainty representations: what does the AI not know?
- Evaluate changes in user trust, confidence, performance, and adoption
- Design interfaces for an AI tool for kidney placement decisions





Designing Explainable AI (XAI) to Improve Human-AI Team Performance: A Medical Stakeholder-Driven Scoping Review

- Contextual use of AI predictions
- Information included in AI predictions
- Personalization of AI prediction for different groups
- Customizing AI prediction for specific cases

User trust will influence adoption and use of OfferAI

System-level

- Helps users develop a **mental model** of the system

(Cai et al. 2019; Gates and Leake 2021)

Prediction-level

- Helps users make decisions and understand the **credibility** and utility of the system

(Barda, Horvat, and Hochheiser 2020)

Barda, Horvat, Hochheiser. (2020). A qualitative research framework for the design of user-centered displays of explanations for machine learning model predictions in healthcare. *BMC Med. Inform. Decis. Mak,* 20(1). doi: 10.1186/s12911-020-01276-x.

Cai, Winter, Steiner, Wilcox, Terry. (2019). “Hello AI”: Uncovering the Onboarding Needs of Medical Practitioners for Human-AI Collaborative Decision-Making. *Proc. ACM Human-Computer Interact.*, 3(CSCW), 1–24. doi: 10.1145/3359206.

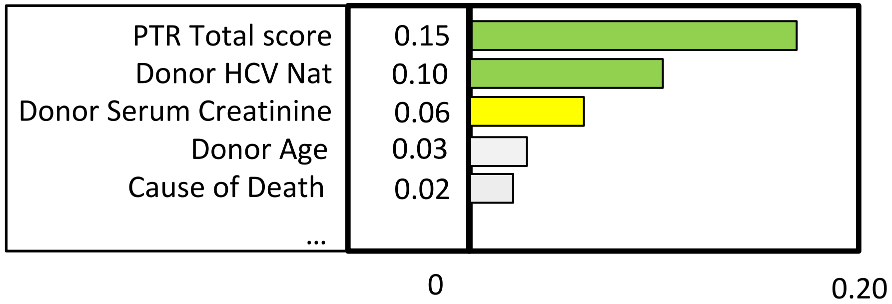
Gates, Leake. (2021). Evaluating CBR explanation capabilities: Survey and next steps. *CEUR Workshop Proceedings*, 3017, 40–51. https://ceur-ws.org/Vol-3017/97.pdf

System-level information should include measures beyond accuracy

Confusion Matrix

	Actual Transplant	Actual Not Transplanted
Predicted Transplant	**5272**	**343**
Predicted Not Transplant	**49**	**5170**

Feature Importance



Training Information

Training Dataset Information	
Dataset Timeline	2016-2021
No. of observations	1,300,000
No. of features	29
…	…

Effect of prediction-level XAI on trust depends on the task and individual

Counterfactual/Contrastive

For this use-case, our model would have made the opposite prediction (i.e., predict "Not Transplant") in each of the following cases:

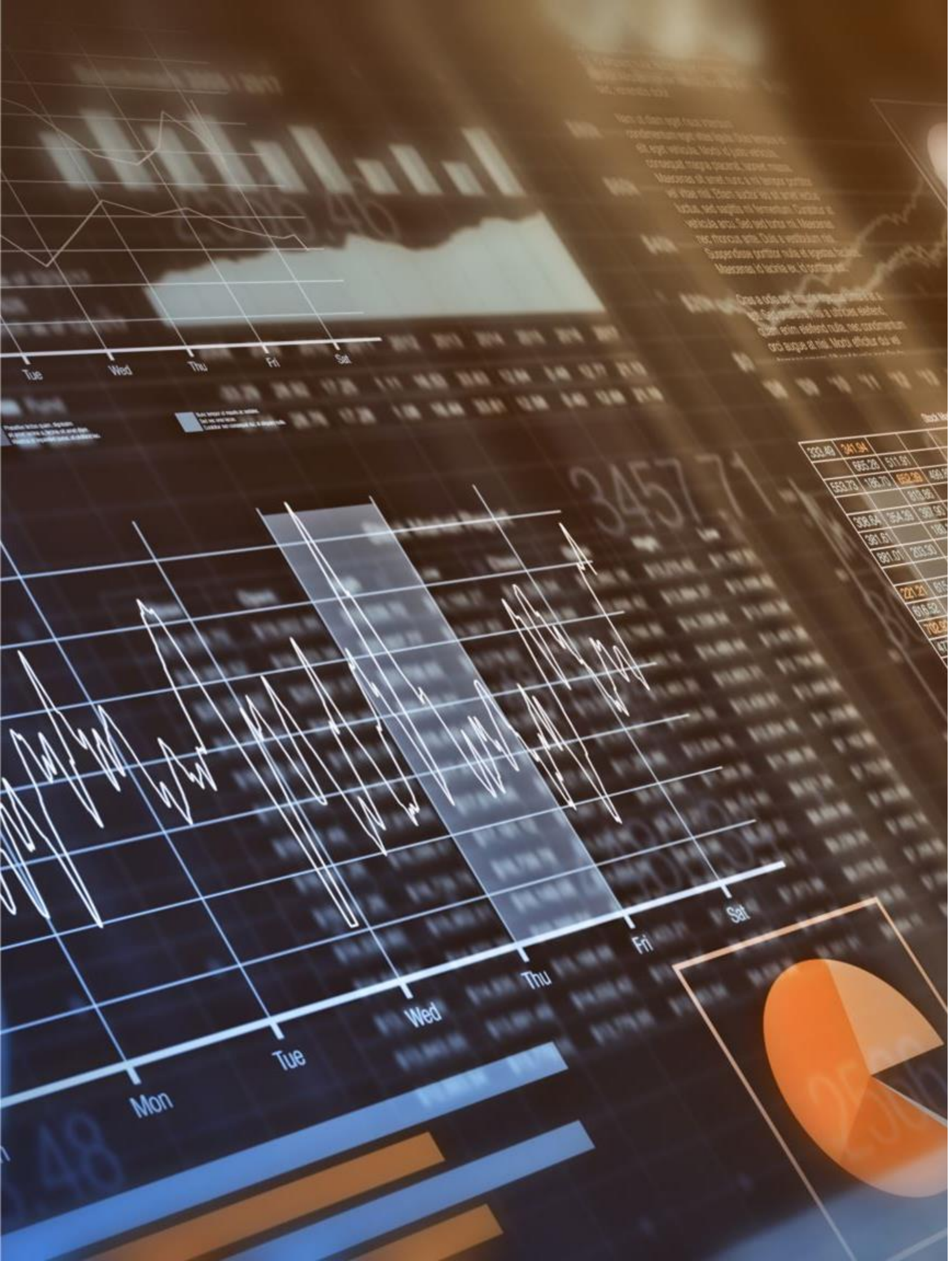
- Donor Serum Creatinine: If the donor serum creatinine had been **2.0** instead of 1.1
- Donor Age: If the donor's age had been **75** instead of 65
- Donor HCV Nat: If the donor's HCV Nat had been **Positive** instead of Negative

Nearest Neighbors

	Current offer	Alternate offer A	Alternate offer B
Prediction	**Transplant**	**Transplant**	**Not Transplant**
Donor Serum Creatinine	1.1	1.6	1.1
Donor Age	65	65	65
Donor HCV Nat	Negative	Negative	Positive
…			

Users want more XAI information for difficult decisions.

- Pathologists preferred AI system to provide **both conservative and liberal predictions** for difficult decisions (Cai et al. 2019).
- Different expertise levels need **different amounts of information** with experts wanting less (Merry et al., 2021).
- User **customization and interaction** improves initial adoption and trust (Barda, Horvat, and Hochheiser 2020).



Next Steps: Experimental Work

How is trust and performance affected when:

- system-level information is provided with and without prediction-level information?
- users customize their XAI interface?
- XAI is tailored based on user expertise?

Analogous task: matching customer with used car.

Pre-requisites:

- Measurable expertise
- Tabulated information
- Predictable outcome
- Subjective truth

Car Features			
Manufacturer	Hyundai	**Shift (auto/manual)**	Auto
Model	Tucson	**Miles**	100,452
Color	Black	**City Mileage**	20 mpg
Back-up camera	Yes	**Highway Mileage**	25 mpg
Doors	4	**Cylinder**	4
Seats	5	**Inspection**	Passed
Rear air vents	Yes	**Power controlled seats**	2
Heated seats	2	**Hands free calling**	Yes
Apple CarPlay	No	**Bluetooth**	Yes
Accident history	No		

Potential Customer			
Age	24	**Budget**	12,000
Gender	F	**Accidental History**	0
…	…	**…**	…

AI prediction: **Accept (80%)**
Based on the information provided, the car is likely to be purchased by this customer.

Please Share your thoughts!!!

