Accountability by Dialogue: a new approach to data protection

Joris Hulstijn,
Tilburg School of Economics and Management
j.hulstijn@uvt.nl

ackn: Stéphanie van Gulijk (TLS)

Example

- **ANWB Drive Safely** is a car insurance offered by Dutch motorist association ANWB. The driver gets a discount on the premiums when driving safely. The safer the driving, the higher the discount.
- Driving behavior is measured by a device that is placed in the car, connected to the car management system. It measures four driving aspects: speed, acceleration, slowing down and taking turns. The assessment of these four aspects is converted into a driving score. Only the score is sent to the insurance company's servers.
- The discount is based on the average driving score calculated over three months. Safe driving is defined as follows: the customer should observe the rules of the road (such as speed limits), adapt to other road users, adapt to circumstances (e.g. construction works or slippery road), take turns calmly, accelerate smoothly and do not brake abruptly.

Example

- This Drive Safely initiative sounds like a good idea. Such initiatives can make the road more safe. However, there are concerns about privacy and data protection.
- The new General Data Protection Regulation (GDPR) that will enter into force in the EU 2018, defines **profiling** as "any form of automated processing of personal data consisting of the use of personal data to evaluate certain personal aspects relating to a natural person, in particular to analyze or predict aspects concerning that natural person's performance at work, economic situation, health, personal preferences, interests, reliability, behavior, location or movements." (par 71)

  Q1. Is the Drive Safe initiative considered profiling?

  Q2. Is it allowed?]

  Q3. If allowed, how can ANWB justify the premium, later?

Problem

e.g. (Nissenbaum 1994;1996; Diakopoulos 2016)

- Algorithmic accountability. Information systems are more and more pervasive; making decisions that crucially affect people's lives.
- Some person (human or legal) is responsible for the decisions being made by the system. Looking back, she is also accountable for the outcomes. Legally, she may even by liable, in case of damages.
- That means that the system should be designed in such a way that decisions can be justified later, so the person can be held accountable:
  - evidence of the data and decision rule that were actually used to generate a decision (outcome control),
  - or, evidence that the system is set-up in such a way that it follows proper processes and uses proper data (behavioural control).

(compare Eisenhardt 1985)

Some terminology

- Agents are **responsible** for their actions, because they have a choice (free will). Responsibility is forward looking.
- Agents are **accountable** for their actions to others, who depend on the outcome of the actions. **Accountability** is backward looking.
- Van De Poel (2011): necessary conditions to be held accountable
  a. Capacity. The agent must be able to act responsibly (including 'free will'). But when the agent is under pressure or coerced, or mentally or physically disabled, she is no longer accountable.
  b. Wrongdoing. Something must be at fault. The agent failed to avoid undesired condition X. Or, the agent transgressed duty D.
  c. Causality. The agent is instrumental in causing some undesired condition X. Either the agent is generally able to influence occurrence of X or transgression of duty D causes X to occur.
- The notions of **blame** and **liability** look similar, but are different.

Computing and Accountability

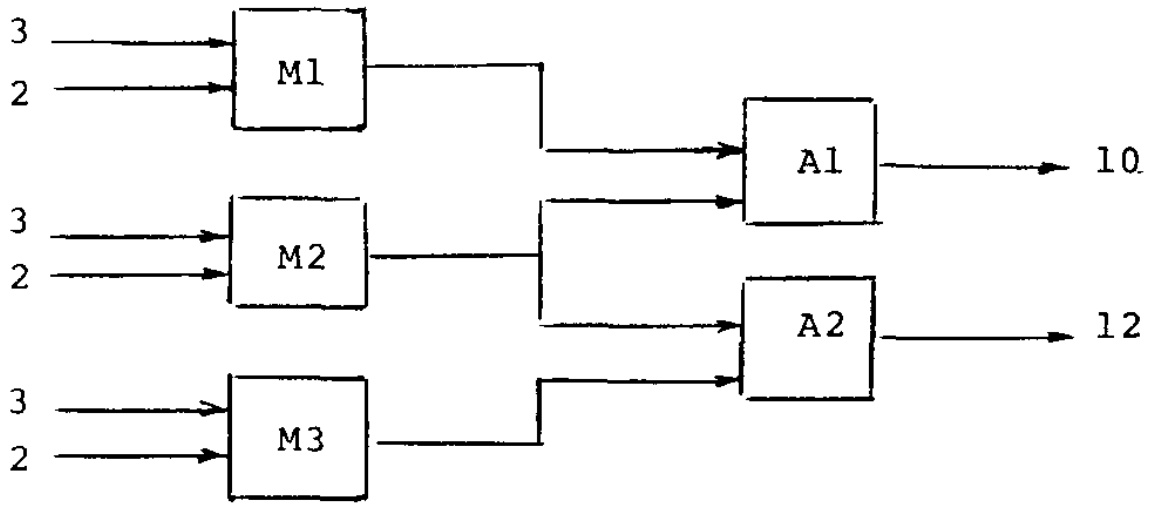
- Nissenbaum (1994): “accountability is systematically undermined in our computerized society” (p 73).
- Four barriers to accountability:
  1. problem of many hands: software is produced by groups, collectives or corporations, and, software itself is modular.
  2. bugs: errors are inevitable, for a complex system. This attitude makes it easy to avoid accountability; it can’t be helped.
  3. computer as scapegoat: this creates sympathy, but avoids the role of the responsible people behind the error.
  4. ownership without liability: debate about software ownership (e.g. patents) avoids responsibilities! (compare aircraft)

