list
cea tech

Séminaire LSL
1er décembre 2022

Model Specification Language for Formal Verification of Consent Properties on Models and Code

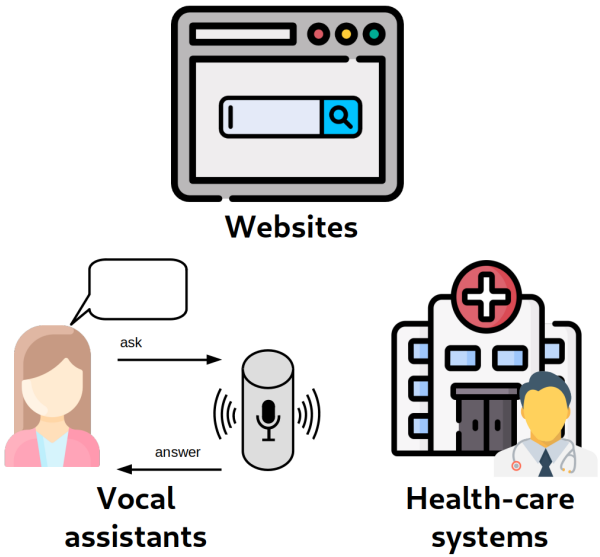
Myriam Clouet

INSTITUT CARNOT TN@UPSaclay
université PARIS-SACLAY

Introduction

Privacy in information systems

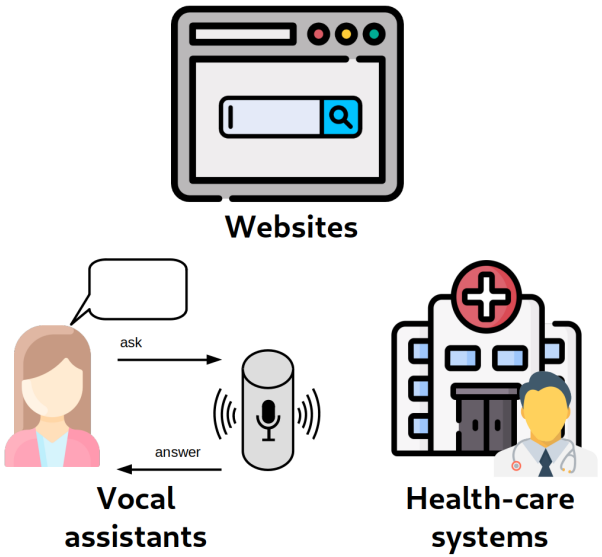
Application areas



Personal data processing

Privacy in information systems

Application areas



Laws & Regulations

- GDPR
- Australian Privacy Act
- Japanese Privacy Act
- ...

Personal data processing

Mandatory!

Judgments for non-compliance with the GDPR

Google

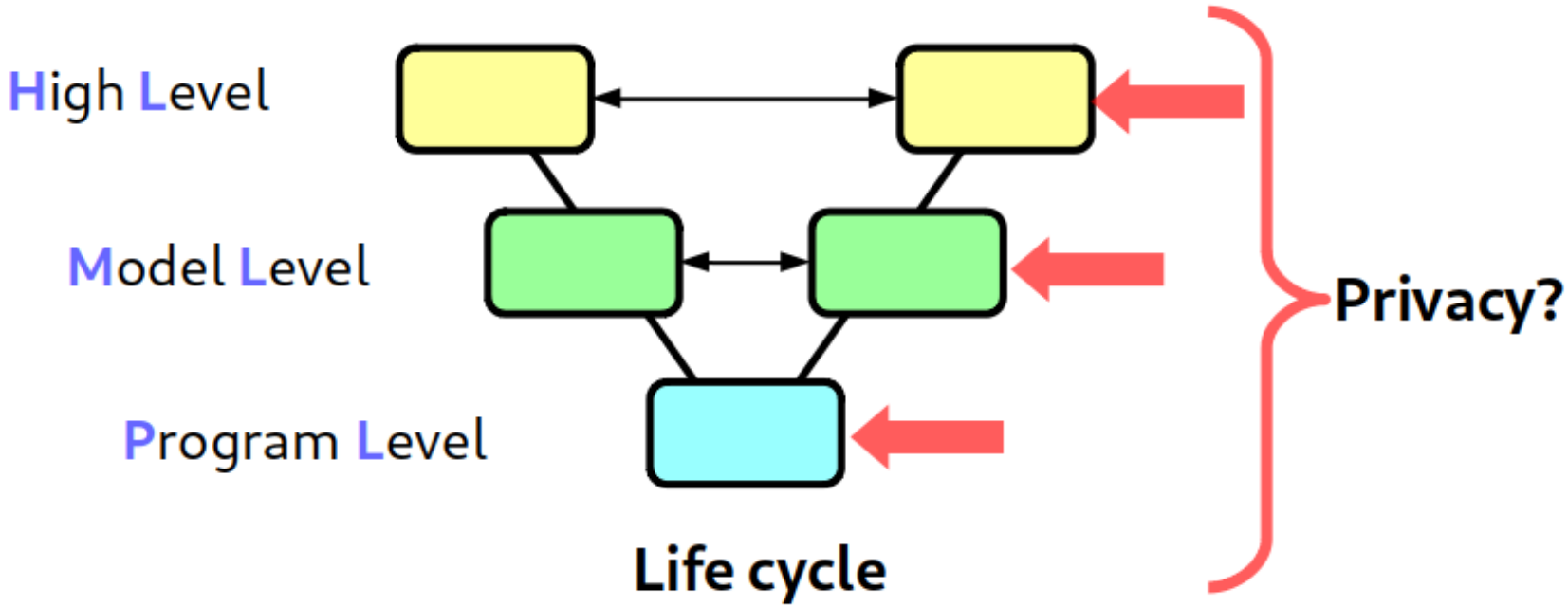
50 million € (2019) [11]

