

Data Protection Vocabulary

<http://w3.org/ns/dpv>

W3C Data Privacy Vocabularies and Controls CG

Harshvardhan J. Pandit

PhD Researcher

ADAPT Centre, Trinity College Dublin

me@harshp.com

Data Privacy Vocabularies and Controls Community Group

<https://www.w3.org/community/dpvcg/>

W3C Community Group made up of academia, industry, legal experts

Objective: to provide Vocabularies based on the heterogeneity or homogeneity of the agreed upon use cases and requirements, we will define a single or a modular set of vocabularies for exchanging and representing interoperably: personal data, purposes/processing, disclosure/consent, anonymisation, and transparency logs.

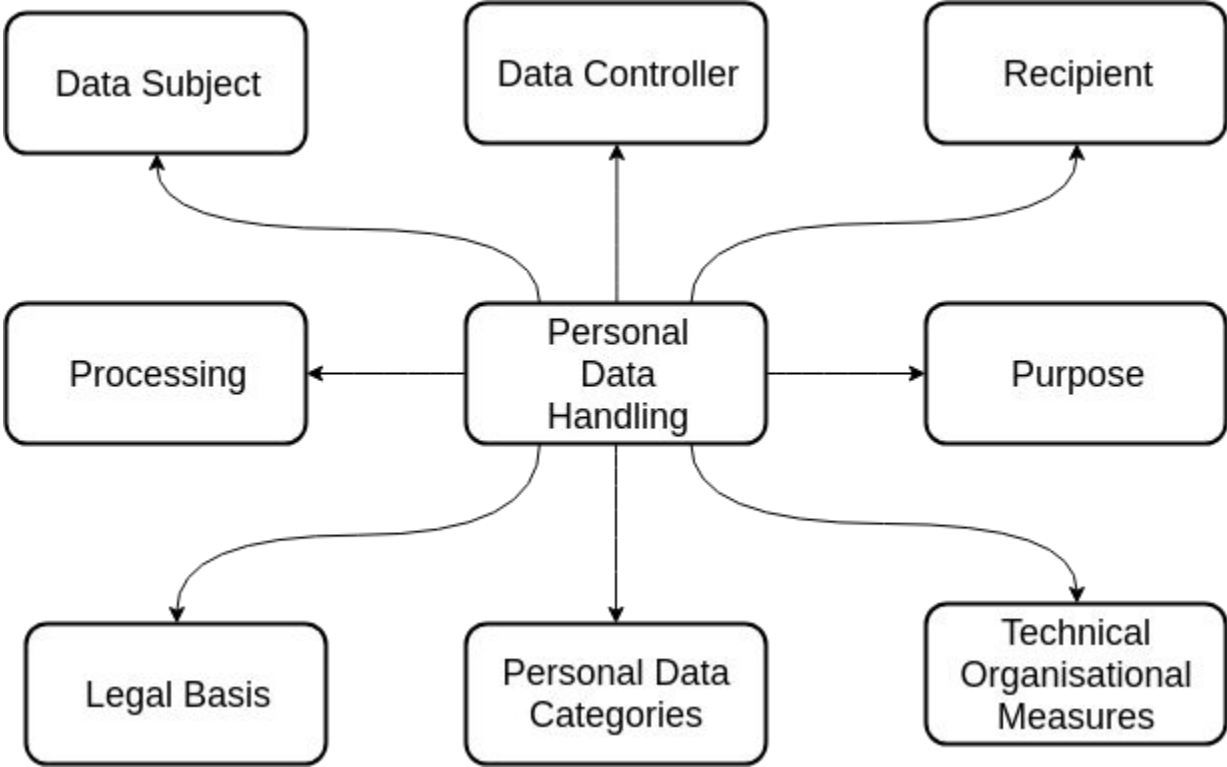
Released first draft of Data Protection Vocabulary (dpv)