Accountability and MAS

- Accountability is about the attitude of social agents towards technology; it concerns socio-technical systems. These can be fruitfully studied and understood using MAS techniques.
- Accountability is especially relevant to agent systems
  - Accountability requires a reflective agent architecture
  - Logics for practical reasoning, are especially useful when studying accountability conditions (Broersen)
  - Fault detection (diagnosis) is AI (Reiter 1987; De Kleer 1987)
- Accountability is especially relevant to multi-agent systems
  - MAS are distributed systems: problem of many hands
  - MAS display emergent behaviour; faults are hard to trace
- Accountability relations involve social interaction.

Fault Detection (Diagnosis)

(Reiter 1987; De Kleer 1987)



$$\text{MULTIPLIER}(m) \wedge \neg \text{AB}(m) \supset \text{out}(m) = \text{in1}(m) * \text{in2}(m) .$$

What is the minimal set of observations, in order to determine which component is at fault?

Agent Architectures for Compliance

Burgemeestre et al 2009

- Norm as a ‘filter condition’

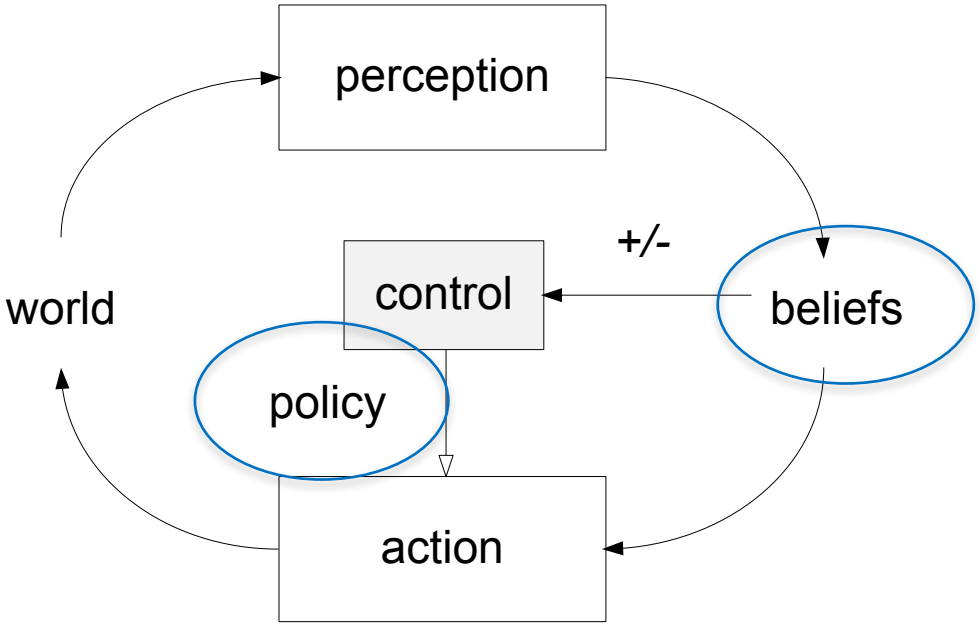


Fig. 1. Perception and Action

Agent Architectures for Compliance

Burgemeestre et al 2009

- Norm as a constraint on plans in the plan library
- Norm as a filter on adoption