WhatsApp

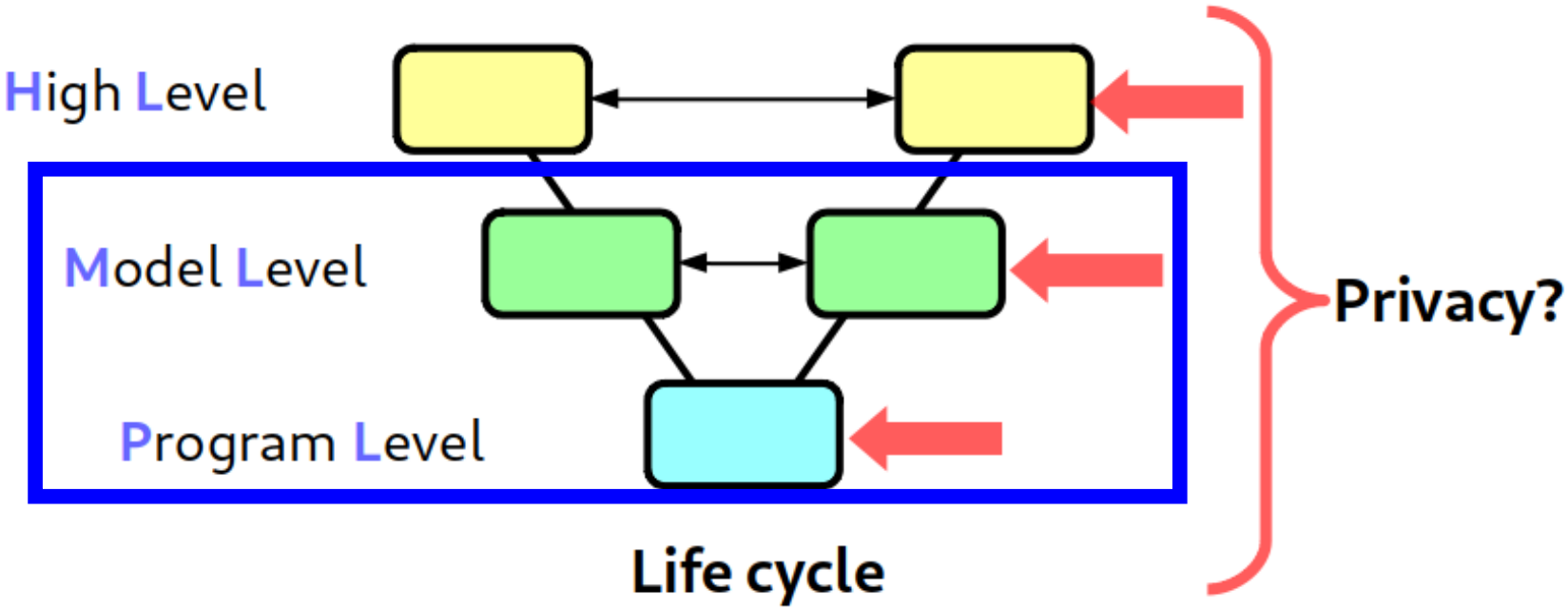
225 million € (2021) [8]

Fines for **invalid consent**

Privacy verification - Complex



Privacy verification - Complex



Goal & Contributions

🎯 ML & PL consent properties: same **formal language**

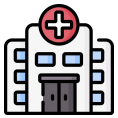
Contributions

- **CSpeL**: **C**onsent **S**pecification **L**anguage
- **CASTT**: **C**onsent **A**nnotation **S**pecification **T**ranslation **T**ool

Running Example

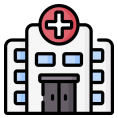
Healthcare

Based on an example of Petkovic et al. [12]

 Hospital information system processing patients's data

Healthcare

Based on an example of Petkovic et al. [12]

 Hospital information system processing patients's data

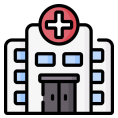
For two **purposes**:

 providing treatment to patients (**Treatment**)

 performing a clinical trial (**Research**).