Access at <http://w3.org/ns/dpv>

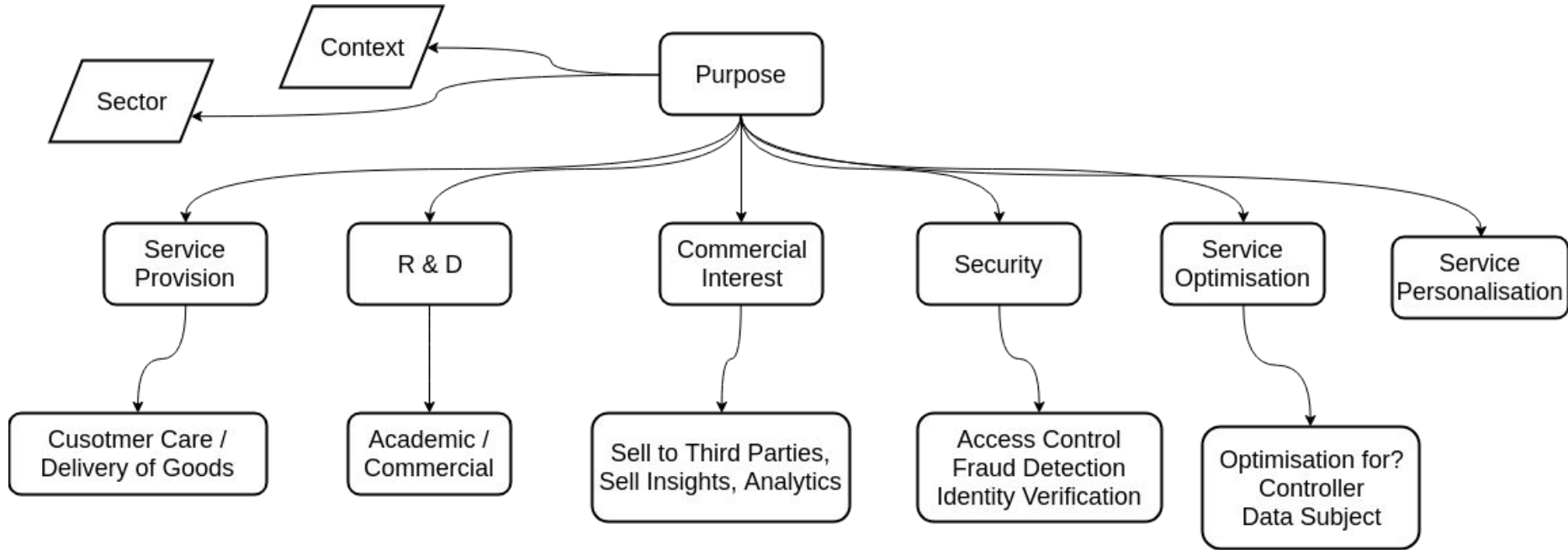
Development at <http://github.com/dpvcg/dpv>

Welcome feedback and comments !!!