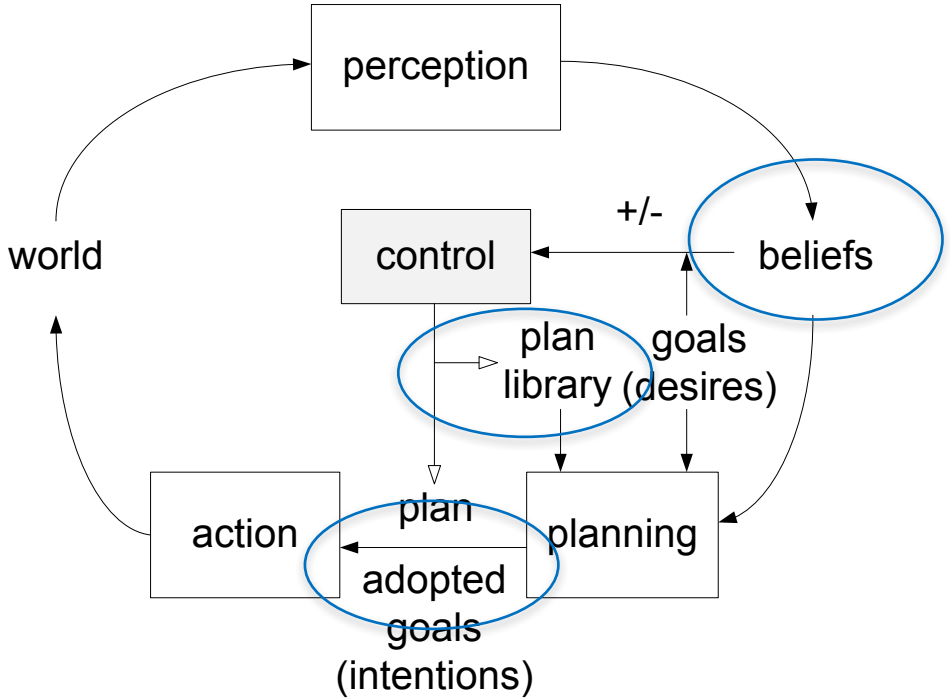


Fig. 2. Perception, planning and action

TILBURG UNIVERSITY

Agent Architectures for Compliance

Burgemeestre et al 2009

- Norm as a maintenance goal

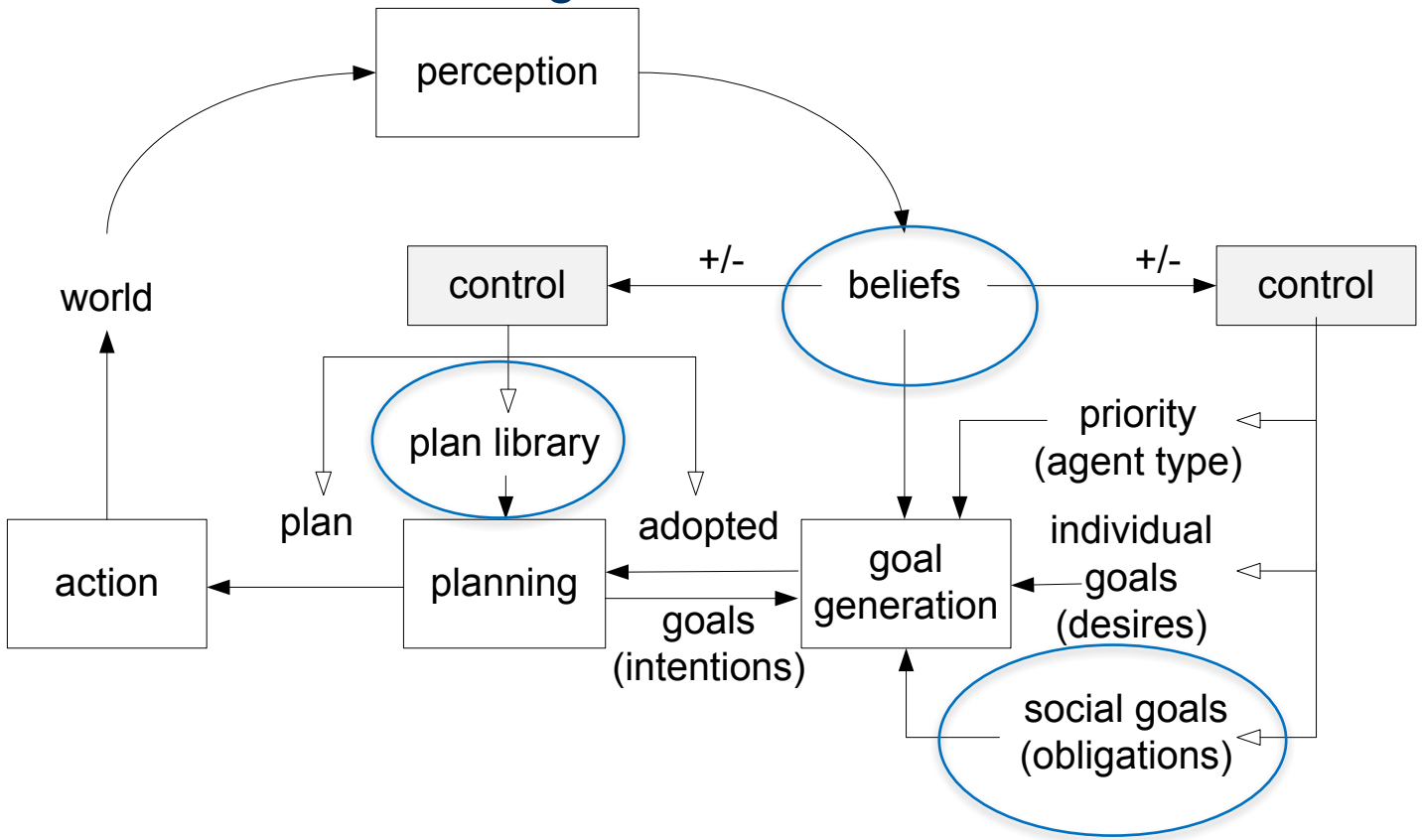


Fig. 3. Perception, goal generation, planning and action

MAS and Monitoring

- What kinds of properties can be monitored?
- Monitoring norms under the condition of imperfect observability (Alechina, Dastani, Logan 2014)