Healthcare

Based on an example of Petkovic et al. [12]

 Hospital information system processing patients's data

For two **purposes**:

 providing treatment to patients (**Treatment**)

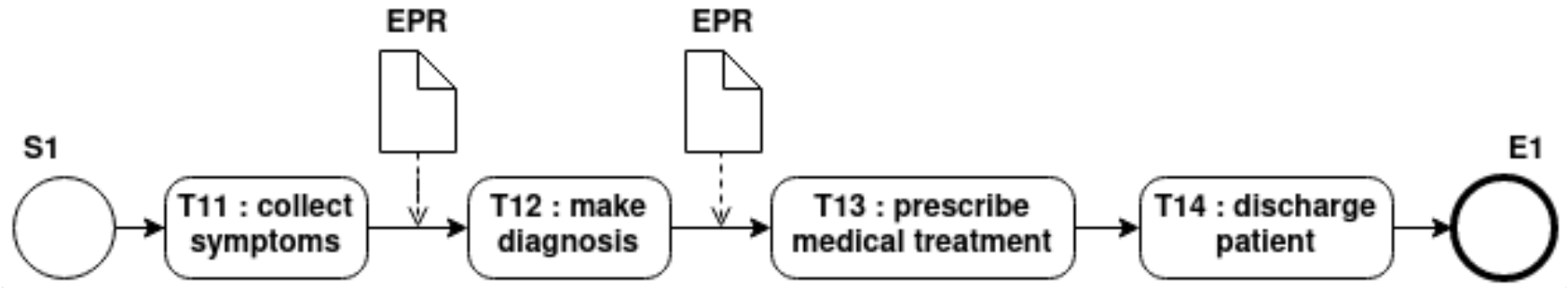
 performing a clinical trial (**Research**).

Consented: processing personal data only for providing treatment

Running Example

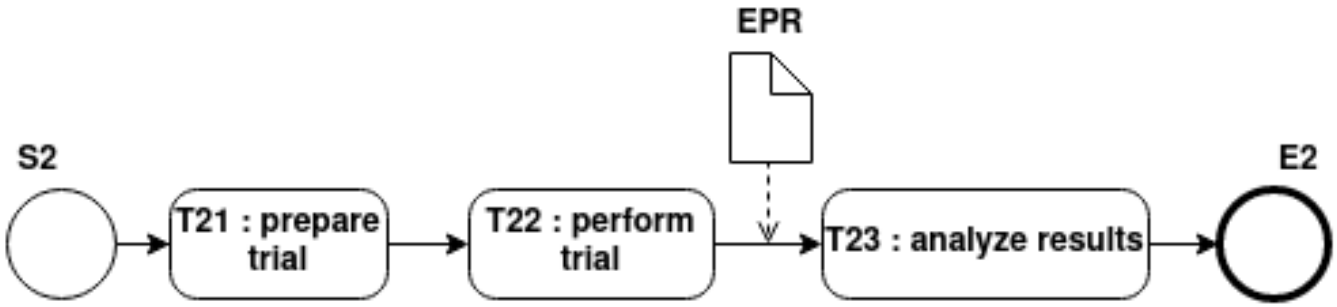
At Model Level - BPMN [5]

Process P1 - Treatment Patient



EPR: **E**lectronic **P**atient **R**ecord

Process P2 - Clinical Trial



EPR: **E**lectronic **P**atient **R**ecord

Running Example

At Program Level - C code

Code - Data

Refinements

ML	ELEMENT	NAME
	Data Object	EPR

PL	ELEMENT	NAME
	typedef struct	Patient(1), TrialData(1), Prescr

Data

```
typedef struct{
   char date[S_SIZE];
   char medecine[S_SIZE];
} Prescr ;

typedef struct{
   int id;
   char name[S_SIZE];
   char birthdate[S_SIZE];
   Prescr prescrList[NB_PRESCR];
   ...
} Patient ;

typedef struct{
   char birthdateList[NB_PATIENT][S_SIZE];
   char sexList[NB_PATIENT][S_SIZE];
   Prescr prescrList[NB_PRESCR];
} TrialData ;
```

