H Hybrid Intelligence

NWO

Project #22:
Articulating power relations through affordance theory
for the design of AI systems for public administration

TU Delft   UMEÅ UNIVERSITY

PhD candidate: Íñigo Martínez de Rituerto de Troya[1]   [1] TU Delft

    Promotors: Neelke Doorn[1], Virginia Dignum[1][2]   [2] Umeå University
 Co-promotor: Roel Dobbe[1]

1. Introduction

1.1 AI in public policy: promises & problems

Data science is increasingly being used to address challenges in public policy. Unfortunately, many such attempts have had **undesirable social consequences**:
> criminal justice:
  racial disparities in predictive policing, facial recognition, recidivism prediction
> child welfare services:
  misidentification of fraudulent child benefits claims (see: Toeslagenaffaire)
> public employment services:
  predicting risk of long-term unemployment rather than proactively recommending interventions

1.2 Power relations in technology: critiques & their shortcomings

Power is a central issue in contemporary discussions about the effects of technology on society.

How can data-driven insights **inform** and **enable** better policy?
How can we ensure affected stakeholders **have a voice**?
How do social media platforms **influence** political discourse?
How can tech workers organise to **influence** decisions within big tech?
How is state **control** facilitated by surveillance technology?

However, critiques often…
… **lack a precise vocabulary** with which to articulate power relations;
… **are not situated** in the practical and institutional realities of the sociotechnical systems;
… **do not point towards constructive responses** to unjust power relations.

1.3 Research Questions

RQ1.a. <u>What</u> power relations are encoded by the sociotechnical affordances of AI systems in public policy contexts?
   b. <u>How</u> can we describe the power relations of a system in terms of its affordances?

RQ2.   <u>How</u> do the practices of system designers reproduce or reconfigure these power relations?

RQ3.   <u>How can</u> a framework for articulating power relations in terms of affordances help designers to build more just systems (and develop institutions that support those efforts)?

2. Theoretical Grounding

2.1 Affordance Theory

Affordance theory originated in perceptual psychology to describe how properties of an environment are perceived by animals as **enabling and constraining different possible actions**. It has since become a prominent concept in literature on design and human-computer interaction.

> "Technologies don't *make* people do things but instead, *push, pull, enable*, and *constrain*. Affordances are *how* objects shape action for socially situated subjects."
> — p.6, Jenny Davis (2020) How Artifacts Afford: The Power and Politics of Everyday Things. MIT Press.

I want to explore how technological affordances (as well as social affordances, institutional affordances, etc.) can be used **to articulate power relations in terms of system specifications**.

2.2 Epistemologies of Power

> **Political philosophy** has described power analytically as *expressions* (A has *power over* B; A has *power to* do X, etc.) and *dimensions* (power to make decisions, to set agendas, to control awareness about relevant issues, to constitute social roles).
> **Sociology** has described how social structures reproduce power relations.
> **Feminist studies** has described how power is experienced by oppressed minorities.

3. Research Plan

3.1 RQ1 plan: literature review & case study

We will conduct a **literature review** to synthesize current theories about power and affordances into an interoperable ontology.

A hypothetical **case study** (see **4.**) will help us to develop a **prototype diagnostic framework** to describe power/affordances.

3.2 RQ2 plan: interviews with practitioners

We will conduct **interviews** with practitioners (data scientists, software developers, managers, PIs, policymakers) to understand how they configure the power relations of the use case through their practices and institutional cultures, e.g.
   "How are users empowered or constrained?"
   "How are decisions made about technical specifications and social interactions?"
   "How can system designers and users dissent constructively?"

3.3 RQ3 plan: diagnostic & design framework

We will work together with public sector partners to develop and evaluate the framework in practice. The framework should:
(i) **enable a diagnosis of problematic power relations** in terms of technical and institutional affordances of the system, and
(ii) point towards ways of **resolving these tensions**.

4. Case Study

Reorienting public employment services from risk prediction to intervention recommendation.

Risk prediction: dispositive

Caseworkers make **"forward"** judgements of a predictive model (low/med/high risk) to a jobseeker when planning their interventions.

Jobseeker   Caseworker   Model

Recommendation: collaborative

Caseworkers can **work together** with jobseekers to synthesize personal development plans through a **shared interface**.