CA-CI: A Normative Framework for Evaluating Privacy and Dignity in AI Governance

Kat Roemmich, Kirsten Martin, Florian Schaub

PEPR '26:
2026 USENIX Conference on
Privacy Engineering Practice and Respect

AI privacy risks

- Multi-purpose AI systems pose new privacy risks
- Interaction data in one context serves as training data for models used across other contexts/purposes
- …

Risk-based AI regulation (EU AI Act)

- Prohibit AI systems posing **unacceptable risk to fundamental rights**
- High-risk AI systems require **fundamental rights impact assessment**
- Risk classification factors: deployment context, intended purpose, technical characteristics, nature/severity of potential harm

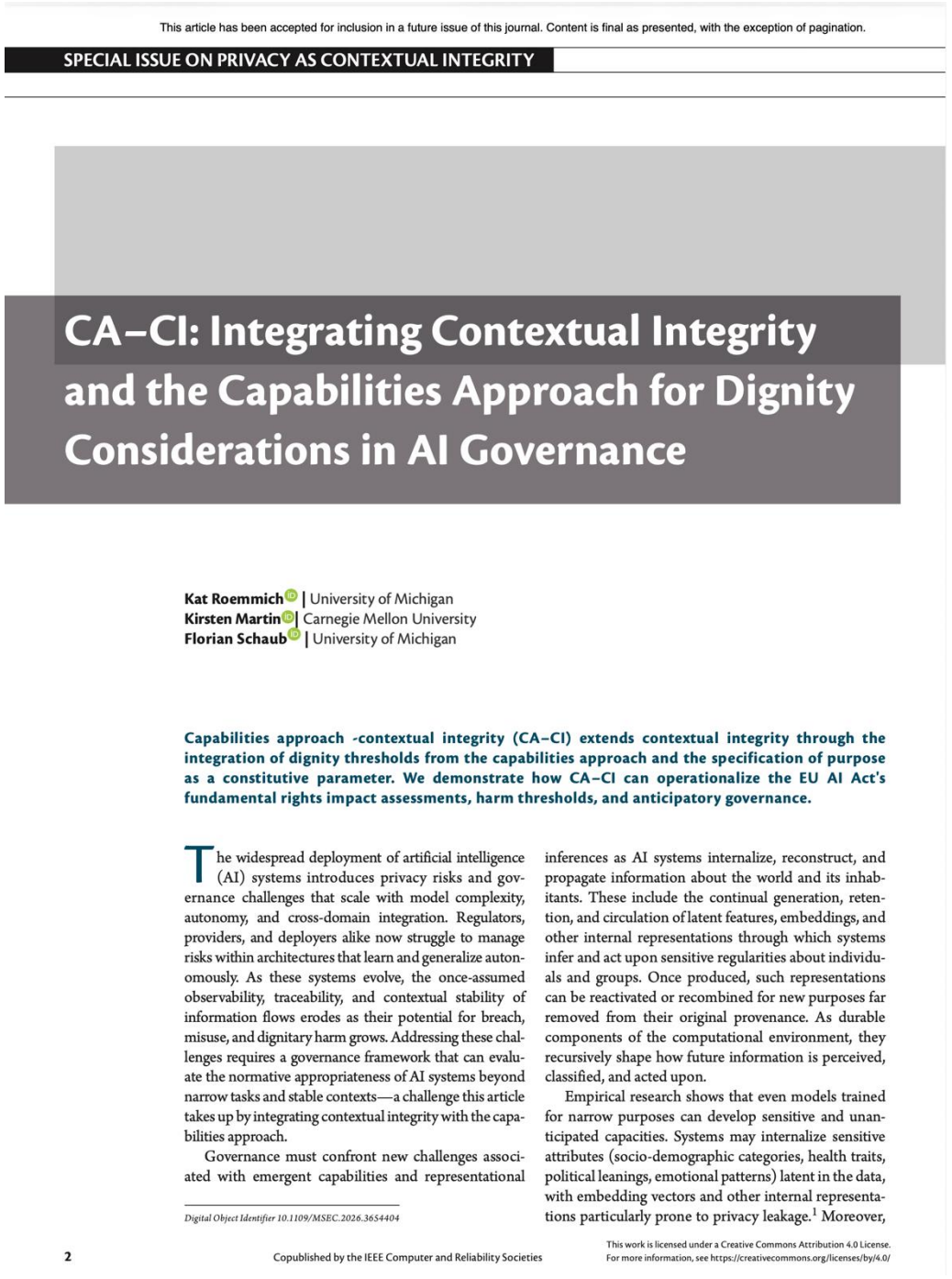
Image source: https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai

- How to evaluate contextual AI risk?
- How to assess impact on fundamental rights?

CA-CI Framework

Dignity considerations in
AI governance and risk assessment

Contextual Integrity (CI)
+ Capabilities Approach (CA)



This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

SPECIAL ISSUE ON PRIVACY AS CONTEXTUAL INTEGRITY

CA–CI: Integrating Contextual Integrity and the Capabilities Approach for Dignity Considerations in AI Governance

Kat Roemmich | University of Michigan
Kirsten Martin | Carnegie Mellon University
Florian Schaub | University of Michigan

Capabilities approach -contextual integrity (CA–CI) extends contextual integrity through the integration of dignity thresholds from the capabilities approach and the specification of purpose as a constitutive parameter. We demonstrate how CA–CI can operationalize the EU AI Act's fundamental rights impact assessments, harm thresholds, and anticipatory governance.

The widespread deployment of artificial intelligence (AI) systems introduces privacy risks and governance challenges that scale with model complexity, autonomy, and cross-domain integration. Regulators, providers, and deployers alike now struggle to manage risks within architectures that learn and generalize autonomously. As these systems evolve, the once-assumed observability, traceability, and contextual stability of information flows erodes as their potential for breach, misuse, and dignitary harm grows. Addressing these challenges requires a governance framework that can evaluate the normative appropriateness of AI systems beyond narrow tasks and stable contexts—a challenge this article takes up by integrating contextual integrity with the capabilities approach.

Governance must confront new challenges associated with emergent capabilities and representational inferences as AI systems internalize, reconstruct, and propagate information about the world and its inhabitants. These include the continual generation, retention, and circulation of latent features, embeddings, and other internal representations through which systems infer and act upon sensitive regularities about individuals and groups. Once produced, such representations can be reactivated or recombined for new purposes far removed from their original provenance. As durable components of the computational environment, they recursively shape how future information is perceived, classified, and acted upon.

Empirical research shows that even models trained for narrow purposes can develop sensitive and unanticipated capacities. Systems may internalize sensitive attributes (socio-demographic categories, health traits, political leanings, emotional patterns) latent in the data, with embedding vectors and other internal representations particularly prone to privacy leakage.[1] Moreover,

Digital Object Identifier 10.1109/MSEC.2026.3654404

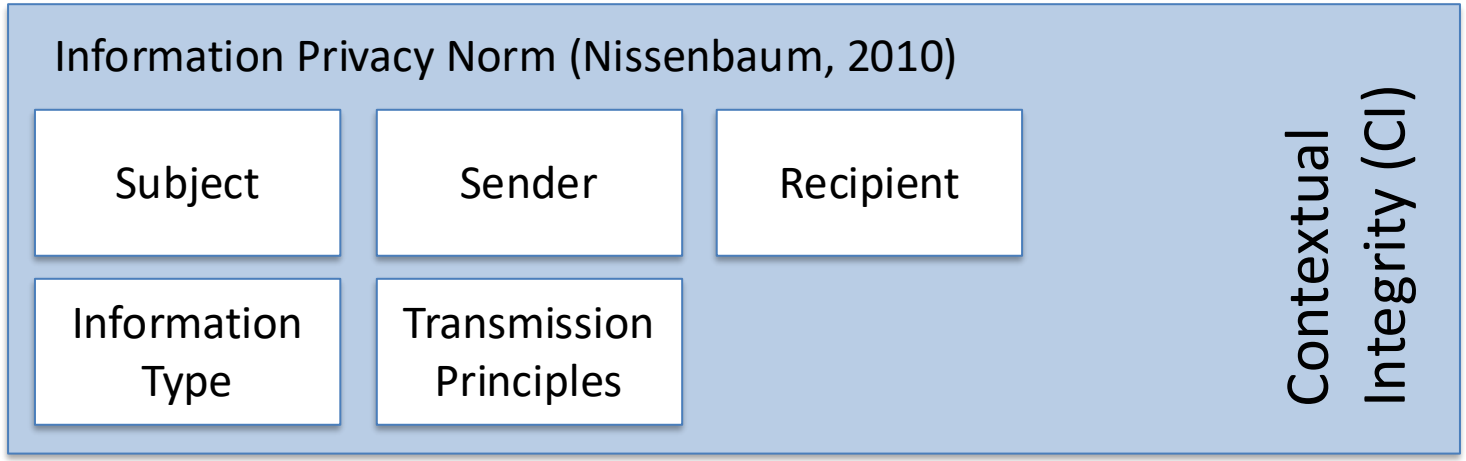
Copublished by the IEEE Computer and Reliability Societies