(1) personal data

Code - Functions

Refinements

	ELEMENT	NAME
ML	BPM(1)	P1
	BPM(1)	P2

	ELEMENT	NAME
PL	function	makePrescr
	function	getData, computeStats

Functions

```
Patient makePrescr(Prescr  newp, Patient p){
   ...
}

TrialData getData(Patient  patientList[NB_PATIENT]){
   ...
}

int computeStats(TrialData  data){
   ...
}
```

(1) Business Process Model

CSpeL

Goal

CSpeL Use

CSpeL

Context model

Context model - Definition

$S \triangleq$ set of the system's processes

$D \triangleq$ set of the personal data

$P \triangleq$ set of the purposes

Context model - Definition

$$
\begin{aligned}
S &\triangleq \text{set of the system's processes} \\
D &\triangleq \text{set of the personal data} \\
P &\triangleq \text{set of the purposes} \\
\gamma &\triangleq D \times P \rightarrow \text{bool} \\
\pi &\triangleq S \rightarrow \mathcal{P}(P) \\
\nu &\triangleq S \rightarrow \mathcal{P}(D)
\end{aligned}
$$

γ: Consent granting

π: Purposes of processing

ν: Data needed for processing

Running Example - ML

Running Example - ML

$$S \triangleq \{\text{P1;P2}\}$$

$$D \triangleq \{\text{EPR}\}$$

$$P \triangleq \{\text{Treatment;Research}\}$$

Running Example - ML

$$S \triangleq \{\text{P1;P2}\}$$

$$D \triangleq \{\text{EPR}\}$$

$$P \triangleq \{\text{Treatment;Research}\}$$

$$\gamma \triangleq \begin{cases} (EPR, Treatment) \mapsto \top \\ (EPR, Research) \mapsto \bot; \end{cases}$$

Running Example - ML

$$S \triangleq \{\text{P1;P2}\}$$

$$D \triangleq \{\text{EPR}\}$$

$$P \triangleq \{\text{Treatment;Research}\}$$

$$\gamma \triangleq \begin{cases} (EPR, Treatment) \mapsto \top \\ (EPR, Research) \mapsto \bot; \end{cases}$$

$$\pi \triangleq \begin{cases} P1 \mapsto \{Treatment\} \\ P2 \mapsto \{Research\}; \end{cases}$$

Running Example - ML

$$S \triangleq \{\text{P1;P2}\}$$

$$D \triangleq \{\text{EPR}\}$$

$$P \triangleq \{\text{Treatment;Research}\}$$

$$\gamma \triangleq \begin{cases} (EPR, Treatment) \mapsto \top \\ (EPR, Research) \mapsto \bot; \end{cases}$$

$$\pi \triangleq \begin{cases} P1 \mapsto \{Treatment\} \\ P2 \mapsto \{Research\}; \end{cases}$$

$$\nu \triangleq \begin{cases} P1 \mapsto \{EPR\} \\ P2 \mapsto \{EPR\}. \end{cases}$$

Running Example - PL

Running Example - PL

$$S \triangleq \{\texttt{makePrescr;getData;computeStats}\}$$

$$D \triangleq \{\text{Patient, TrialData}\}$$

$$P \triangleq \{\text{Treatment,Research}\}$$

Running Example - PL

$$S \triangleq \{\texttt{makePrescr;getData;computeStats}\}$$

$$D \triangleq \{\text{Patient, TrialData}\}$$

$$P \triangleq \{\text{Treatment,Research}\}$$

$$\gamma \triangleq \begin{cases} (\mathit{Patient}, \mathit{Treatment}) & \mapsto \top \\ (\mathit{Patient}, \mathit{Research}) & \mapsto \bot \\ (\mathit{TrialData}, \mathit{Treatment}) & \mapsto \bot \\ (\mathit{TrialData}, \mathit{Research}) & \mapsto \bot \end{cases}$$

Running Example - PL

$$S \triangleq \{\texttt{makePrescr;getData;computeStats}\}$$

$$D \triangleq \{\text{Patient, TrialData}\}$$

$$P \triangleq \{\text{Treatment,Research}\}$$

$$\gamma \triangleq \begin{cases} (\mathit{Patient}, \mathit{Treatment}) & \mapsto \top \\ (\mathit{Patient}, \mathit{Research}) & \mapsto \bot \\ (\mathit{TrialData}, \mathit{Treatment}) & \mapsto \bot \\ (\mathit{TrialData}, \mathit{Research}) & \mapsto \bot \end{cases}$$

$$\pi \triangleq \begin{cases} \texttt{makePrescr} & \mapsto \{\mathit{Treatment}\} \\ \texttt{getData} & \mapsto \{\mathit{Research}\} \\ \texttt{computeStats} & \mapsto \{\mathit{Research}\}; \end{cases}$$

