Consent as a Foundation for Responsible Autonomy

Munindar P. Singh
North Carolina State University
singh@ncsu.edu

AAAI 2022 Blue Sky
Consent for AI: A Foundation for Governance

Societal and individual objectives

Sociotechnical systems
Consent for AI: A Foundation for Governance

Societal and individual objectives

Sociotechnical systems

Autonomy
Consent for AI: A Foundation for Governance

Societal and individual objectives

Sociotechnical systems

Autonomy

Responsibility
Consent for AI: A Foundation for Governance

- Societal and individual objectives
- Sociotechnical systems
  - Autonomy
  - Responsibility
    - Governance
    - Consent reasoning
Domains of Consent

Personal

Sex
Risk

Government
Explicit
Tacit

Health care
Treatment
Information

Privacy
Computing
Contracts

Singh (NC State)
Consent as a Foundation for Responsible Autonomy
AAAI 2022 Blue Sky
Domains of Consent

- Personal
- Risk
- Privacy
- Computing
- Contracts
- Health care
- Information
- Treatment

Government

Explicit
Tacit
Domains of Consent

- Personal
- Government
- Explicit
- Tacit
- Health care
  - Treatment
  - Information
- Privacy
- Computing
- Contracts
- Risk
- Sex
Domains of Consent

Personal

Government

Explicit

Tacit

Risk

Sex

Health care

Treatment

Information

Privacy

Contracts

Computing

Consent as a Foundation for Responsible Autonomy

AAAI 2022 Blue Sky
What Makes Consent Worthwhile?

Moral legitimacy

Help or intrude

Moral magic

Legal legitimacy

Borrow or steal

Transfer accountability

Override defaults
What Makes Consent Worthwhile?

Moral legitimacy

Moral magic

Help or intrude

Transfer accountability

Override defaults

Legal legitimacy

Borrow or steal
What Makes Consent Worthwhile?

Moral legitimacy

Help or intrude

Moral magic

Borrow or steal

Legal legitimacy

Transfer accountability

Override defaults
Parties to Consent

Traditionally

Consenter

Consentee
Parties to Consent

In computing, apparently

Consenter → Software
Parties to Consent

In computing, actually

Consenter \[\text{usr}/\text{cgs}\] 

Software 

Consentee \[\text{usr}/\text{cgs}\]
Parties to Consent

With AI agents, user to agent and agent to agent

Consenter  →  Agent  →  Agent  →  Consentee
Responsibility in AI

Responsible AI

Value-based design

Design methodologies

Static picture

How to build, deploy AI apps

How an agent behaves

Decision making by AI

Responsible autonomy

Follow own values

Respect others’ values

Consent as a Foundation for Responsible Autonomy

Singh (NC State)

Consent is crucial

Understand responsibility

Develop responsible agents

AAAI 2022 Blue Sky
Responsibility in AI

Responsible AI

Value-based design

Design methodologies

Static picture

How to build, deploy AI apps

Consent is crucial

Understand responsibility

Consent as a Foundation for Responsible Autonomy

Responsible autonomy

How an agent behaves

Decision making by AI

Follow own values

Respect others' values

Develop responsible agents

Singh (NC State)
Responsibility in AI

Responsible AI
- How to build, deploy AI apps
- Value-based design
- Design methodologies

Responsible autonomy
- How an agent behaves
- Decision making by AI
- Follow own values
- Consent as a Foundation for Responsible Autonomy
- Respect others' values

Consent is crucial
- Consent as a Foundation for Responsible Autonomy
- Understand responsibility
- Develop responsible agents

Singh (NC State) Consent as a Foundation for Responsible Autonomy AAAI 2022 Blue Sky
Why Consent in AI?

- Exercise one’s autonomy
- Discretion to violate norms
- Grant powers
- Receive power
- Respect others’ autonomy
- Strengthen norms
- Societal benefits
- Align norms and values
- Engender trust in AI
- Align norms and values
Why Consent in AI?

- Exercise one's autonomy
- Grant powers
- Discretion to violate norms

Respect others’ autonomy
- Receive power
- Strengthen norms

Societal benefits
- Align norms and values
- Engender trust in AI
- Singha (NC State)

Consent as a Foundation for Responsible Autonomy

AAAI 2022 Blue Sky
Why Consent in AI?