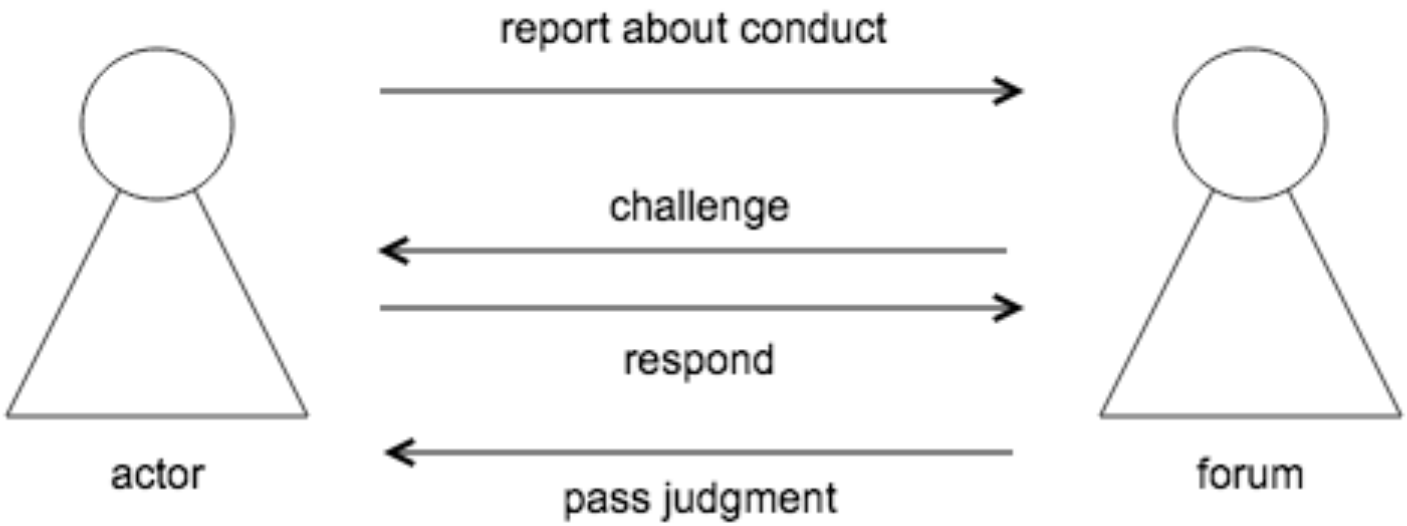
Accountability and Traceability

Chopra and Singh 2014

- A lot of emphasis on an audit trail, that allows to trace who did what.
- Accountability =\= traceability
- First, traceability of actions is not always necessary. One can also use the outcomes of a process, rather than the way it was carried out, to hold someone accountable. Compare the difference between outcome control and behavioral control (Eisenhardt 1985).
- Second, traceability is not enough. What is needed in addition is a social mechanism of holding the agent accountable: someone must evaluate the audit trail and confront the agent with possible faults.

Accountability as a dialogue

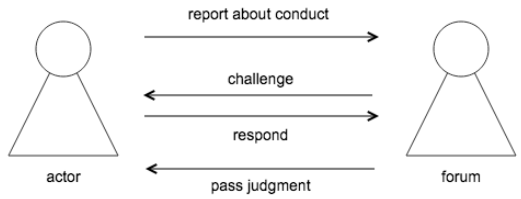
(Bovens 2005; 2007)



- Actor: feels obliged to report about conduct, to forum
- Forum: some ‘significant other’, e.g. clients, general public, God, ...

- Can be specified as a dialogue game (Burgemeestre et al 2011)

Accountability

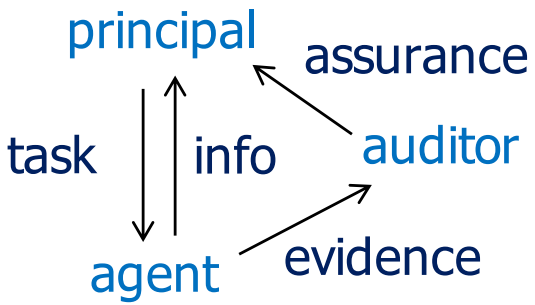


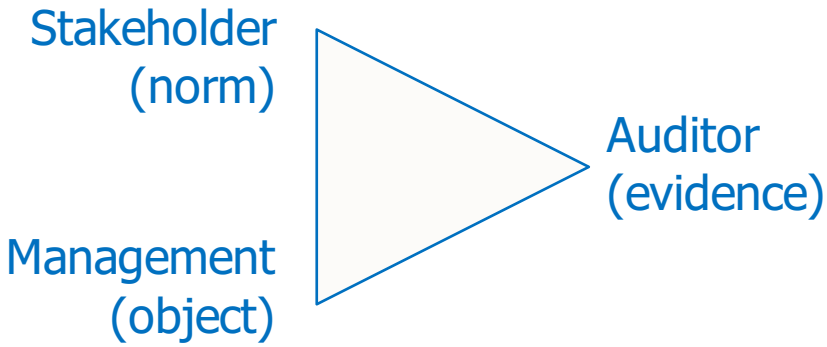
Bovens identifies 5 necessary conditions:

- (1) public accessibility of the account giving—and not purely internal, discrete informing;
- (2) explanation and justification of conduct—and not propaganda, or the provision of information or instructions to the general public;
- (3) the explanation should be directed at a specific forum—and not be given at random;
- (4) the actor must feel obliged to come forward—instead of being at liberty to provide any account whatsoever; and
- (5) there must be a possibility for debate and judgment, including an optional imposition of (informal) sanctions, by the forum—and not a monologue without engagement" (Bovens, 2005) (p 185)