OfferAI

Usability

Adoption

Fairness

Sub-Team

Casey Canfield (PI)
Eng Mgt & Sys Eng

Daniel Shank (Co-PI)
Psych

Amaneh (Elham) Babaee
MS Student
Industrial/Organizational Psych

Jason Eberl
Health Care Ethics

Michael Miller
Health Care Ethics

Uniqueness Neglect

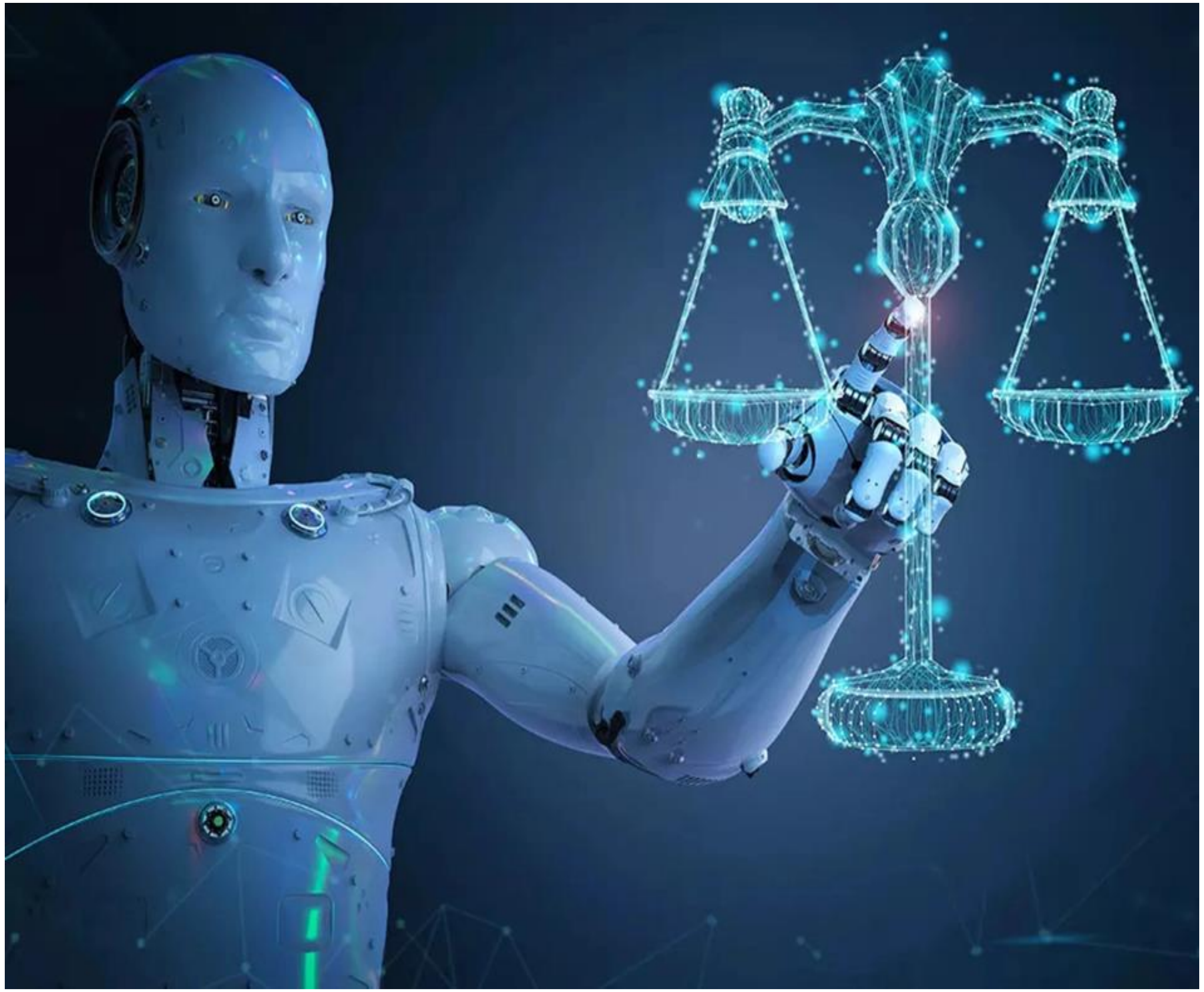
Previous research on "uniqueness neglect" suggests that people resist AI systems that fail to recognize the uniqueness of the user (Longoni et al., 2019).

Longoni, C., Bonezzi, A., & Morewedge, C. K. (2019). Resistance to medical artificial intelligence. *Journal of Consumer Research*, 46(4), 629-650

Trust in AI systems

- AI is perceived as morally wrong if it errs or is biased, unfair, or unethical (Shank et al., 2019).