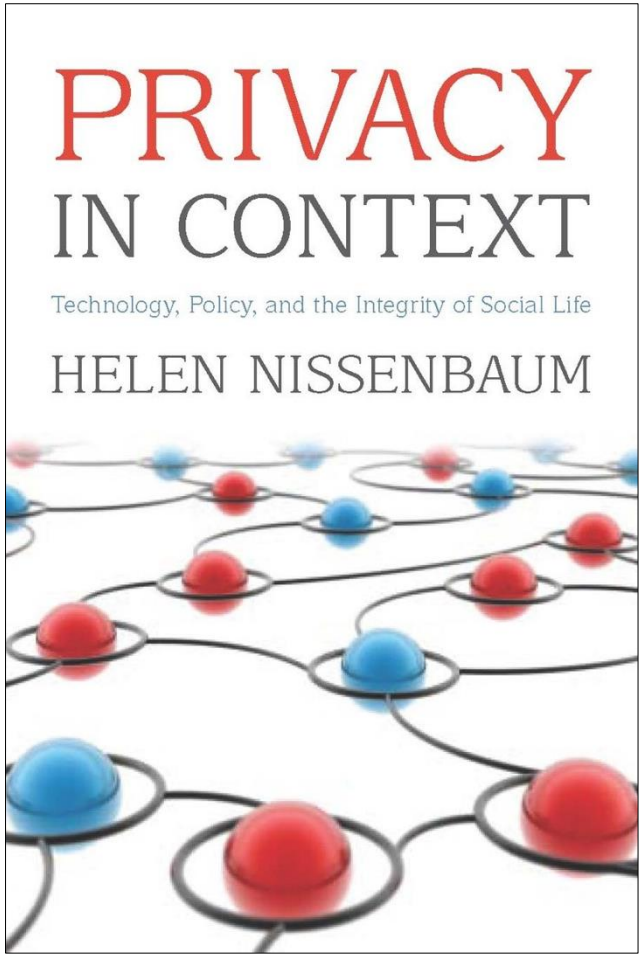
IEEE Security & Privacy, 2026

DOI: 10.1109/msec.2026.3654404

Contextual integrity (CI)

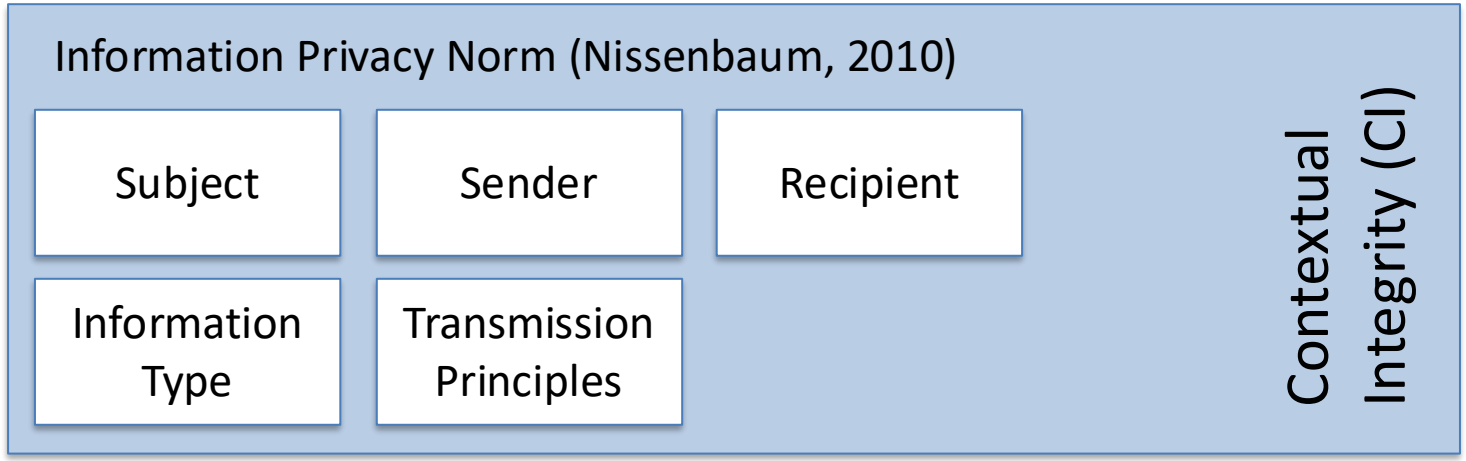
- Evaluate appropriateness of information flows relative to social context norms





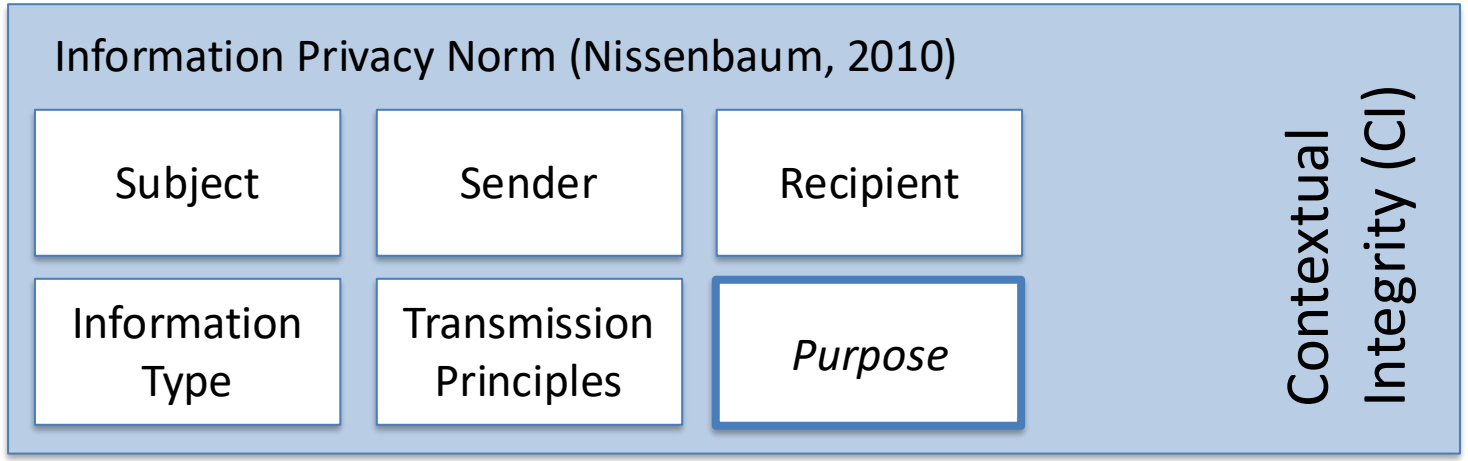
Contextual integrity (CI)

- **Purpose** only implicitly encoded in transmission principle
- Makes it difficult to track shifting use in same context
- Purpose specification is core privacy principle



CA-CI makes purpose explicit in CI

- Purpose added as explicit parameter
- Helps differentiate practices
- Identify cases when info flow/use diverges from original intent



CA-CI: Integrating Contextual Integrity and the Capabilities Approach for Dignity Considerations in AI Governance
Kat Roemmich, Kirsten Martin, Florian Schaub
IEEE Security & Privacy, 2026

Contextual integrity (CI)

- CI encodes social context norms for privacy

- What are privacy norms for AI systems?
- How do we define high / unacceptable risk thresholds?

Human dignity

- “Inherent worth and value of a human being”
- Shared moral agreement across societies and social domains
- Dignity protected by fundamental human rights
- How to operationalize dignity?