Running Example - PL

$$S \triangleq \{\texttt{makePrescr;getData;computeStats}\}$$

$$D \triangleq \{\text{Patient, TrialData}\}$$

$$P \triangleq \{\text{Treatment,Research}\}$$

$$\gamma \triangleq \begin{cases} (\mathit{Patient}, \mathit{Treatment}) & \mapsto \top \\ (\mathit{Patient}, \mathit{Research}) & \mapsto \bot \\ (\mathit{TrialData}, \mathit{Treatment}) & \mapsto \bot \\ (\mathit{TrialData}, \mathit{Research}) & \mapsto \bot \end{cases}$$

$$\pi \triangleq \begin{cases} \texttt{makePrescr} & \mapsto \{\mathit{Treatment}\} \\ \texttt{getData} & \mapsto \{\mathit{Research}\} \\ \texttt{computeStats} & \mapsto \{\mathit{Research}\}; \end{cases}$$

$$\nu \triangleq \begin{cases} \texttt{makePrescr} & \mapsto \{\mathit{Patient}\} \\ \texttt{getData} & \mapsto \{\mathit{Patient}, \mathit{TrialData}\} \\ \texttt{computeStats} & \mapsto \{\mathit{TrialData}\}. \end{cases}$$

CSpeL

Execution Trace

Execution Trace - Definition

$$T ::= \epsilon \mid \mathit{Handle}(\sigma, d); T.$$

$\epsilon \triangleq$ **empty trace**

$\mathit{Handle}(\sigma, d) \triangleq$ σ **process the data** d

$; \triangleq$ **concatenation operator**

Running Example - ML

Execution **P1**:

$Handle(P1, EPR); \epsilon$

Execution **P2**:

$Handle(P2, EPR); \epsilon$

Running Example - PL

 Execution of **`makePrescr`**:
Handle(`makePrescr`, *Patient*); *Handle*(`makePrescr`, *Prescr*); ϵ

 Execution of **`getData`** then execution of **`computeStats`**:
Handle(`getData`, *Patient*); *Handle*(`getData`, *TrialData*);
Handle(`computeStats`, *TrialData*); ϵ

CSpeL

Property

Property - Definition

Purpose Compliance: "a user **agreed** to **at least one** of the **process's purposes** before his personal data is **handled** by this process."

$$\mathcal{C} \vdash \mathit{Handle}(\sigma, d) \iff \begin{cases} d \notin D \\ \exists p \in \pi(\sigma), \gamma(d, p) = \top. \end{cases} \text{; or}$$

Property - Inference Rules

$$\frac{}{\mathcal{C} \vdash \epsilon}$$

Empty Trace

$$\frac{\mathcal{C} \vdash \mathit{Handle}(\sigma, d) \quad \mathcal{C} \vdash T}{\mathcal{C} \vdash \mathit{Handle}(\sigma, d); T}$$

Trace **Concatenation**

Running Example - ML

$Handle(P1, EPR); \epsilon \rightarrow$ purpose compliant

$Handle(P2, EPR); \epsilon \rightarrow$ not purpose compliant

Because

- $D = \{EPR\}$
- $\pi(P1) = \{Treatment\}$
- $\pi(P2) = \{Research\}$
- $\gamma(EPR, Treatment) = \top$
- $\gamma(EPR, Research) = \bot$

CSpeL - Syntax

Running Example - ML

Model	Syntax
$S \triangleq \{P1;P2\}$	\process{P1,P2}
$D \triangleq \{EPR\}$	\personalData{EPR}
$P \triangleq \{Treatment;Research\}$	\purposes{Treatment, Research}
$\gamma \triangleq \begin{cases} (EPR, Treatment) \mapsto \top \\ (EPR, Research) \mapsto \bot; \end{cases}$	\isGranted{(EPR: Treatment)}
$\pi \triangleq \begin{cases} P1 \mapsto \{Treatment\} \\ P2 \mapsto \{Research\}; \end{cases}$	\hasPurposes{(P1:{Treatment}), (P2:{Research})}
$\nu \triangleq \begin{cases} P1 \mapsto \{EPR\} \\ P2 \mapsto \{EPR\}. \end{cases}$	\needData{(P1:{EPR}), (P2:{EPR})}
$Handle$(P1,EPR);ϵ	\handle(P1, EPR); VOID

CASTT

Goal

Goal

Two uses:

- Trace Analysis
- Translation for a PL Tool

CASTT

Trace Analysis

Use

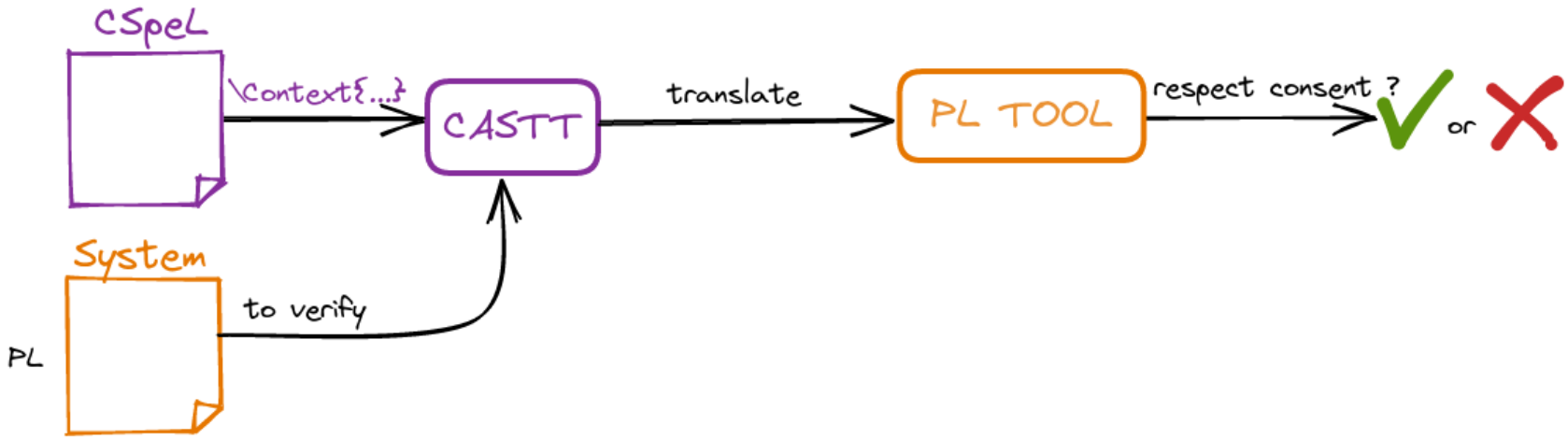
Experiment Results - Trace Analysis

Example	LVL	nbTests	valid traces detected	invalid traces detected
Healthcare	ML	4	✔	✔
	PL	4	✔	✔
Website	ML	5	✔	✔
	PL	3	✔	✔
total		16	8/8	8/8

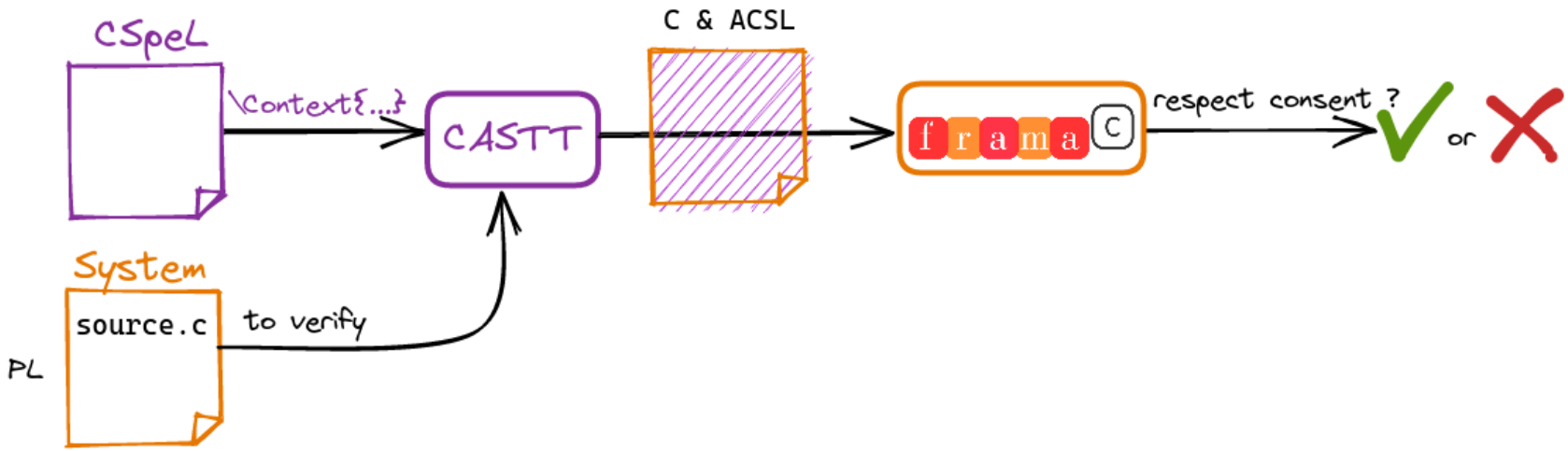
CASTT

CASTT- Translation

Use


CSpeL
\context{...}
CASTT
translate
PL TOOL
respect consent ?
or
System
to verify
PL