- However, when users can make **slight changes to the AI after it errs**, those users continue to use it.



Shank, D. B., DeSanti, A., & Maninger, T. (2019). When are Artificial Intelligence versus Human Agents Faulted for Wrongdoing? Moral Attributions after Individual and Joint Decisions. *Information, Communication, and Society*, 22(5), 648-663.

Adoption of OfferAI

- OfferAI needs to be responsive to user preferences to support adoption

- Little work has focused on integrating diverse sources of preference information to minimize human effort (Zhang et al., 2019)

Zhang, S., Yao, L., Sun, A., & Tay, Y. (2019). Deep learning based recommender system: A survey and new perspectives. ACM Computing Surveys (CSUR), 52(1), 1-38.

Development Stages

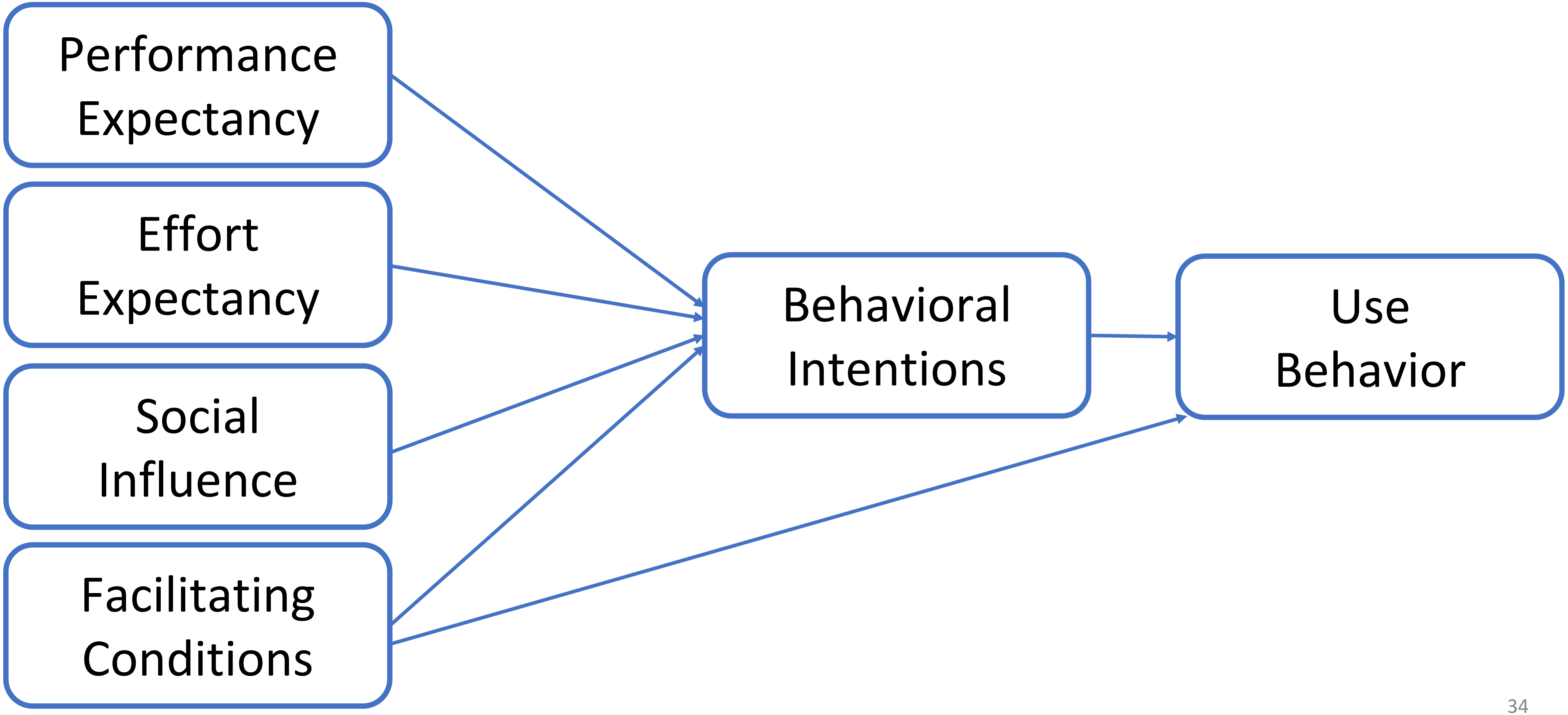
Develop Stakeholder Survey

Develop System-Level Training Materials

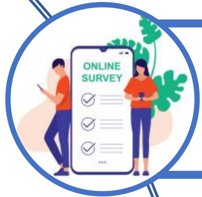
Conduct Customization Experiments

Integrate Preference Learning in OfferAI

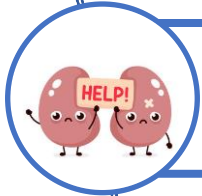
We are building on Technology Adoption theories