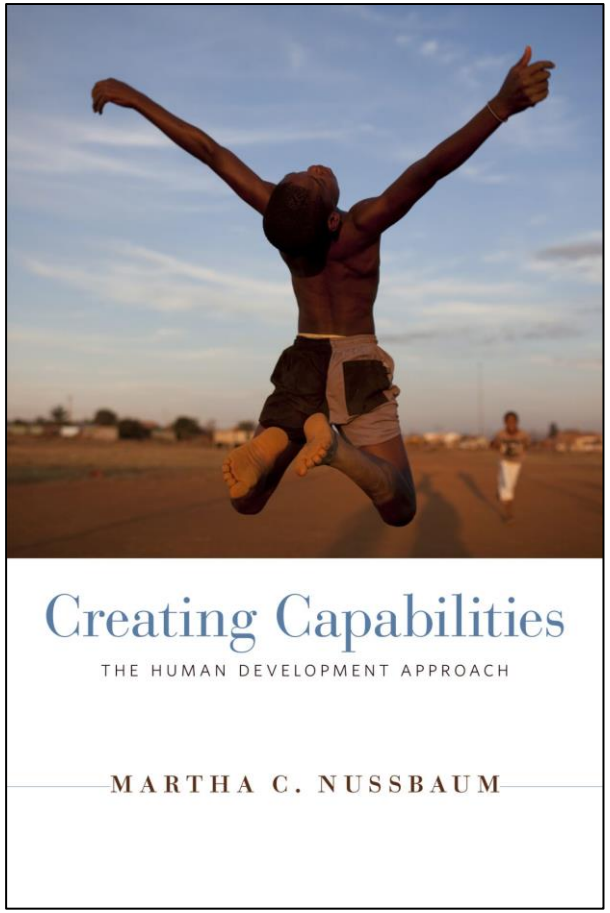
Capabilities Approach (CA)

- Ten core capabilities define *minimum* for a dignified life

Human Dignity, defined by core capabilities (Nussbaum, 2000)

Life	Bodily Health	Bodily Integrity	Emotions	Senses, Imagination, Thought
Practical Reason	Affiliation	Play	Other Species	Control over one's Environment

Capabilities Approach (CA)



Ten Central Capabilities

Life
Living to the end of a human life of normal length, not dying prematurely

Bodily integrity
Moving freely, secure against violent assault, choice in matters of reproduction and sexuality

Emotions
Attachments to things and people outside ourselves, free from emotional development blighted by fear and anxiety

Affiliation
Living with and toward others on the social bases of self-respect, treated as a dignified being of equal worth

Play
Being able to laugh, to play, to enjoy recreational activities

Bodily health
Good health, adequate nourishment, adequate shelter

Senses, imagination, and thought
Using one's senses, imagination, and thought informed by education and protected by freedom of expression

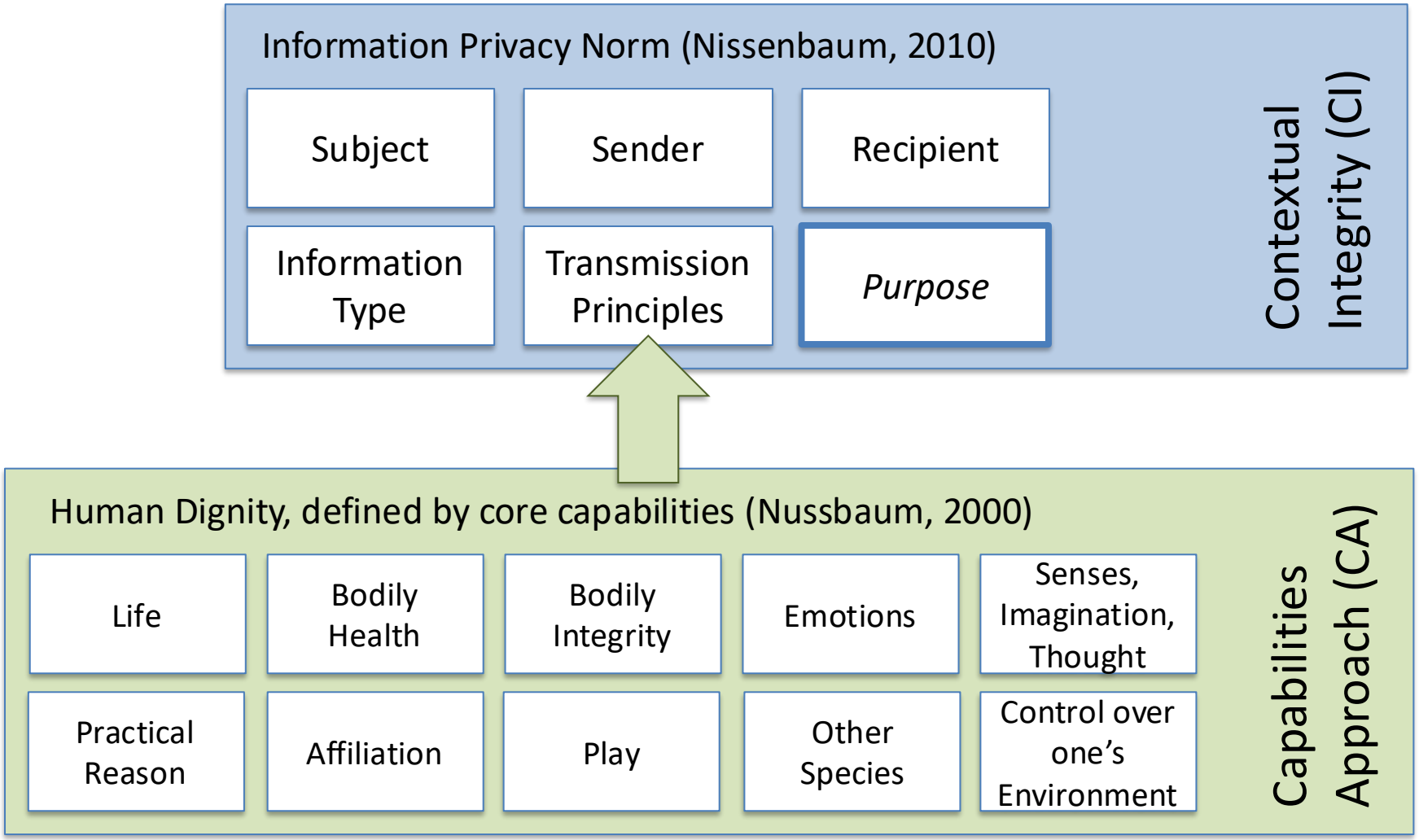
Practical reason
Forming a conception of the good and engaging in critical reflection about the planning of one's life

Other species
Living with concern for and in relation to animals, plants, and the world of nature

Control over one's environment
Participating effectively in political choices; holding property and seeking employment on an equal basis with others

Nussbaum, Creating Capabilities, 2011

CA-CI: Capabilities Approach–Contextual Integrity

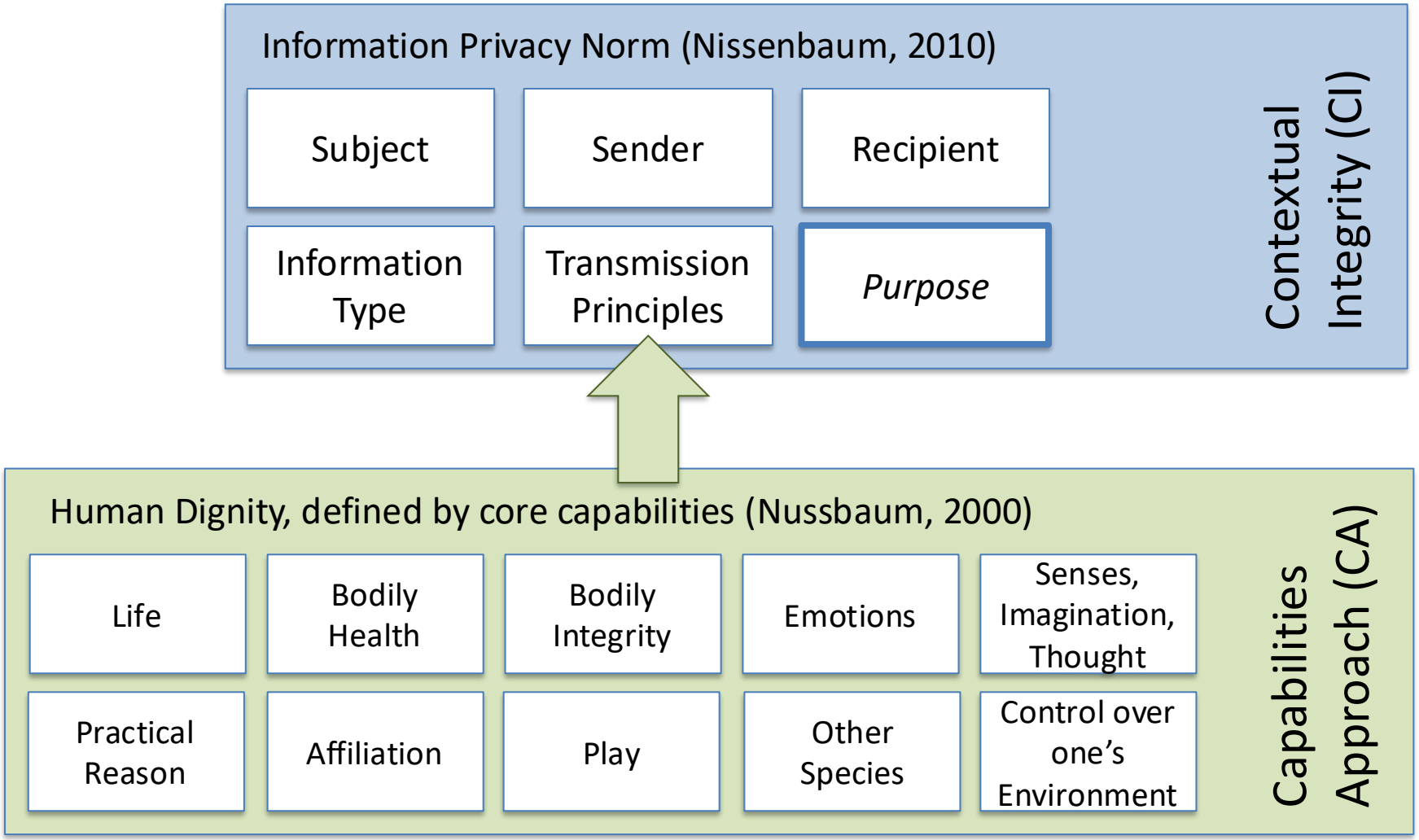


Two extensions to CI:

1. **Purpose explicit** to track information use
2. **Minimum dignity thresholds** based on core capabilities

CA-CI: Integrating Contextual Integrity and the Capabilities Approach for Dignity Considerations in AI Governance
Kat Roemmich, Kirsten Martin, Florian Schaub. IEEE Security & Privacy, 2026

CA-CI: Capabilities Approach–Contextual Integrity



AN INFO FLOW IS PERMISSIBLE WHEN…

1. contextually appropriate
2. dignity is secured

CA-CI: Integrating Contextual Integrity and the Capabilities Approach for Dignity Considerations in AI Governance
Kat Roemmich, Kirsten Martin, Florian Schaub. IEEE Security & Privacy, 2026

Case study:
Foodinho and worker rights

Case study: Foodinho and worker rights

2021

Italian Supervisory Authority Fines Foodinho Over Its Use of Performance Management Algorithms

By Dan Cooper on July 13, 2021

2024

Italy watchdog fines Foodinho 5 million euros for rider data breaches

By **Reuters**

November 22, 2024 3:24 PM EST · Updated November 22, 2024

Case study: Foodinho and worker rights

Transparency and Fairness

Under-specified and inaccurate disclosures of personal **data processing**:

- Riders' geo-location data
- Riders' conversation text and audio
- "Excellence" rider rating system

GDPR TEXT

Article 5.1(a) GDPR. Core principles: lawful, fair, transparent

Case study: Foodinho and worker rights

Transparency and Fairness

Under-specified and inaccurate disclosures of personal **profiling** with “excellence” system:

- Existence
- purpose
- consequences

GDPR TEXT

Article 13 GDPR. Right to be informed of automated decision-making

Case study: Foodinho and worker rights

Transparency and Fairness

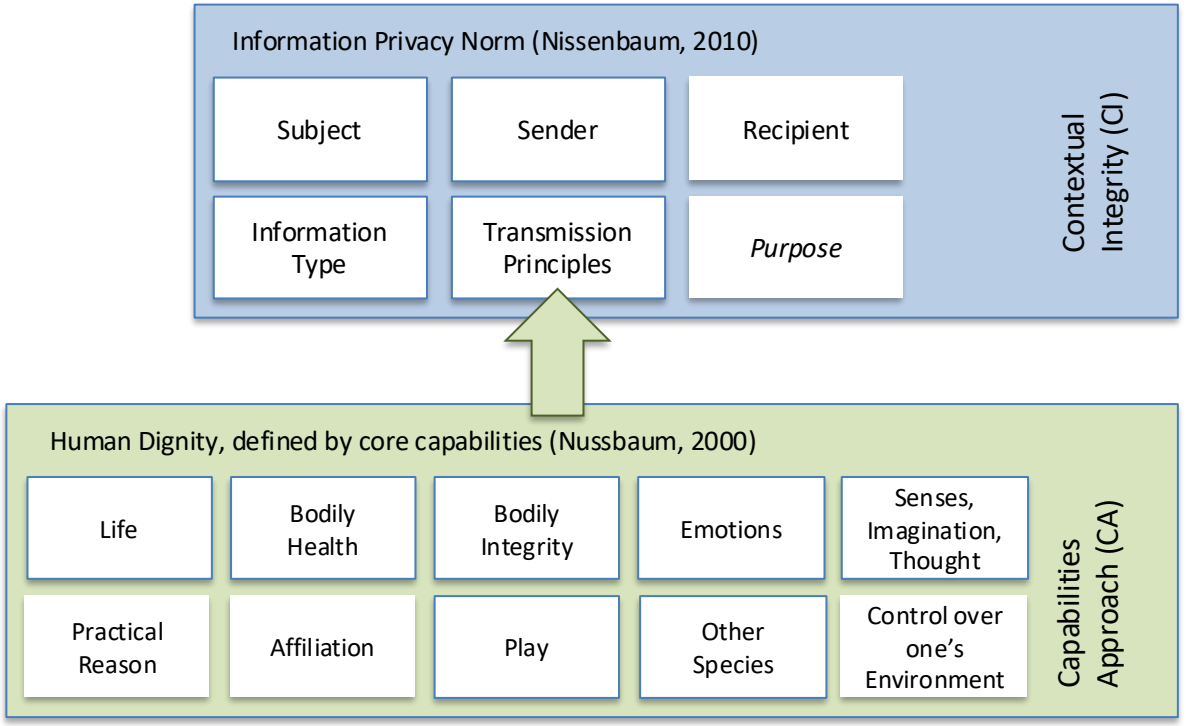
Insufficient safeguards for riders' rights, freedoms, and interests under personal **profiling**:

- human oversight
- self-expression
- contest decisions

GDPR TEXT

Article 22 GDPR. Right not to be subject to automated decision making, including profiling …and safeguards to protect rights and freedoms for exception cases

Case study: Foodinho and worker rights

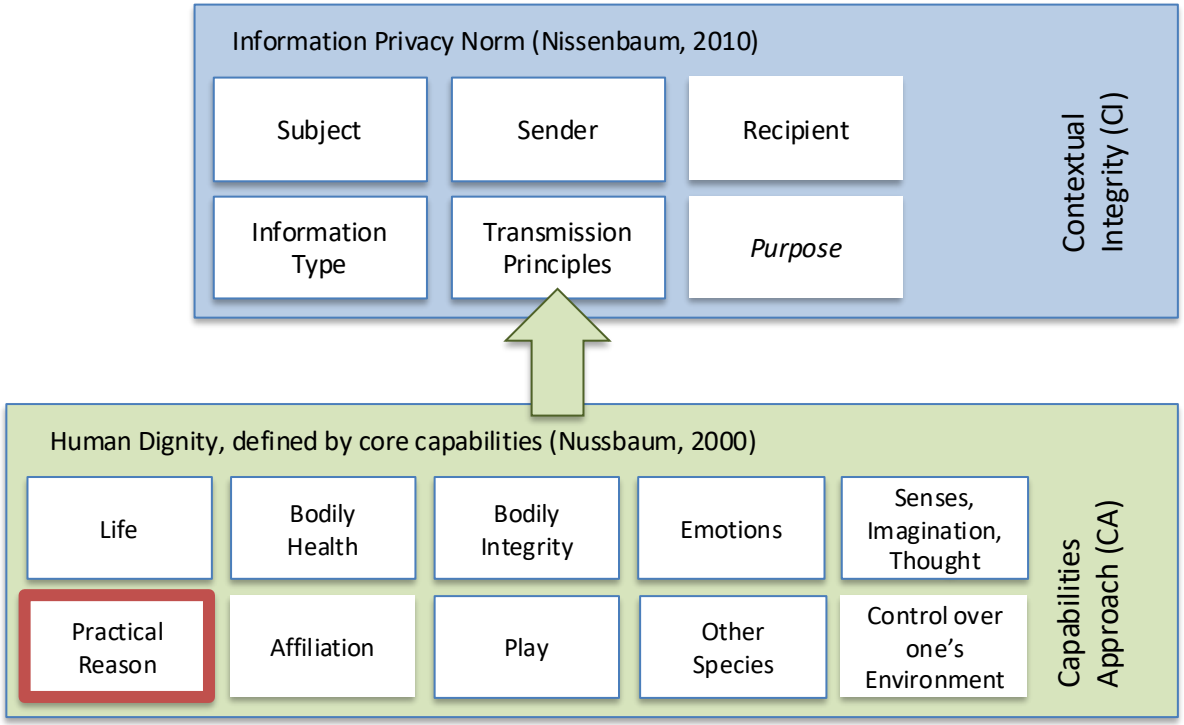


ARTICLE 5(a) Uphold principles of fairness and transparency in data processing

ARTICLE 13 Disclose the existence, purpose, and consequences of personal profiling

