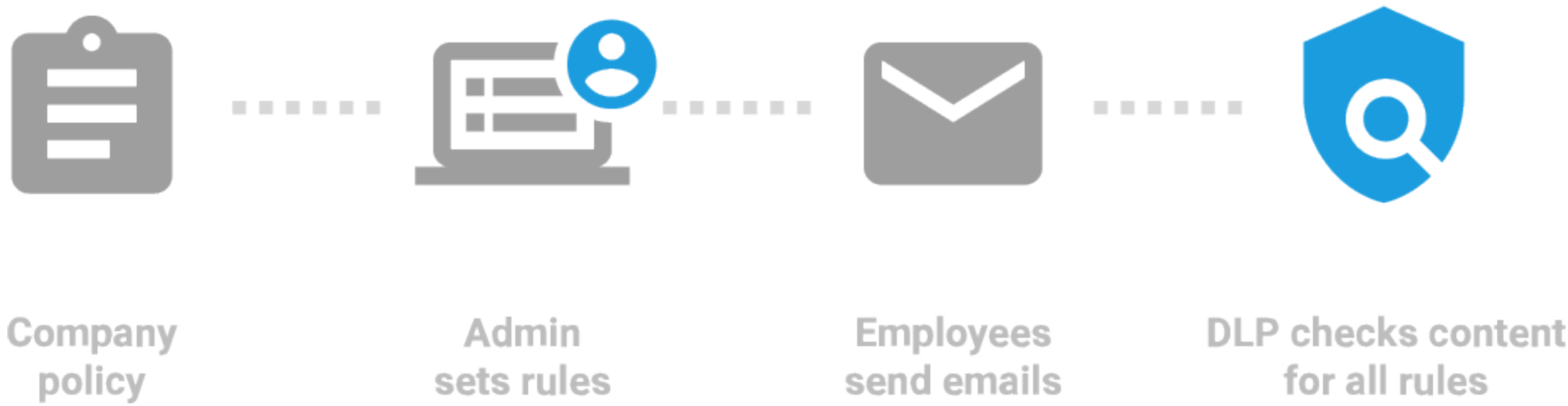
The VACCINE Framework for Building DLP Systems

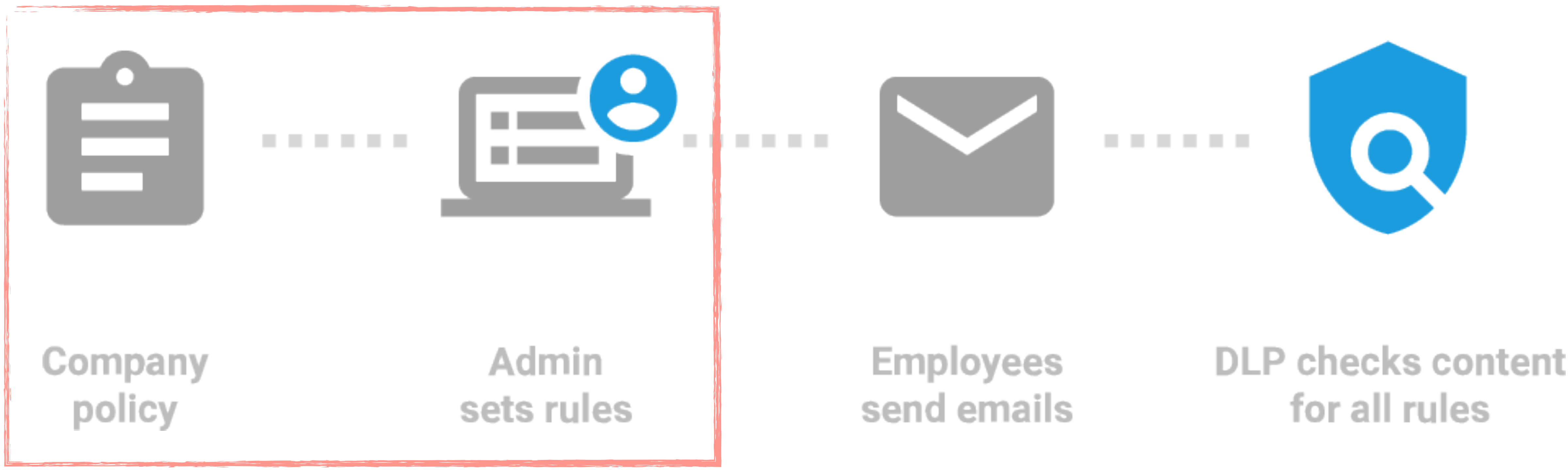
Yan Shvartzshnaider (NYU/CITP, Princeton), Zvonimir Pavlinovic (NYU),
Thomas Wies (NYU), Lakshminarayanan Subramanian (NYU),
Prateek Mittal (Princeton), and Helen Nissenbaum (Cornell Tech)