Stakeholder Survey

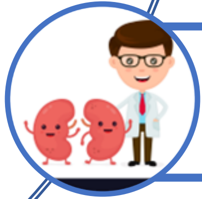


Public



Candidates/Recipients

OPO Staff



Transplant Surgeons

3 parts of the survey:

1. Estimate willingness to adopt OfferAI
2. Evaluate ethics
3. Report demographics

Note: This is 1 of 2 surveys that will be distributed

Proposed Recruitment Strategy

	Public	Patients	OPO Staff	Transplant Surgeons
Target Groups		Candidates on waiting list, transplant recipients	Involved in allocation	Involved in offer decisions
Recruitment	Prolific (online participant platform)	American Association of Kidney Patients (AAKP) National Kidney Foundation (NKF)	Association of Organ Procurement Organizations (AOPO)	American Society of Transplant Surgeons (ASTS) American Society of Transplantation (AST)

Please let us know if there are other organizations we should reach out to!

OfferAI

Usability

Adoption

Fairness

Sub-Team

Sid Nadendla
Computer Science

Mukund Telukunta
PhD Student
Computer Science

Gabriella Stickney
Interns

Sukruth Rao
Interns

Discrimination in Kidney Placement

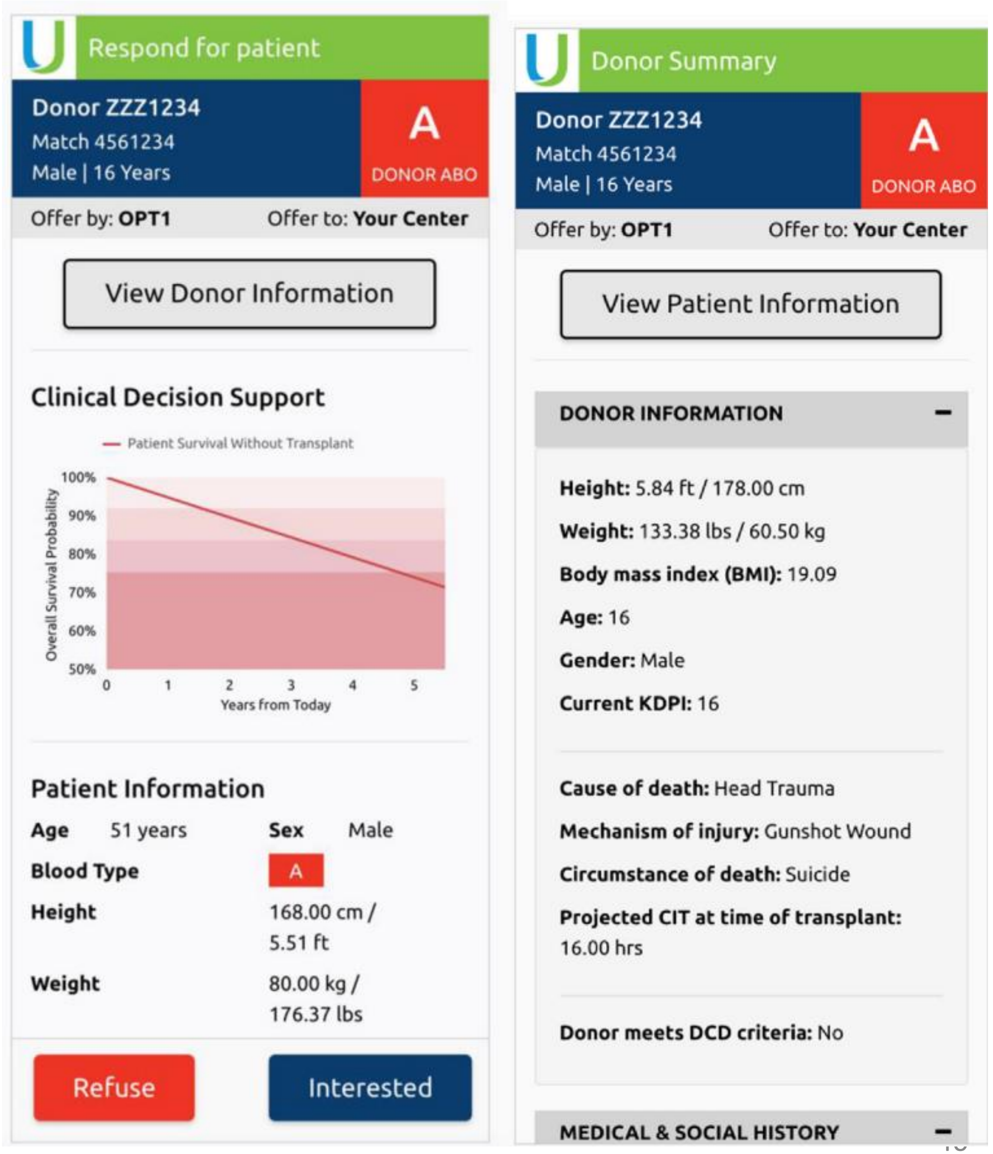
- Racial discrimination
  - calculation of creatinine-based estimated glomerular filtration rate (eGFR)
    - Black patients had higher creatinine levels (Chong et al., 2021)
  - calculation of the Kidney Donor Risk Index (KDRI)
    - Black donors have historically been shown to be associated with lower allograft quality (Chong et al., 2021)
- Gender discrimination
  - women with a primary diagnosis of diabetes was higher than males (Wells, 2009)
- Location discrimination
  - mandatory medical evaluation is only available in tertiary care centers (Tonelli et al., 2006)