ARTICLE 22 Safeguard rights, freedoms, and legitimate interests under profiling

Case study: Foodinho and worker rights



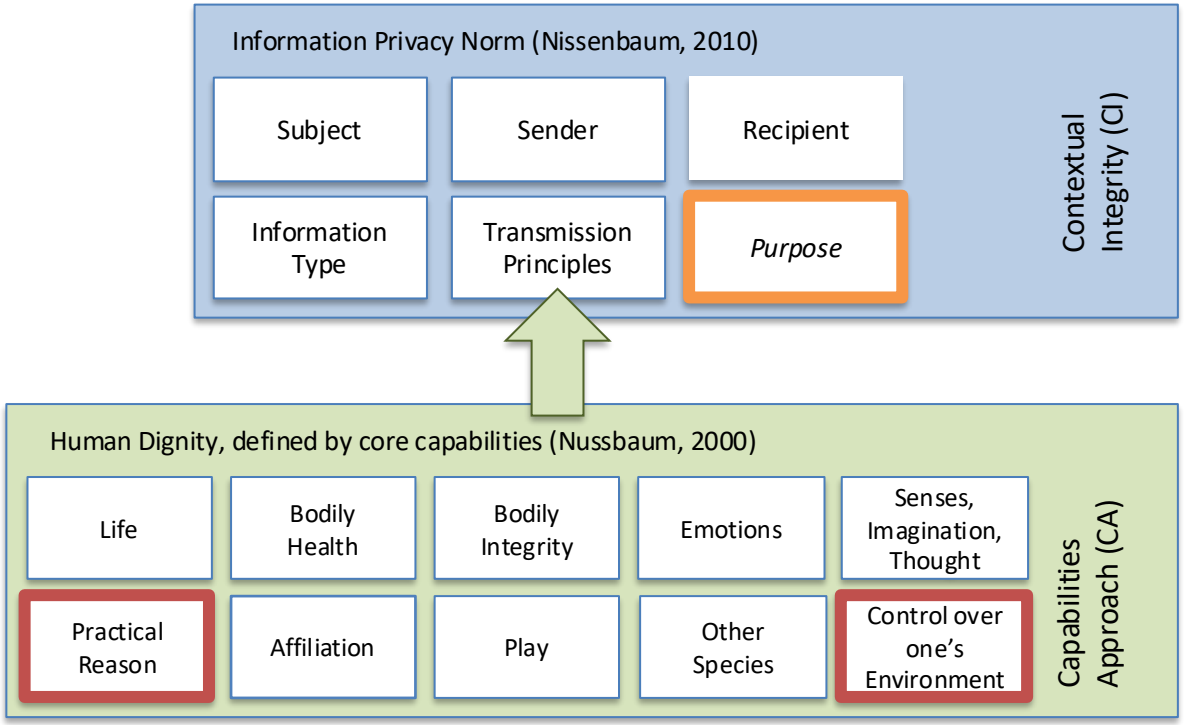
ARTICLE 5(a) Uphold principles of fairness and transparency in data processing

Failure to disclose processing limits capacity to choose and pursue dignified work

ARTICLE 13 Disclose the existence, purpose, and consequences of personal profiling

ARTICLE 22 Safeguard rights, freedoms, and legitimate interests under profiling

Case study: Foodinho and worker rights



ARTICLE 5(a) Uphold principles of fairness and transparency in data processing

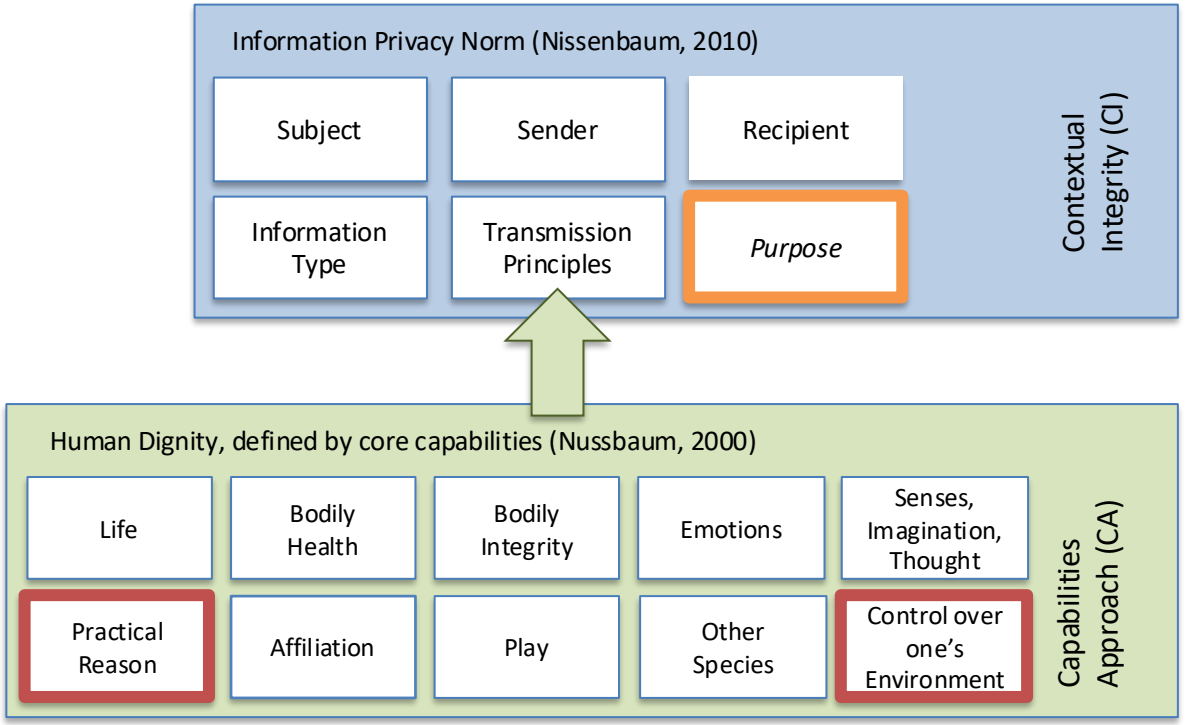
Failure to disclose processing limits capacity to choose and pursue dignified work

ARTICLE 13 Disclose the existence, purpose, and consequences of personal profiling

Shadow profiling of workers for automated employment decisions is indignifying.

ARTICLE 22 Safeguard rights, freedoms, and legitimate interests under profiling

Case study: Foodinho and worker rights



ARTICLE 5(a) Uphold principles of fairness and transparency in data processing

Failure to disclose processing limits capacity to choose and pursue dignified work

ARTICLE 13 Disclose the existence, purpose, and consequences of personal profiling

Shadow profiling of workers for automated employment decisions is indignifying.

ARTICLE 22 Safeguard rights, freedoms, and legitimate interests under profiling

Practical reasoning → meaningful disclosure

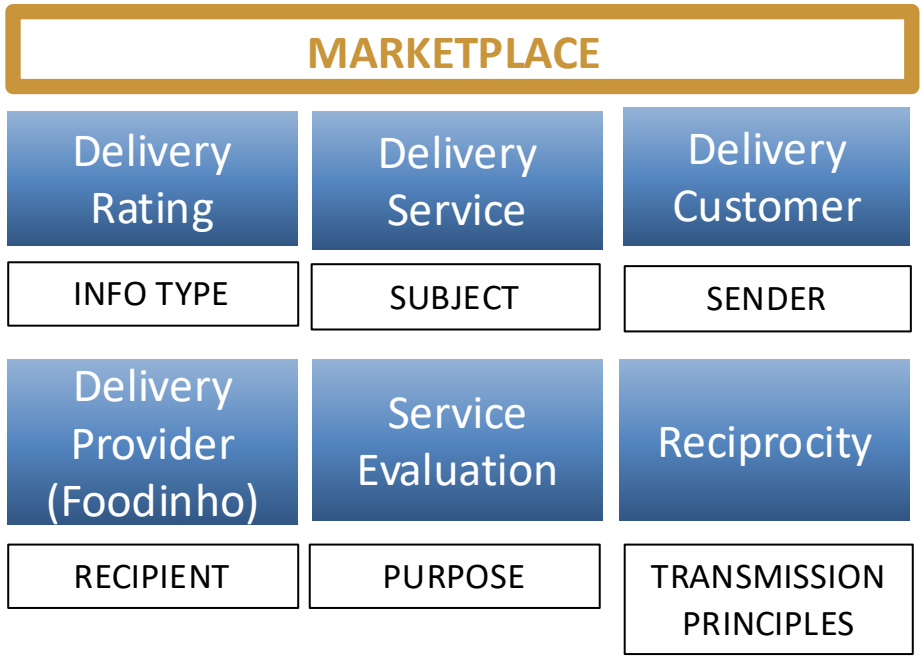
Control over environment → ability to contest decisions