Grant powers
Receive power
Exercise one’s autonomy
Respect others’ autonomy
Discretion to violate norms
Strengthen norms

Societal benefits
Align norms and values
Engender trust in AI

Singh (NC State) Consent as a Foundation for Responsible Autonomy AAAI 2022 Blue Sky
Adequacy of Representation

Meta-physical

Others choose their values & interests

Gain legitimacy

Heterogeneity

Epistemological

Don't know others' values & interests

Gain confidence

Heuristic

Autonomy

Efficiency

Consent as a Foundation for Responsible Autonomy

[McCarthy & Hayes, 1969]
Adequacy of Representation

Meta-physical

Others choose their values & interests

Gain legitimacy

Autonomy

Epistemological

Heterogeneity

Don’t know others’ values & interests

Gain confidence

Heuristic

Gain contestability

Efficiency

Ensure legitimacy

Volenti maxim

[McCarthy & Hayes, 1969]
Adequacy of Representation

Meta-physical
- Others choose their values & interests

Epistemological
- Don’t know others’ values & interests
- Gain confidence

Autonomy

Gain legitimacy

Heterogeneity

Gain contestability

Efficiency

Ensure legitimacy

Volenti maxim

Heuristic

Singh (NC State) Consent as a Foundation for Responsible Autonomy

[McCarthy & Hayes, 1969]
Consent as a Foundation for Responsible Autonomy
Where Do We See Consent?

- Vacuity through tacit consent
- Consent of the governed
- Legitimacy from value alignment
- Not intent
- Not efficiency or fairness
- Not enforcement
- Renegotiate if exceptions
- Emphasize trust
- Business contracts
- Political science
- Social relations

Singh (NC State) Consent as a Foundation for Responsible Autonomy
Where Do We See Consent?

- Vacuity through tacit consent
- Consent of the governed
- Legitimacy from value alignment
- Not intent
- Not efficiency or fairness
- Not enforcement

Political science

Business contracts

Social relations

Renegotiate if exceptions
Emphasize trust

Singh (NC State) Consent as a Foundation for Responsible Autonomy AAAI 2022 Blue Sky
Popular in computing

Obtain approval
Inform user
Easy to implement

Notice and Choice

Problems
Power difference
Lack of understanding
Hides context

Singh (NC State) Consent as a Foundation for Responsible Autonomy AAAI 2022 Blue Sky
Notice and Choice

Popular in computing

Obtain approval
Inform user
Easy to implement

Problems

Power difference
Lack of understanding
Hides context

Singh (NC State)
Consent as a Foundation for Responsible Autonomy
AAAI 2022 Blue Sky
Misuses of Consent

- Mislead consenter
- Disregard consenter’s interests
- Absolve consentee
Misuses of Consent

- Mislead consenter
- Disregard consenter’s interests
- Absolve consentee
Misuses of Consent

Mislead consenter
Disregard consenter’s interests

Absolve consentee
Misuses of Consent

- Mislead consenter
- Disregard consenter’s interests
- Absolve consentee

Just a means to be irresponsible
What is Consent?

- Mental action
- How can norms arise?
- Communicative act
- How can there be mistakes?
- Pragmatic, social action
- Can be evaluated for legitimacy
- Mental and communicative
What is Consent?

Mental action

Internal choice

Pragmatic, social action

Can be evaluated for legitimacy

Mental and communicative

Communicative act

Confers powers

How can norms arise?

How can there be mistakes?
What is Consent?

- Mental action
- Internal choice
- Communicative act
- Confers powers
- How can there be mistakes?

Pragmatic, social action

Can be evaluated for legitimacy

Mental and communicative
Habermas' Validity Claims

Objective

- Is it true?

Subjective

- Is it sincere?
- Appropriate beliefs, intentions

Practical

- Suited to its context
- Is it justified?
Habermas’ Validity Claims

Objective
- Is it true?

Subjective
- Is it sincere?
- Appropriate beliefs, intentions

Practical
- Suited to its context
- Is it justified?
Habermas' Validity Claims

Objective

Is it true?

Empirical

Subjective

Is it sincere?

Appropriate beliefs, intentions

Practical

Is it justified?