Audit and Agency Theory

- Accountability: management must provide reliable evidence of financial results (compliance) to stakeholder (regulator)
- Paradox: evidence is generated by procedures and information systems, which are controlled by the party being regulated
- Internal controls: precautions built into the processes, information systems and governance structure to ensure reliability
- Audit: provide assurance over reliability (accuracy and completeness) of the evidence, and hence over reliability of internal controls





Agency theory: e.g. (Eisenhard 1989)

Designing for Accountability

- Requirements for an accountability relation (compare Grice)
  - Reliable reporting (accurate and complete). Effectuated by internal controls, built into processes and systems.
  - Existence of a some standard or norm, to compare.
  - From Bovens: Specific, independent and critical forum. Powerful enough to have a credible claim to sanctions. Expertise to evaluate and pass judgment.

Case

Ensuring Data Protection by Contract Monitoring

Stéphanie van Gulijk, Joris Hulstijn

Tilburg University

S.vanGulijk@uvt.nl, J.Hulstijn@uvt.nl

TILBURG UNIVERSITY

Problem

- Data drives modern society.
- Citizens' rights are being threatened; citizens feel uninformed or even powerless towards these large organizations when consenting to new data practices
- General Data Protection Regulation
  - compliance by design
  - duty to report breaches of security
  - right to be forgotten
  - privacy officer
  - profiling forbidden, unless …
- But ...

GDPR has a number of flaws (i)

- The notion of informed consent, on which much legal doctrines to protect personal data are based, is ineffective in practice (Schermer, Custers, & van der Hof, 2014; van Boom, Giesen, & Verheij, 2008).
  - contract terms on data protection are often hidden in complex sets of standard terms.
  - consumers tend to easily accept standard contract terms, regardless whether they really understand the consequences.
  - consumers often agree because there is no real alternative.

GDPR has a number of flaws (ii)

- Current public law legislation only covers the collection and processing of personal data, i.e. about individuals. Data concerning groups is not regulated (Taylor, Floridi, & van der Sloot, 2017).
- Control over how data is being used is not covered either.
- Practices like automated decision making (profiling) are obscure and very hard to regulate (accountability of algorithms)

GDPR has a number of flaws (iii)

- Supervision is largely based on corporate compliance, embedded in a public law framework.
  - Companies make privacy policies, but compliance to these policies is hardly ever tested.
  - In case of violations, regulators can only take action after the fact and in individual situations measures can be taken.
  - There is insufficient enforcement capacity
- No role for individual end users

Proposal

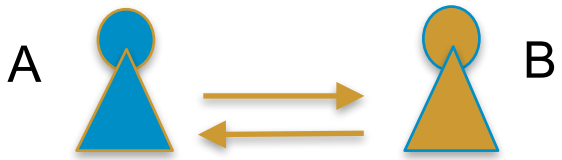
- We argue that private law instruments could empower consumers to take data protection into their own hands. The existing public law data protection framework remains valuable as a ‘last resort’
  - **Purpose binding**. The notion of ‘purpose’ remains central in the GDPR (minimize data collecting; minimize storage duration, unless needed for given purpose). What evidence? The contract.
  - **Counter offer** In many business models, consumers essentially ‘pay’ for services by providing personal data.
  - **Consumer empowerment**. Collectively consumers have power over an organization. The organization needs the data. When contracts are individually negotiable, the user remains in power. He or she could withdraw consent at any moment. When users organize themselves in collectives, they can leverage that power

Strengthening data protection rights in contracts

- Currently, contract law falls short, since the notions consent and considerations are not fit for purpose when payment with personal data instead of money is at stake.
- Clauses in standard contract terms (“small print”) that include the collection or use of personal data are not yet protected by law. For instance, clauses in standard contract terms that exclude or limit damages in relation to consumer, or clauses in standard contract terms that force consumer to bring their dispute to a court of arbitration instead of the civil court, are unfair to consumers (art. 6:236/237 BW).
- Unfair contract terms that lead to consumers being in a possible worse situation than professional contract parties are not allowed according to Dutch and European contract law [HvJ Océano, Pannon, Mostaza Claro etc.]