Using CA-CI for assessing information flows

AN INFO FLOW IS PERMISSIBLE WHEN…

1. contextually appropriate
2. dignity is secured

CA-CI Assessment



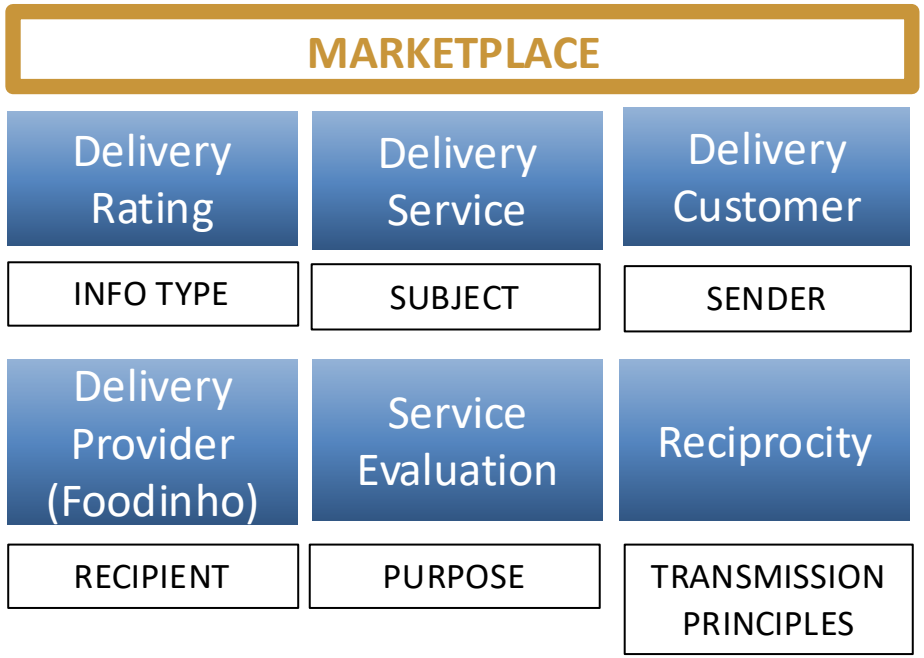
Delivery Service Provider

Relationship between flow and contextual purpose

Data Flow
- Service evaluation

Market Context
- Exchange
- Provisioning
- Coordination
- Mutual benefit

CA-CI Assessment

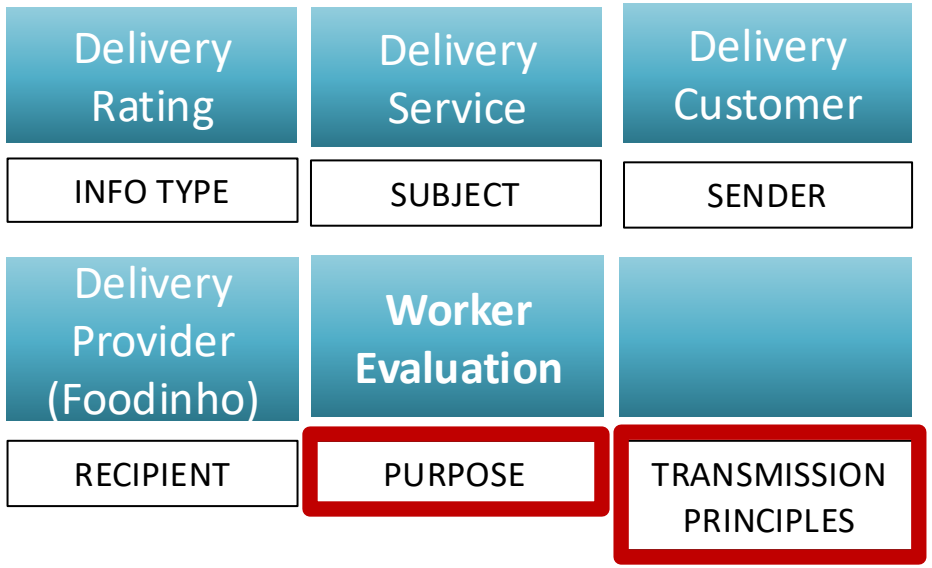


Delivery Service Provider

Relationship between flow and contextual purpose

Data Flow
- Service evaluation

Market Context
- Exchange
- Provisioning
- Coordination
- Mutual benefit



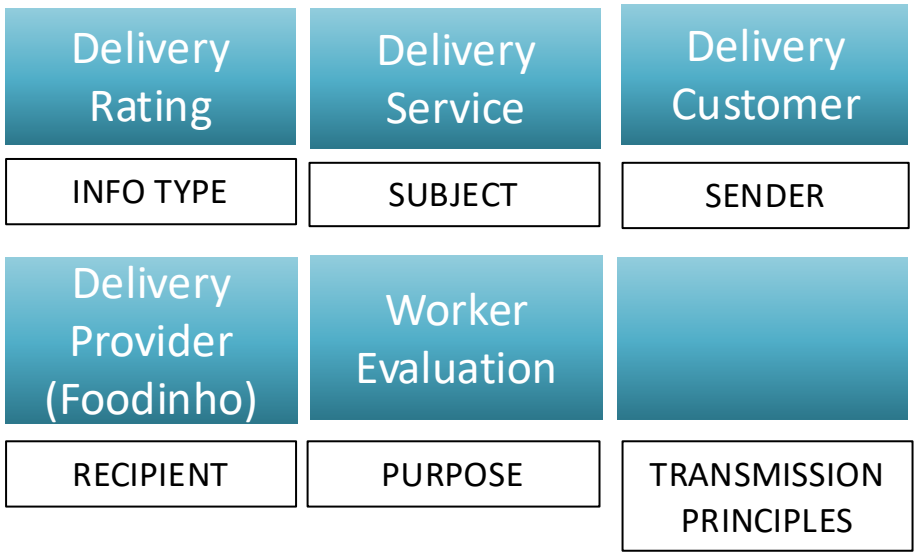
Employer

Relationship between flow and contextual purpose