Chong, Kelly, et al. (2021). Reconsidering donor race in predicting allograft and patient survival among kidney transplant recipients. *Kidney360,* 2(11).
Wells, Shirley A. (2009). Health disparities in kidney transplantation: an equity analysis. *Journal of Health Disparities Research and Practice,* 3(2), 1.
Tonelli, Marcello, et al. (2006). Residence location and likelihood of kidney transplantation. *Cmaj,* 175(5), 478-482.

AI-based Decision Support in Kidney Placement

UNOS's Predictive Analytics

- Time to better offer and Patient survival without transplant
- Deployed in January 2023
- Cox Hazards model trained to predict patient outcomes from historical data
  - Inherits biases and discrimination present in legacy kidney placement

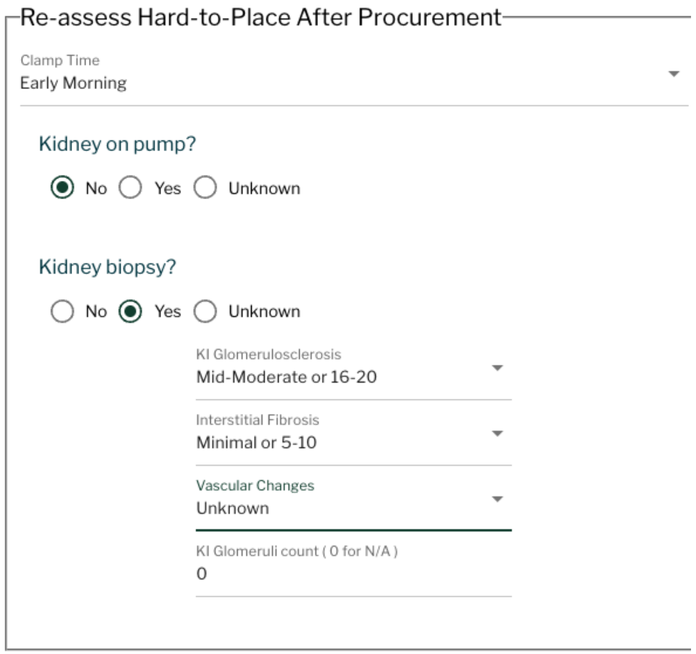


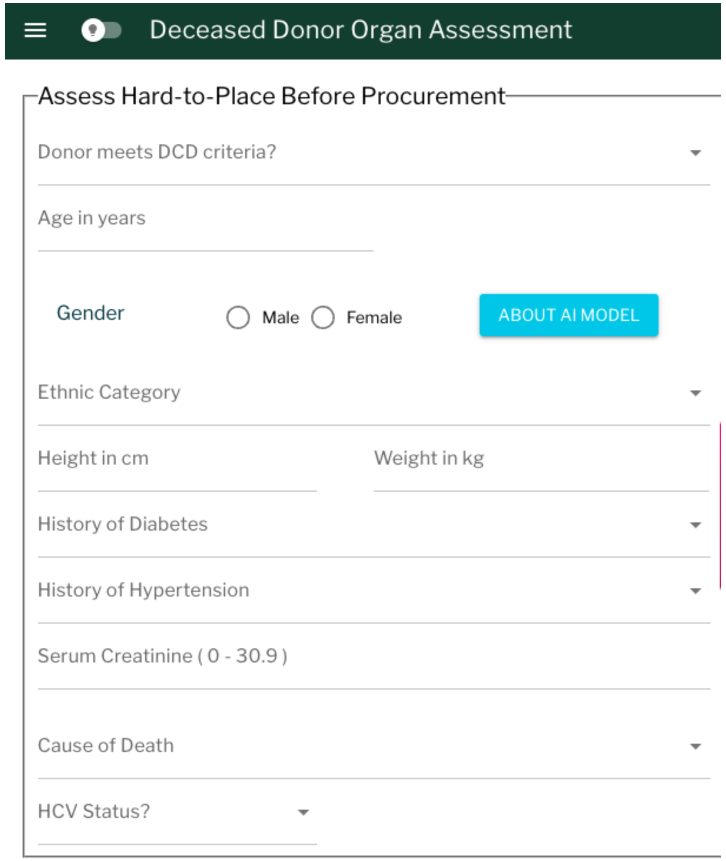
McCulloh, I. et al. (2023). An experiment on the impact of predictive analytics on kidney offers acceptance decisions." *American Journal of Transplantation*.

AI-based Decision Support in Kidney Placement (contd.)

Missouri S&T's OfferAI (ddoa.mst.hekademeia.org/#/)

- Neural Network based model trained to predict decisions made by transplant centers
  - Inherits biases and discrimination present in historic data





Ashiku, Dagli. (2023). Identify Hard-to-Place Kidneys for Early Engagement in Accelerated Placement With a Deep Learning Optimization Approach. *Transplantation Proceedings*, 55(1).

Measuring Discrimination in Decision Support Systems

- **Discrimination**: Disparate Treatment vs. Disparate Impact
- *Disparate Treatment*: Are similar individuals treated similarly?
  - **Individual Fairness**: How can we measure similarity between individuals?
- *Disparate Impact*: Does individuals across different social groups have similar outcomes?