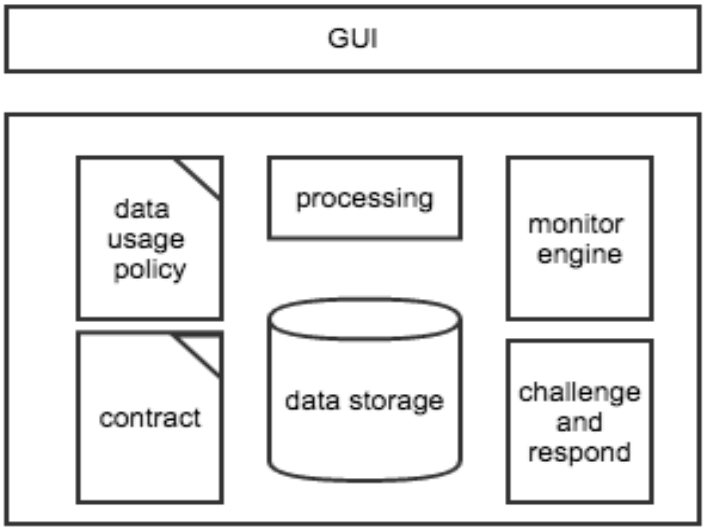
Strengthening enforcement

- An organization and a user jointly agree on a data policy. The policy clearly states which data are collected, for which purpose, and against which compensation.
- Some trusted infrastructure monitors the way data is actually being used and provides frequent reports, for example through a dashboard.
- Using these reports, stakeholders can determine whether organizations conform to the policies.
- If not, stakeholders can take action: raise awareness, withdraw consent, or negotiate a better compensation. Collectively stakeholders have a lot of power.



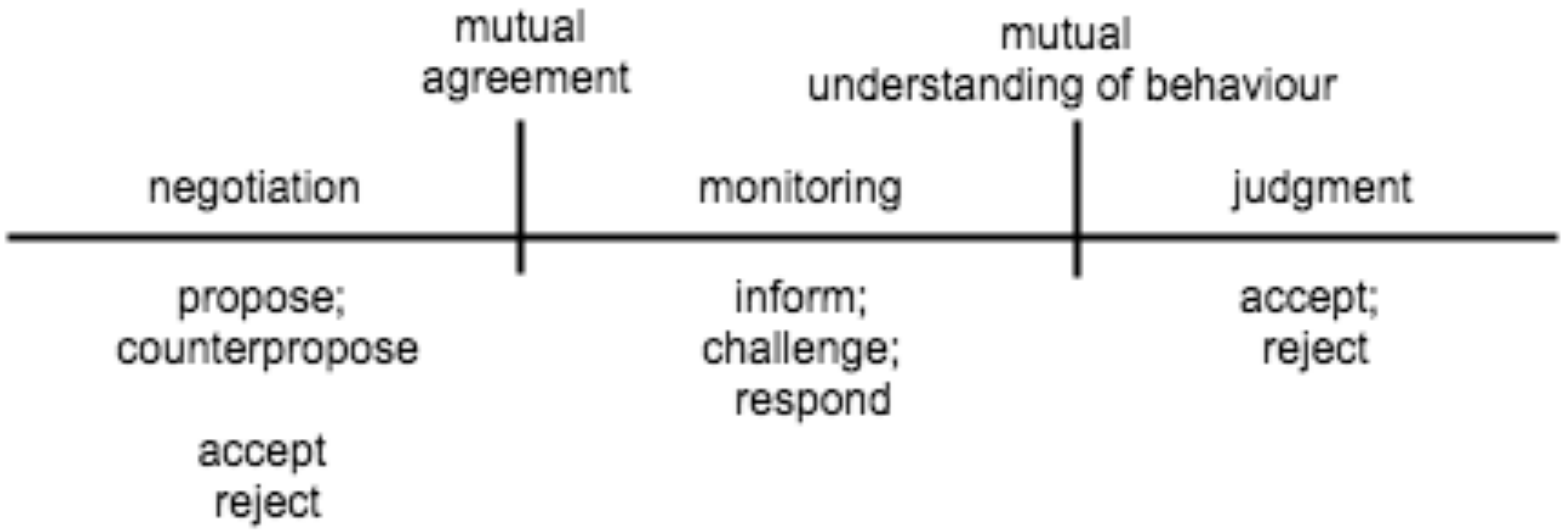
System requirements

- What is needed in terms of software support and environment?
  - A framework to help users negotiate sensible contracts that take data collection and usage into account.
  - A data infrastructure with built-in mechanisms for monitoring adherence to these terms and conditions
  - Governance structure with opportunities for challenging the organization, and passing judgment.
  - Social network tools, to mobilize a forum, of similar users.