Implementation



Code Generation - C

CSpeL	Generated code
\personalData	`enum _PersonalData {TRIALDATA = 0, PATIENT = 1}` `typedef enum _PersonalData PersonalData;`
\purposes	`enum _Purposes {RESEARCH = 0, TREATMENT = 1}` `typedef enum _Purposes Purposes;`
\personalData & \purposes	`/*@ ghost bool Consent[2][2];*/`

\isGranted & \init	`/*@ ghost Consent[PATIENT][RESEARCH] = 0; */`
	`/*@ ghost Consent[PATIENT][TREATMENT] = 1; */`
	`/*@ ghost Consent[TRIALDATA][RESEARCH] = 0; */`
	`/*@ ghost Consent[TRIALDATA][TREATMENT] = 0; */`

Code Generation - ACSL

CSpeL	Generated code
\process & \hasPurposes & \needData	`/*@ requires Consent[PATIENT][TREATMENT] ≡ 1; */`

Functions with ACSL

```
/*@ requires Consent[PATIENT][TREATMENT] ≡1;
   ...
*/
Patient makePrescr(Prescr  newp, Patient p){
   ...
}

/*@ requires
   (Consent[PATIENT][RESEARCH] ≡1 ∧Consent[TRIALDATA][RESEARCH] ≡1);
   ...
 */
TrialData getData(Patient  patientList[NB_PATIENT]){
   ...
}

/*@ requires Consent[TRIALDATA][RESEARCH] ≡1;
  ...
*/
int computeStats(TrialData  data){
   ...
}
```

Experiment Results - Plugin Use

	plugin	valid traces detected	invalid traces detected
static	WP [3]	✔	?
	Eva [4]	✔	✔
runtime	EACSL [13]	✔	✔

Conclusion

Conclusion

ML & PL consent properties: same **formal language**

Conclusion

ML & PL consent properties: same **formal language**

Contributions

- **CSpeL**: **C**onsent **S**pecification **L**anguage
- **CASTT**: **C**onsent **A**nnotation **S**pecification **T**ranslation **T**ool

Conclusion

ML & PL consent properties: same **formal language**

Contributions

- **CSpeL**: **C**onsent **S**pecification **L**anguage
- **CASTT**: **C**onsent **A**nnotation **S**pecification **T**ranslation **T**ool

Currently

- **Purpose Compliance** property
- Trace verification **ML** & **PL**
- Translation to PL Tool

Conclusion

ML & PL consent properties: same **formal language**

Contributions

- **CSpeL**: **C**onsent **S**pecification **L**anguage
- **CASTT**: **C**onsent **A**nnotation **S**pecification **T**ranslation **T**ool

Currently

- **Purpose Compliance** property
- Trace verification **ML** & **PL**
- Translation to PL Tool

Next

- More consent properties
- More complex examples
- Translation to ML tool

Conclusion

Appendix

Experiment Results - Translation

Example		nb lines in source file	nb generated lines by CASTT	nb lines in CSpeL file
Healthcare	T1	51	26	10
	T2	52	26	10
	T3	51	26	10
	T4	53	26	10
Website	T1	85	44	16
	T2	85	44	16
	T3	86	44	16

Goal & Related Work

ML & PL properties: same **formal language** $\rightarrow$ soundness

Goal & Related Work

ML & PL properties: same **formal language** $\rightarrow$ soundness

Consent-related solutions:

Goal & Related Work


ML & PL properties: same **formal language** → soundness

Consent-related solutions:

Solution	Formal properties	ML	PL	Language	Tool
[1]	✔	✔	✘	unnamed	✘
[2]	✔	✔	✘	CAPVerDE	CAPVerDE
[14]	✔	✔	✘	unnamed	DataProVe
[9]	✔	✔	✘	Prolog	Prolog

Goal & Related Work


ML & PL properties: same **formal language** → soundness

Consent-related solutions:

Solution	Formal properties	ML	PL	Language	Tool
[1]	✓	✓	✗	unnamed	✗
[2]	✓	✓	✗	CAPVerDE	CAPVerDE
[14]	✓	✓	✗	unnamed	DataProVe
[9]	✓	✓	✗	Prolog	Prolog
[10]	✗	✗	✓	unnamed	Poly
[15]	✗	✗	✓	unnamed	CASTOR
[6]	✗	✗	✓	OpenAPI	OpenAPI
[7]	✗	✗	✓	JIF	JIF

Goal & Related Work


ML & PL properties: same **formal language** → soundness

Consent-related solutions:

Solution	Formal properties	ML	PL	Language	Tool
[1]	✓	✓	✗	unnamed	✗
[2]	✓	✓	✗	CAPVerDE	CAPVerDE
[14]	✓	✓	✗	unnamed	DataProVe
[9]	✓	✓	✗	Prolog	Prolog
[10]	✗	✗	✓	unnamed	Poly
[15]	✗	✗	✓	unnamed	CASTOR
[6]	✗	✗	✓	OpenAPI	OpenAPI
[7]	✗	✗	✓	JIF	JIF
[our work]	✓	✓	✓	CSpeL & ACSL	CASTT & Frama-C

Bibliography

References I

[1] Masoud Barati et al. "GDPR Compliance Verification in Internet of Things". In: *IEEE Access* (2020).

[2] K. Bavendiek et al. "Automatically proving purpose limitation in software architectures". In: *IFIP Int. Conf. on ICT Systems Security and Privacy Protection*. Springer. 2019, pp. 345–358.

[3] A. Blanchard. *Introduction to C program proof with Frama-C and its WP plugin*. Tutorial. 2020.

[4] S. Blazy, D. Bühler, and B. Yakobowski. "Structuring abstract interpreters through state and value abstractions". In: *Int. Conf. on Verification, Model Checking, and Abstract Interpretation*. Springer. 2017.

References II

[5] Michele C. and Alberto T. "BPMN: An introduction to the standard". In: *Computer Standards & Interfaces* (2012).

[6] E. Grünewald et al. "TIRA: An OpenAPI Extension and Toolbox for GDPR Transparency in RESTful Architectures". In: *arXiv preprint arXiv:2106.06001* (2021).

[7] Katia Hayati and Martin Abadi. "Language-based enforcement of privacy policies". In: *International Workshop on Privacy Enhancing Technologies*. 2004.

[8] Jacques Cheminat. *RGPD : WhatsApp condamné à 225 millions d'euro d'amende*. URL: https://www.lemondeinformatique.fr/actualites/lire-rgpd-whatsapp-condamne-a-225-meteuro-d-amende-84044.html (visited on 09/06/2021).

References III

[9] Colombe de Montety, Thibaud Antignac, and Christophe Slim. "GDPR Modelling for Log-Based Compliance Checking". In: *Trust Management XIII*. Ed. by Weizhi Meng et al. 2019.

[10] Mohammad Nauman, Sohail Khan, and Xinwen Zhang. "Apex: extending android permission model and enforcement with user-defined runtime constraints". In: *Proceedings of the 5th ACM symposium on information, computer and communications security*. 2010.

[11] Nicolas Certes. *RGPD : Google condamné à 50 millions d'euro par la CNIL*. URL: https://www.lemondeinformatique.fr/actualites/lire-rgpd-google-condamne-a-50-meteuro-par-la-cnil-74062.html (visited on 12/10/2020).

References IV

[12] Milan Petkovic, Davide Prandi, and Nicola Zannone. "Purpose control: Did you process the data for the intended purpose?" In: *Workshop on Secure Data Management*. 2011.

[13] J. Signoles, N. Kosmatov, and K. Vorobyov. "E-ACSL, a Runtime Verification Tool for Safety and Security of C Programs. Tool Paper". In: *Int. Workshop on Competitions, Usability, Benchmarks, Evaluation, and Standardisation for Runtime Verification Tools (RV-CuBES)*. 2017.

[14] V. Ta and M. Eiza. "DataProVe: Fully Automated Conformance Verification Between Data Protection Policies and System Architectures". In: *Proceedings on Privacy Enhancing Technologies* (2022).

References V

[15] Shukun Tokas, Olaf Owe, and Toktam Ramezanifarkhani. "Language-based mechanisms for privacy-by-design". In: *IFIP International Summer School on Privacy and Identity Management*. 2019.

Credits - Images

- **Freepik**
- **Smashicons**
- **Ian June**
- **Md Tanvirul Haque**

Available on **www.flaticon.com**

Commissariat à l'énergie atomique et aux énergies alternatives
CEA Tech List
Centre de Saclay | 91191 Gif-sur-Yvette Cedex
www-list.cea.fr

Etablissement public à caractère industriel et commercial | RCS Paris B 775 685 019