  - **Group Fairness**: What parity (i.e. statistical measure) is appropriate for a given application domain (e.g. kidney placement)?

Accuracy-Fairness Tradeoff: Decision Correctness vs. Social Discrimination

Is it feasible to mitigate discrimination with little/no compromise on decision correctness?

Dwork, Cynthia, et al. (2012). Fairness through awareness. *Proceedings of the 3rd innovations in theoretical computer science conference*.

Our Focus: Group Fairness in Kidney Placement

Equal Opportunity: $TP_A = TP_{A'}$

Equal Accuracy: $(TP_A + TN_A) = (TP_{A'} + TN_{A'})$

Predictive Equality: $FN_A = FN_{A'}$

and many more…

Fundamental Trade Offs: not all fairness notions can be guaranteed at the same time unless

- the base rates are equal or
- the classifier is perfect

		Final Transplant Decision	
		Transplant	No Transplant
OfferAI Prediction	Transplant	TP_A	FP_A
	No Transplant	FN_A	TN_A

		Final Transplant Decision	
		Transplant	No Transplant
OfferAI Prediction	Transplant	$TP_{A'}$	$FP_{A'}$
	No Transplant	$FN_{A'}$	$TN_{A'}$

Kleinberg, Mullainathan, Raghavan. (2017). Inherent Trade-Offs in the Fair Determination of Risk Scores. *8th Innovations in Theoretical Computer Science Conference*, 43(1).

Chouldechova. (2017). Fair prediction with disparate impact: A study of bias in recidivism prediction instruments. *Big Data.*

Fairness Opinions from Diverse Stakeholders

Clinical Experience

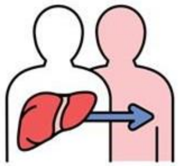
- Surgeons
- OPOs

Characteristics:

- Extensive clinical expertise
- Recommends final decision regarding the kidney offer
- Feasible to collect *exact* classification

Personal Experience

- Patients
- Donors
- Family and friends



Characteristics:

- Rich personal experience from being involved in kidney placement process
- Ultimately decides based on surgeon's recommendation
- Not feasible to collect exact classification