TILBURG UNIVERSITY

Towards a framework

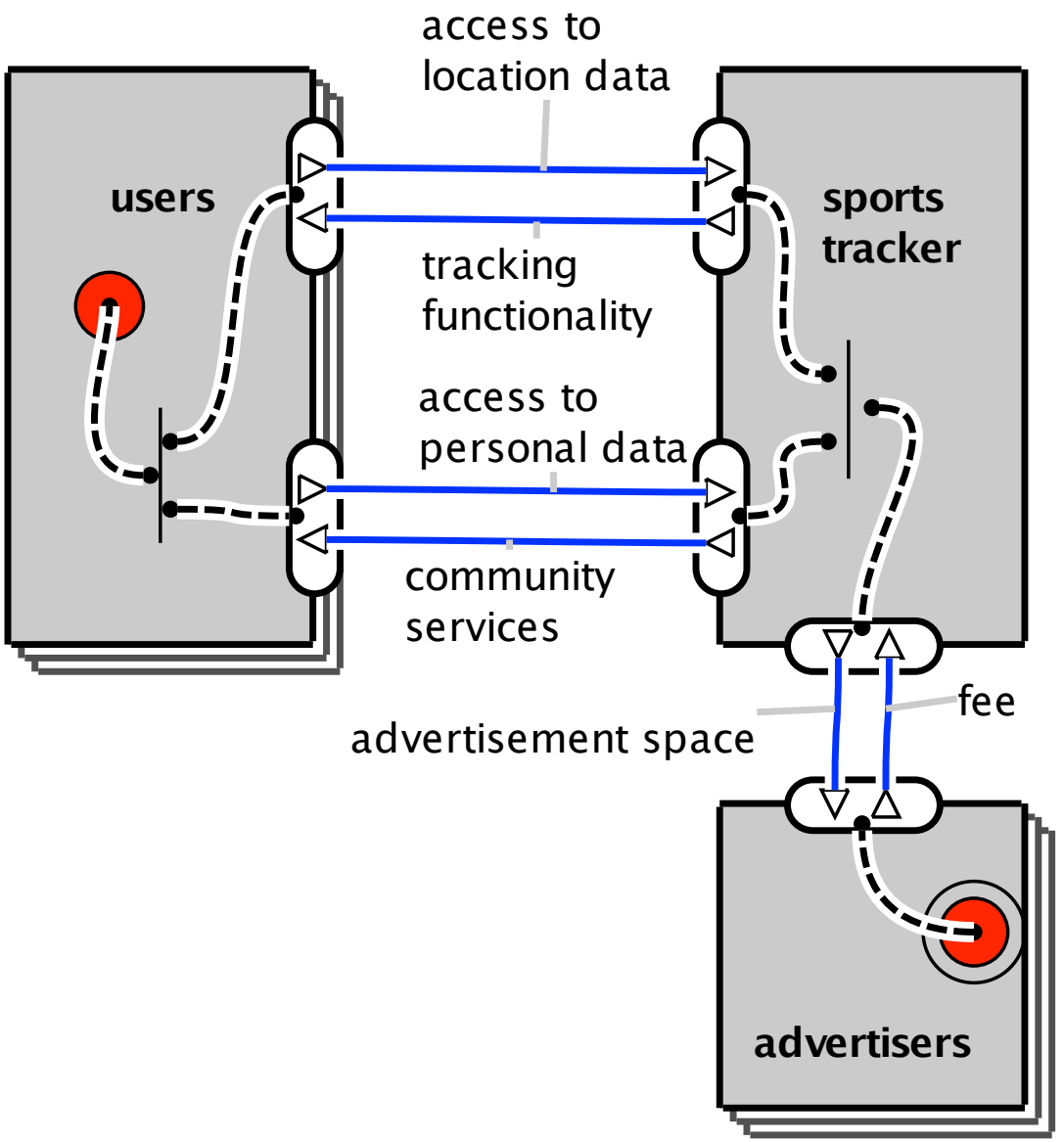


Is this a case of accountability?

- According to Bovens it is, when
  1. public accessibility: access limited to meta-data, not content, and only within a closed community [feasible]
  2. explanation and justification of conduct: relates to objectives, values and policies that were agreed. [functionality]
  3. directed at specific forum: must be mobilized; can be done, based on data-driven business model [functionality]
  4. actor must feel obliged: powerful enough? [legal action]
  5. possibility for debate and possibly sanctions [ability to block]

Example

Compliant Norwegian Consumer Council against Runkeeper (2016)