Data Leakage Protection



Aimed at prevent an accidental or unintentional distribution of private or sensitive data to an unauthorized entity.

Data Leakage Protection



1. Admin self interpreting policies from handbooks to specify rules (which is error-prone)

2. Conflating extraction sensitive data (regular expressions templates, keywords, or patterns) and enforcement of policies

VACCINE: Verifiable and ACtionable Contextual Integrity Norms Engine

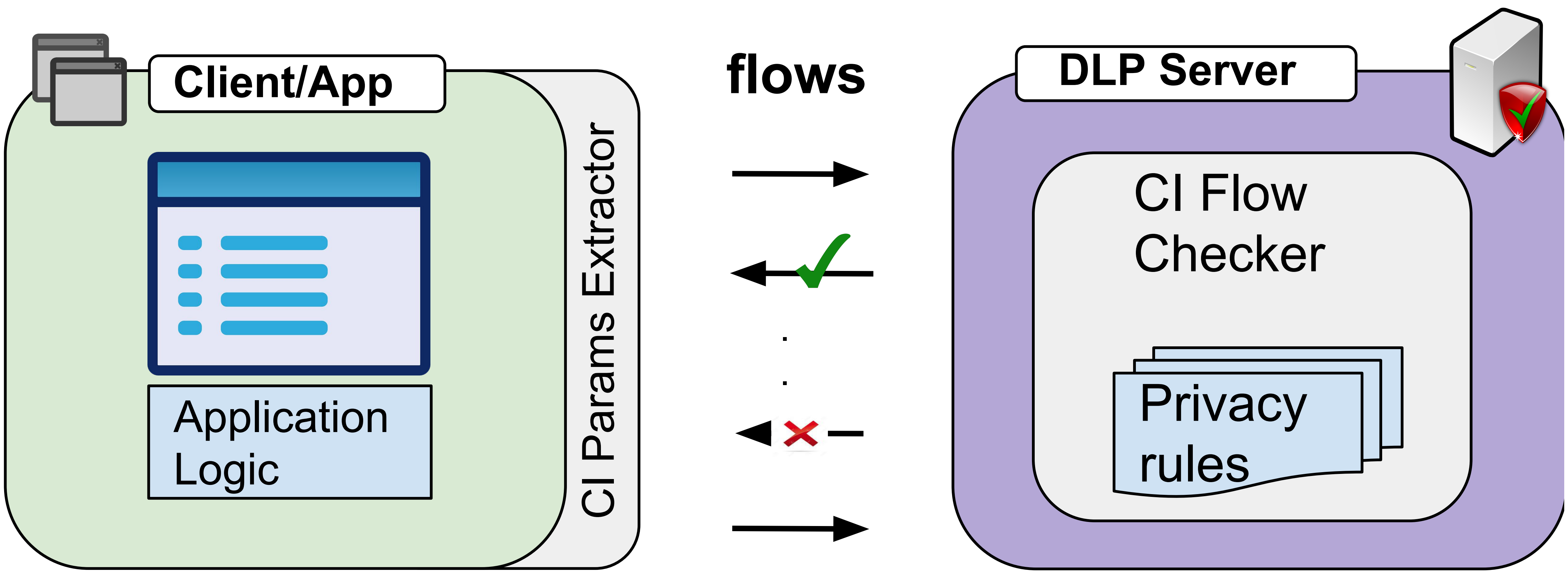
- Uses Contextual Integrity to model the information flows and the notion of information leakage
    - A flow represents an atomic unit of an information exchange

      <sender, recipient, subject, attribute>
    - Contextual Informational Norms specify what flows are allowed in a given privacy context constrained by transmission principles.

      <Sender, Recipient, Subject, Attribute, Transmission principle>
- Norm violation serves as the definition of information leakage.