Our Methodology

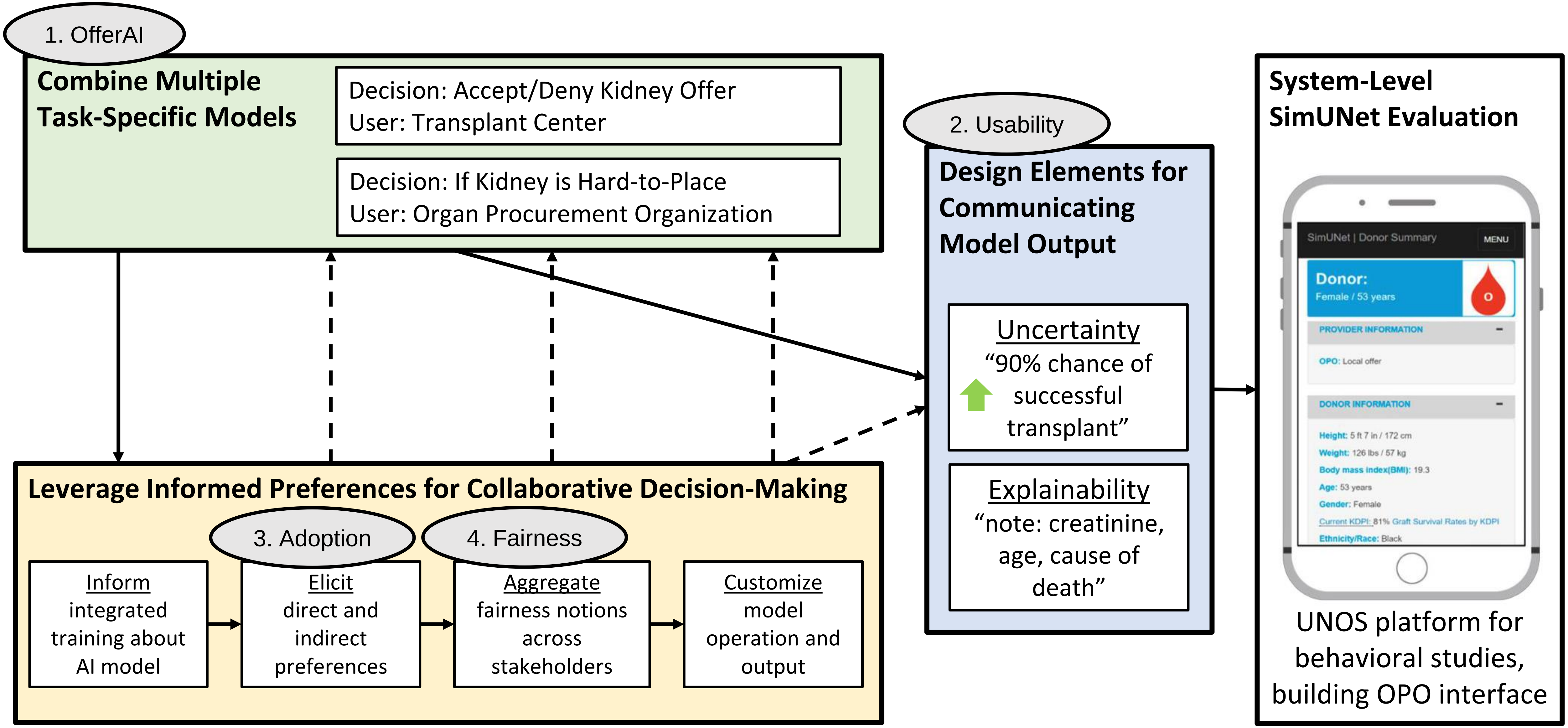
1. *Design survey and elicit feedback responses from diverse stakeholders*
2. *Estimate preferences over group fairness notion from elicited fairness feedback*
3. *Aggregate preferences using diverse voting rules*

OfferAI

Usability

Adoption

Fairness



2022-2026. National Science Foundation, Collaborative Research: FW-HTF-R: Embedding Preferences in Adaptable Artificial Intelligence Decision Support for Transplant Healthcare to Reduce Kidney Discard

Please Share Your Questions & Comments!

Stay Tuned:

Participate in the 2 Stakeholder Surveys (Adoption + Fairness)

Try out OfferAI

2024 Ethics Symposium

Funding: National Science Foundation (#2222801 and #2026324)

More Information: https://sites.mst.edu/aifortransplant/

Contact: Casey Canfield, canfieldci@mst.edu

Discussion

- What parts of the proposed research are confusing?
- To what extent do you think OfferAI will increase kidney utilization? Increase equity and fairness?
- What would make you more or less comfortable with OfferAI being implemented in the transplant system?
- Are there other issues/challenges that we should be incorporating into this research?