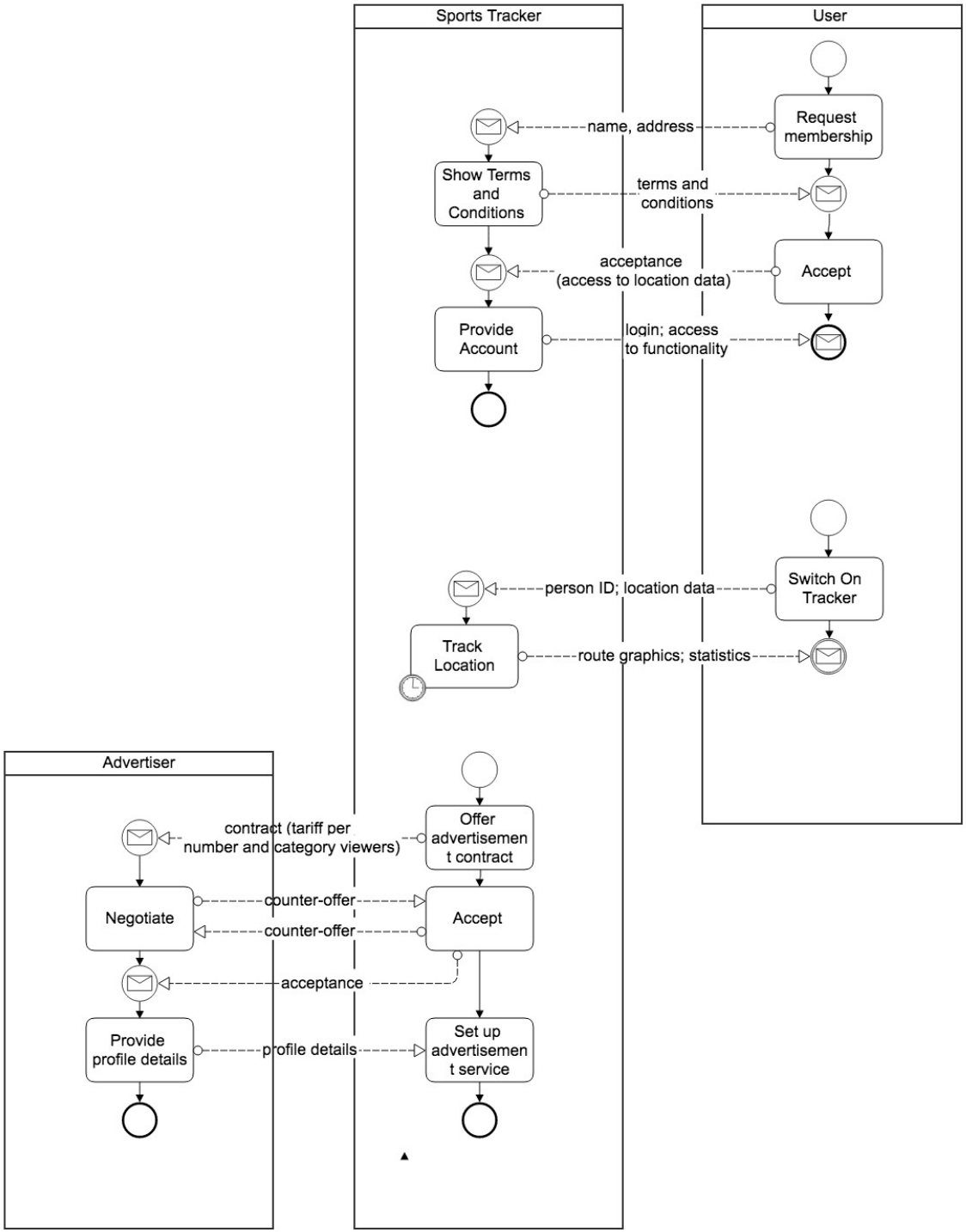


Example

- User may provide:
  - Access to location data [Y/N]
  - Personal data [Y/N]
  - Right to share aggregated data in community (anonymous) [Y/N]
  - Right to share sports related data in community (traceable) [Y/N]
  - Contributions to community [text]
- Tracker may provide
  - Tracker functionality, only if user provides access to location data
  - Additional services, only if user provides access to location and personal data, and allows tracker to share aggregated location data with other users (anonymous)
  - Community services, only if user provides access to location and personal data, and allows tracker to share aggregated location data with other users, under community ID

Example

- Process model
- Need collaboration of mobile platforms (iOS, Android) to install a kind of App-fire-wall:
- monitoring and filtering of different kinds of data according to policy rules



TILBURG UNIVERSITY

Conclusions

- Accountability is crucial for our computerized society, but is harmed.
- Accountability is a real MAS topic.
  - distributed; fault detection and diagnosis; practical reasoning
- A solution lies in treating accountability as a dialogue with a forum.
  - find a powerful forum (e.g. community of users / data subjects)
  - provide tools to negotiate better contracts, and monitor
  - allow a debate, and possibly sanction (collective action)
- Application domain: data protection (give subjects control over data)
  - could be based on private law
  - empower individual end-users
  - simple tools; understandable; rooted in existing platforms

References

- N. Alechina and M. Dastani and B Logan (2014), Norm Approximation for Imperfect Monitors, AAMAS 2014, 117—124.
- van Boom, W. H., I. Giesen and A. J. Verheij (2008). Gedrag en Privaatrecht, over gedragspresumpties en gedragseffecten bij privaatrechtelijk leerstukken. Den Haag, Boom Juridische Uitgeverij (BJu).
- Broersen, J., Dastani, M., Hulstijn, J., and Van der Torre, L. (2002) 'Goal Generation in the BOID Architecture', Cognitive Science Quarterly, 2(3-4): 431-450
- Burgemeestre, B., J. Hulstijn and Y.-H. Tan (2009). Agent Architectures for Compliance. ESAW 2009. LNCS 5881: 68-83.
- Chopra, A. K. and M. P. Singh (2014). The thing itself speaks: Accountability as a foundation for requirements in sociotechnical systems. (RELAW 2014) 22-22
- Eisenhardt, K. M. (1985). Control: Organizational and Economic Approaches. Management Science, 31(2), 134-149.

References

- de Kleer, J. and B. C. Williams (1987). "Diagnosing Multiple Faults." Artificial Intelligence 32(1): 97-130.
- Nissenbaum, H. (1994). Computing and Accountability Communications of the ACM, 37(73-80).
- Nissenbaum, H. (1996). Accountability in a Computerized Society. Science and Engineering Ethics, 2(1), 25-42.
- Reiter, R. (1987). "A theory of diagnosis from first principles." Artificial Intelligence 32(1): 57 – 95.
- Schermer, B. W., B. H. M. Custers and S. van der Hof (2014). "The crisis of consent: how stronger legal protection may lead to weaker consent in data protection." Ethics and Information Technology 16(2): 171-182.
- aylor, L., L. Floridi and B. van der Sloot, Eds. (2017). Group Privacy: new challenges of data technologies. Dordrecht Springer.