Data Flow
- **Worker evaluation**

Market Context
- Labor market

CA-CI Assessment



Employer

Ambiguous

Relationship between flow and contextual purpose

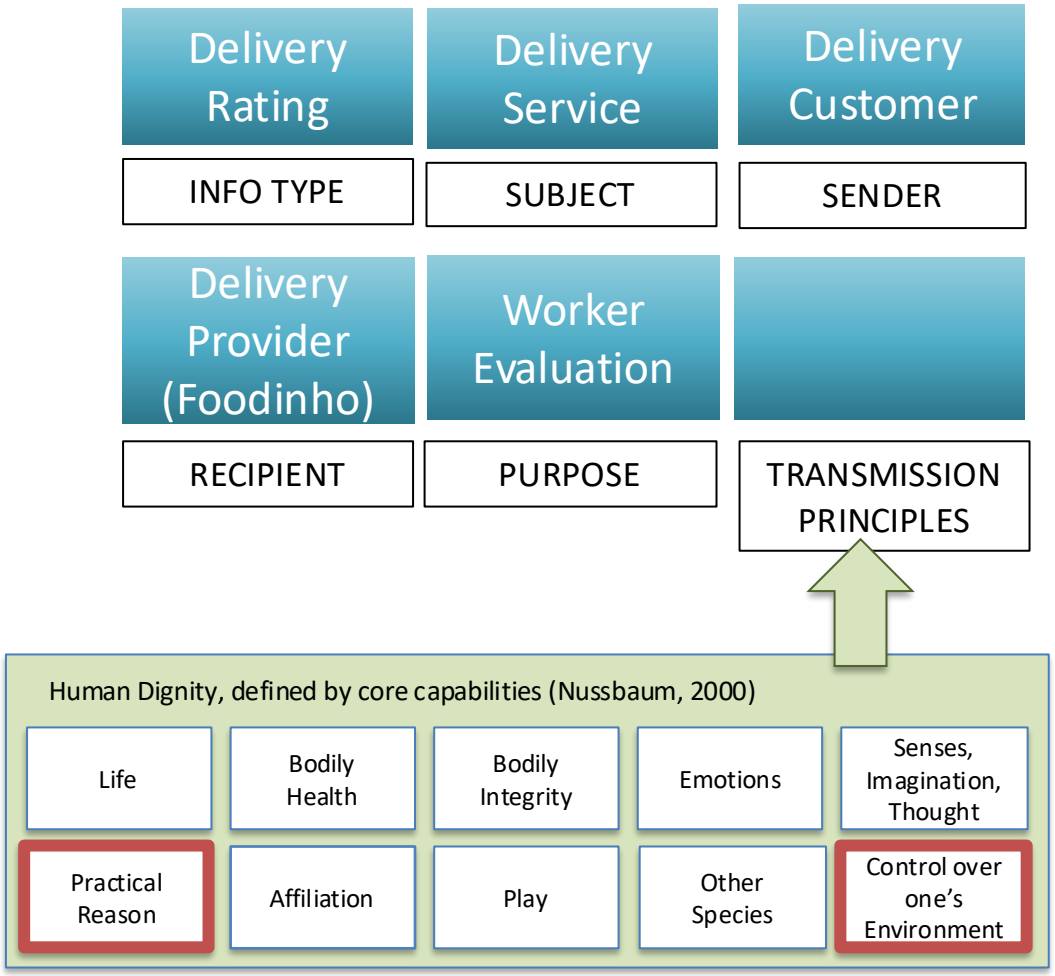
Data Flow

- Worker evaluation

Market Context

- Labor market

CA-CI Assessment



Employer

Ambiguous

Relationship between flow and contextual purpose

Data Flow
- Worker evaluation

Market Context
- Labor market

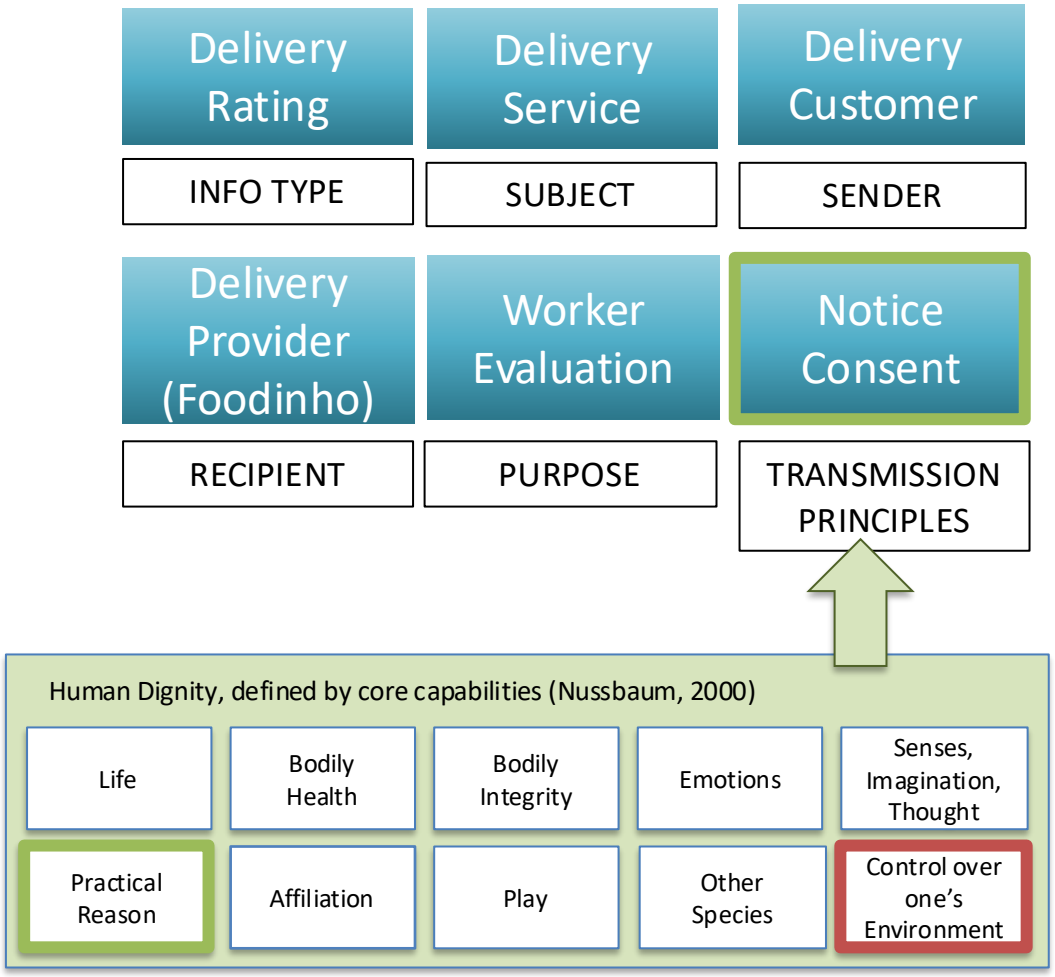
Evaluate dignity risks

- Minimum thresholds
- Mitigation solutions

Design

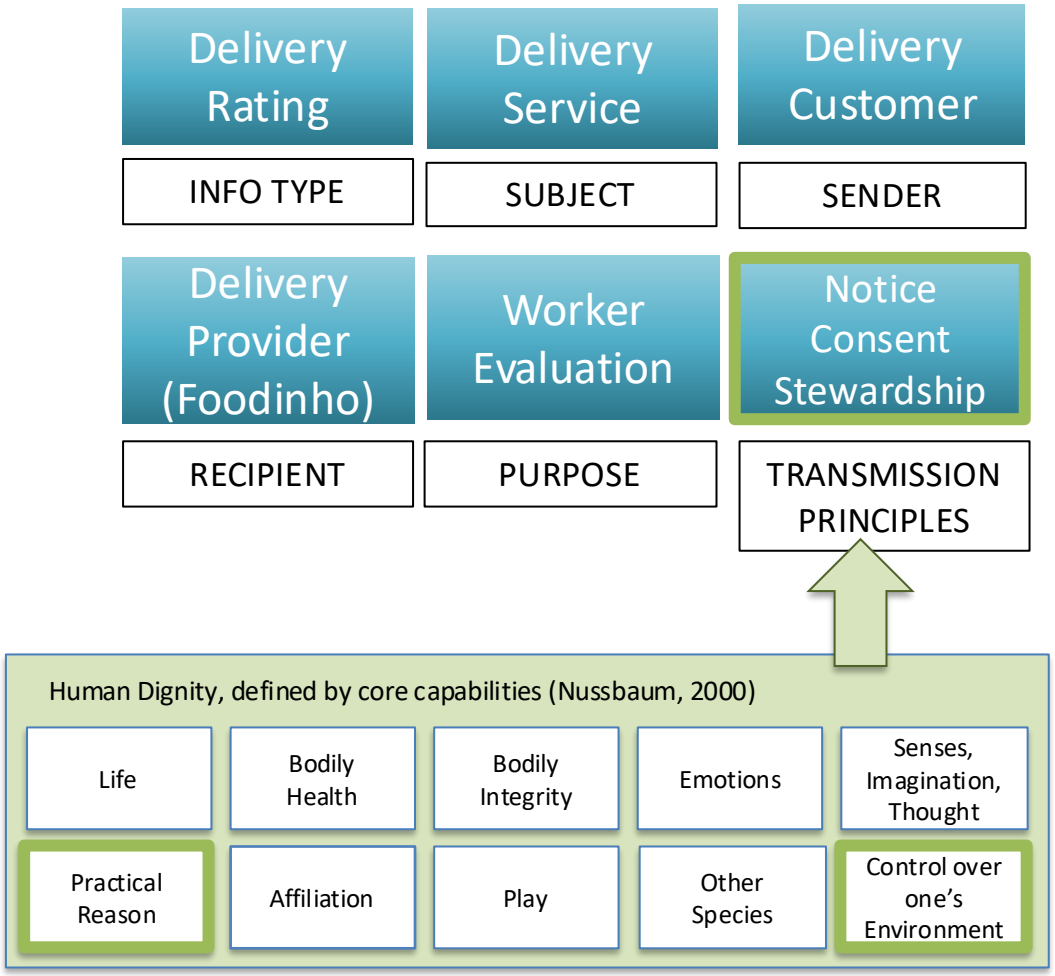
- Re-configure flow parameters
- Add safeguards, privacy enhancements
- Re-evaluate risk
- Develop design requirements & documentation