Suited to its context

Singh (NC State)

Consent as a Foundation for Responsible Autonomy

AAAI 2022 Blue Sky
## Criteria for Valid Consent

<table>
<thead>
<tr>
<th>Criterion</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>Visibility</td>
<td>Consent is observable, i.e., a communication</td>
</tr>
<tr>
<td>Free Will</td>
<td>Consenter acts without being coerced</td>
</tr>
<tr>
<td>Truth</td>
<td>Consenter’s beliefs are true and complete</td>
</tr>
<tr>
<td>Capacity</td>
<td>Consenter is mentally fit</td>
</tr>
<tr>
<td>Cognition</td>
<td>Consenter believes and intends to</td>
</tr>
<tr>
<td>Attention</td>
<td>Consenter exercises mental faculties</td>
</tr>
<tr>
<td>Statutes</td>
<td>Consenter meets statutory criteria, e.g., age</td>
</tr>
<tr>
<td>Power</td>
<td>Consenter is not subjugated by consentee</td>
</tr>
<tr>
<td>Honesty</td>
<td>Consentee does not mislead consenter</td>
</tr>
</tbody>
</table>

Singh (NC State)  
Consent as a Foundation for Responsible Autonomy  
AAAI 2022 Blue Sky
Habermas' Validity Claims for Consent

Objective

Free will
Visibility

Truth

Subjective

Capacity
Cognition
Attention

Practical

Honesty
Statutes
Power

Singh (NC State) Consent as a Foundation for Responsible Autonomy AAAI 2022 Blue Sky
Habermas' Validity Claims for Consent

Objective
- Free will
- Visibility

Subjective
- Capacity
- Cognition
- Attention

Practical
- Honesty
- Statutes

Power

Truth

Singh (NC State)  Consent as a Foundation for Responsible Autonomy  AAAI 2022 Blue Sky
### Criteria for Valid Consent

<table>
<thead>
<tr>
<th>Objective</th>
<th>Subjective</th>
<th>Practical</th>
</tr>
</thead>
<tbody>
<tr>
<td>Visibility</td>
<td>Free Will</td>
<td>Statutes</td>
</tr>
<tr>
<td>Truth</td>
<td>Capacity</td>
<td>Power</td>
</tr>
<tr>
<td>Cognition</td>
<td>Attention</td>
<td>Honesty</td>
</tr>
</tbody>
</table>

**Meaning**

- Consent is observable, i.e., a communication
- Consenter acts without being coerced
- Consenter’s beliefs are true and complete
- Consenter is mentally fit
- Consenter believes and intends to
- Consenter exercises mental faculties
- Consenter meets statutory criteria, e.g., age
- Consenter is not subjugated by consentee
- Consentee does not mislead consenter

Singh (NC State)  Consent as a Foundation for Responsible Autonomy  AAAI 2022 Blue Sky
Challenges for Sociotechnical Systems

- Communication semantics
- Norms
- Argumentation

Bridge AI ethics and law
- Duty
- Discretion

Design methods
- Values
- Consentability

Formal Verification
- Quandaries are not enough
- Procedural justice

Continual learning
- Consent in context
- Dialogue

Singh (NC State) Consent as a Foundation for Responsible Autonomy AAAI 2022 Blue Sky
Challenges for Sociotechnical Systems

- Norms
- Argumentation
- Communication semantics
- Dialogue
- Consent in context
- Dialogue
- Consensual learning
- Procedural justice
- Formal verification
- Quandaries are not enough
- Values
- Design methods
- Consentability
- Duty
- Discretion
- Bridge AI ethics and law
- Duty
- Discretion
- Bridge AI ethics and law

Consent as a Foundation for Responsible Autonomy

Singh (NC State)

AAAI 2022 Blue Sky
Challenges for Sociotechnical Systems

- Norms
- Communication semantics
- Argumentation
- Bridge AI ethics and law
- Duty
- Discretion
- Design methods
- Values
- Consentability
- Continual learning
- Formal Verification
- Procedural justice
- Quandaries are not enough
- Consent in context
- Dialogue
- Consent as a Foundation for Responsible Autonomy

Singh (NC State)
Challenges for Sociotechnical Systems

- Norms
- Communication semantics
- Duty
- Discretion
- Bridge AI ethics and law
- Design methods
- Values
- Consentability
- Dialogue
- Continual learning
- Consent in context
- Argumentation

Formal Verification

- Procedural justice
- Quandaries are not enough

Singh (NC State) Consent as a Foundation for Responsible Autonomy AAAI 2022 Blue Sky
Challenges for Sociotechnical Systems

- Argumentation
- Norms
- Communication semantics
- Bridge AI ethics and law
- Duty
- Discretion
- Design methods
- Values
- Consentability
- Dialogue
- Consensual in context
- Continual learning
- Procedural justice
- Formal Verification
- Quandaries are not enough
- Procedural justice
- Continual learning
- Consent as a Foundation for Responsible Autonomy

Singh (NC State)
Thanks

- US NSF grant IIS-2116751
- DoD (Science of Security Lablet) for support
- Nirav Ajmeri, Pradeep Murukannaiah, Amika Singh