VACCINE Architecture

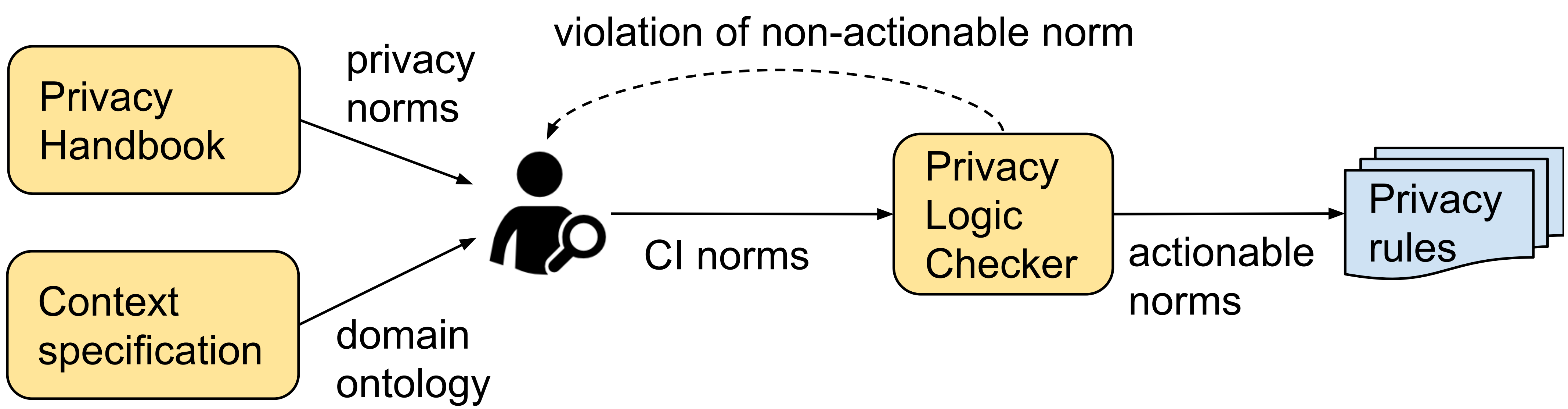
The system allows exactly those flows that adhere to the given contextual information norms



Privacy Logic

- **Actionable norms:** specify the operational rules that define the runtime behavior of the system
  - Example: Professor may not disclose student's educational record to parents without the student’s permission.
    - allowed (FERPA$_{ctx}$, Professor$_{sndr}$, ER$_{attr}$, Student$_{subj}$, Parents$_{rcp}$) explicit_permission (Student$_{subj}$, PO$_{attr}$, Agent$_{rcp}$)
- **Non-actionable norms:** define auxiliary properties that must be guaranteed by the actionable norms e.g., *Implicit Norms, Blocking Norms.*
  - Example (implicit norm): *A student should be able to send herself a message containing her own personal data.*

VACCINE Architecture
Extracting Privacy Logic



- Extracts the privacy logic in the form of actionable and non-actionable norms from a privacy handbook and checking their consistency
  - ie., make sure actionable norms don't violate non-actionable ones
- Norms are translated into operational rules that can be enforced by an engine

Evaluation Questions

- **How formal methods can assist in the creation of a consistent set of privacy rules?**
  - Manually extracted privacy norms from the FERPA summary actionable and non-actionable norms
    - It took three iterations of the check-refine loop to obtain a consistent set of actionable norms
  - If a particular non-actionable norm is violated, the theorem prover (Z3) will produce a model describing a sequence of information flows that respects the rules but violates the norm.
    - Using this model, we can then identify the rules that are responsible for the violation.

Evaluation Questions

- **How efficiently can CI flows be checked against the privacy rules?**
  - Checking whether a flow complies with the privacy logic amounts to performing a single query of the *allowed* predicate.
- **How effective is the VACCINE framework in preventing potentially unauthorized flows in a real-world emulated context?**
  - We created 43 Enron privacy rules that focus on access and disclosure of PII in a corporate setting.

Lessons & Future Work

- Lessons
  - Privacy regulations are not written with CI in mind e.g., lots of assumption about implicit flows
  - CI is not well understood outside the legal and privacy scholars communities
- Future work
  - Automate privacy logic extraction