CA-CI Assessment



Strengthen transmission principles

CA-CI Assessment

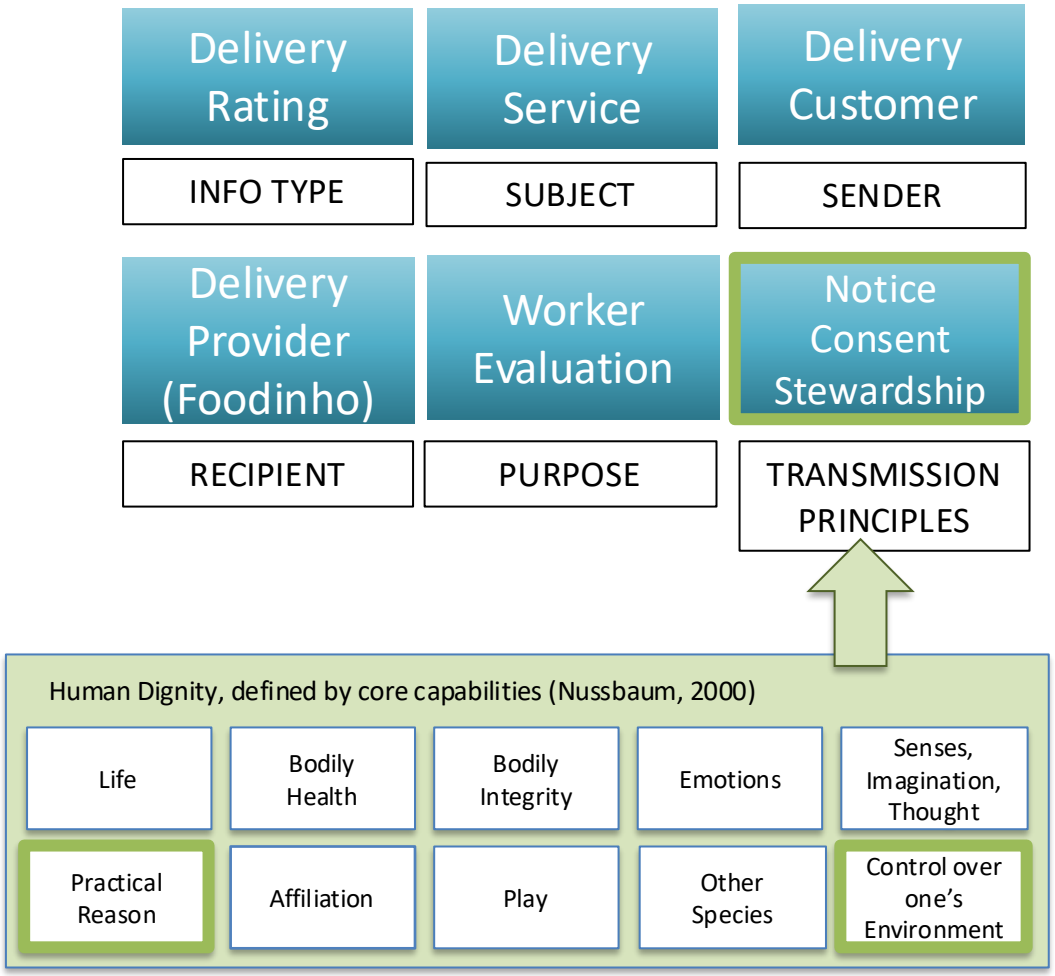


Strengthen transmission principles

Strengthen design requirements

- Worker feedback
- Opportunities to see and contest

CA-CI Assessment



Strengthen transmission principles

Strengthen design requirements

- Worker feedback
- Opportunities to see and contest

Strengthen privacy and AI governance

Towards CA-CI in practice

Dignity as a *minimum* standard

From harm taxonomies to core principles to human rights and freedoms

Dignity = threshold for what counts

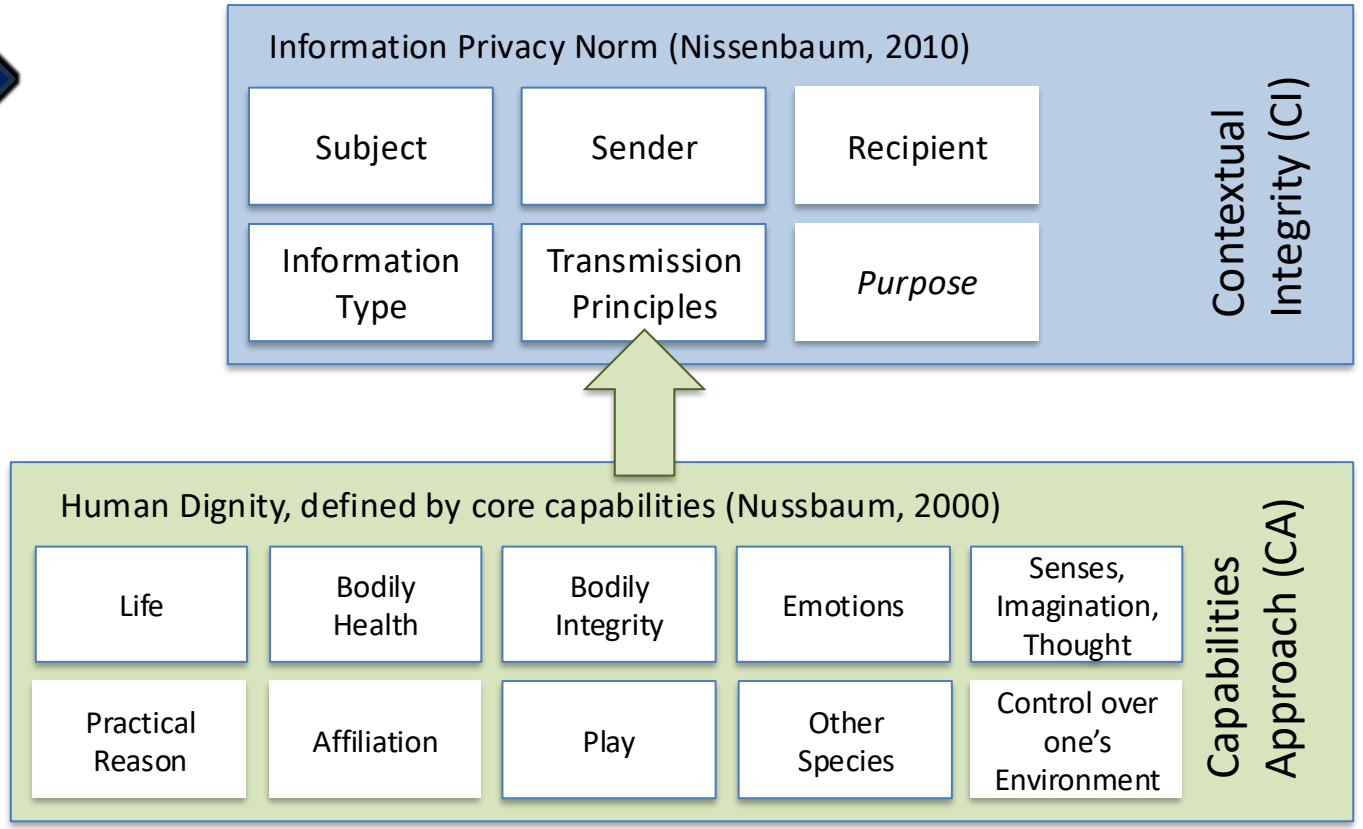
Practical risk assessment and governance

Privacy by design, DPIAs/FRIAs, decision rules for safeguards, risk monitoring & mitigation

Future directions

Impact assessment templates

Purpose limitation enforcement & monitoring



Towards CA-CI in practice

Dignity as a *minimum* standard

From harm taxonomies to core principles to human rights and freedoms

Dignity = threshold for what counts

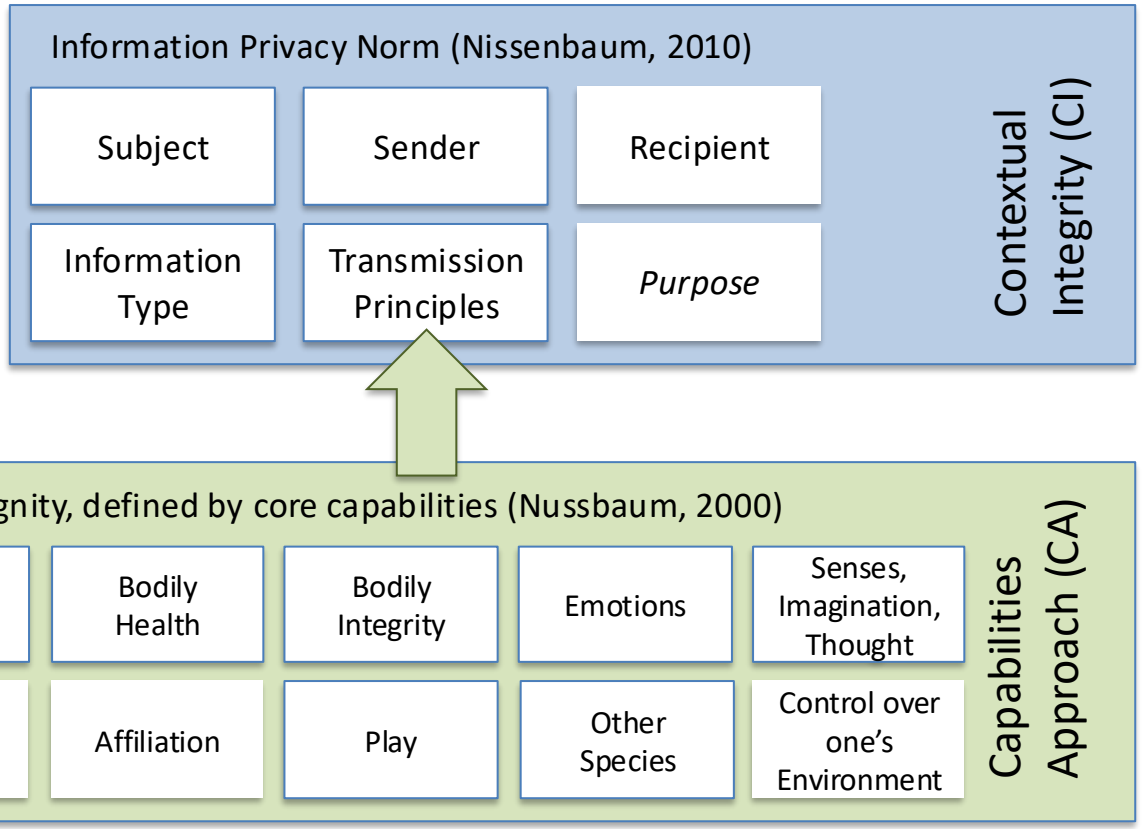
Practical risk assessment and governance

Privacy by design, DPIAs/FRIAs, decision rules for safeguards, risk monitoring & mitigation

Future directions

Impact assessment templates

Purpose limitation enforcement & monitoring



CA-CI: Integrating Contextual Integrity and the Capabilities Approach for Dignity Considerations in AI Governance

Kat Roemmich, Kirsten Martin, Florian Schaub

IEEE Security & Privacy, 2026