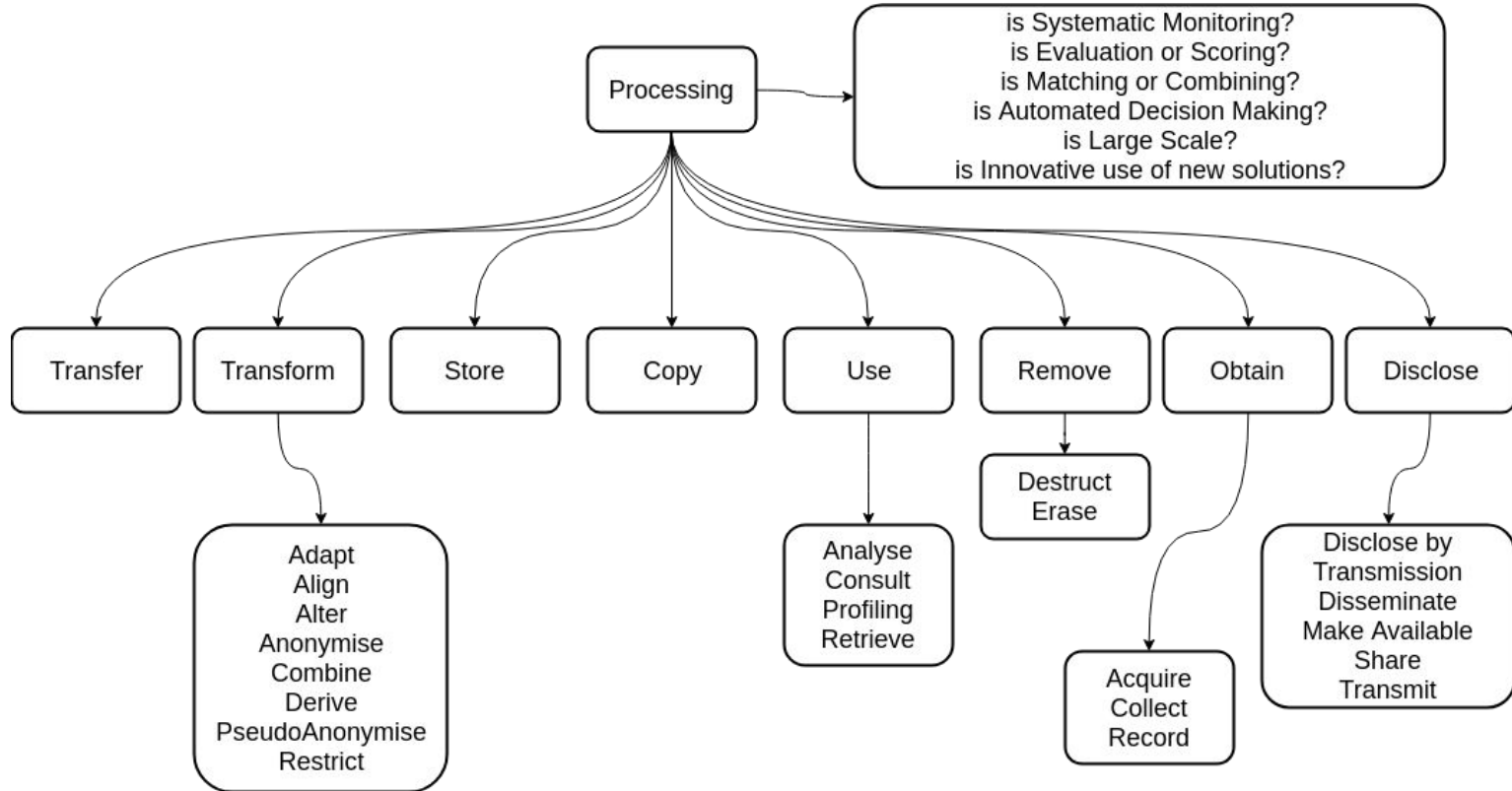
DPV - A collection of vocabularies



Purposes



Processing



Personal Data Categories

CATEGORIES OF PERSONAL INFORMATION

The following are categories of information relating to an individual, whether it relates to his or her private, professional or public life. Categories are not exclusive. Information may transcend multiple categories.



INTERNAL

- Knowledge and Belief**
Information about what a person knows or believes
religious beliefs, philosophical beliefs, thoughts, what they know and don't know, what someone thinks
- Authenticating**
Information used to authenticate an individual with something they know
passwords, PIN numbers, maiden name
- Preference**
Information about an individual's preferences or interests
opinions, intentions, interests
favorite foods, colors, items, dislikes, music

HISTORICAL

- History**
Information about an individual's personal history
events that happened in a person's life, after birth or just around them which might have influenced them (WWII, 9/11)

FINANCIAL

- Account**
Information that identifies an individual's financial account
credit card number, bank account
- Ownership**
Information about things an individual has owned, rented, borrowed, possessed
cars, houses, apartments, personal possessions
- Transactional**
Information about an individual's purchasing, spending or income
purchase orders, credit records, loan records, transactions, loans, purchases and spending habits
- Credit**
Information about an individual's reputation with regards to money
credit records, creditworthiness, credit standing, credit capacity

EXTERNAL

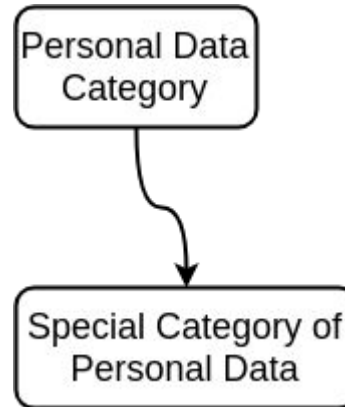
- Identifying**
Information that uniquely or semi-uniquely identifies a specific individual
name, user name, unique identifier, government issued identification, picture, biometric data
- Ethnicity**
Information that describes an individual's origins and lineage
race, national or ethnic origin, languages spoken, dialects, accents
- Sexual**
Information that describes an individual's sexual life
gender identity, preferences, proclivities, fetishes, history, etc.
- Behavioral**
Information that describes an individual's behavior or activity, on-line or off
browsing/behavioral logs, links clicked, demeanor, attitude
- Demographic**
Information that describes an individual's characteristics shared with others
age ranges, physical traits, income brackets, geographic
- Medical and Health**
Information that describes an individual's health, medical conditions or health care
physical and mental health, drug test results, disabilities, family or individual health history, health records, blood type, DNA code, prescriptions
- Physical Characteristic**
Information that describes an individual's physical characteristics
height, weight, age, hair color, skin tone, tattoos, gender, piercings

SOCIAL

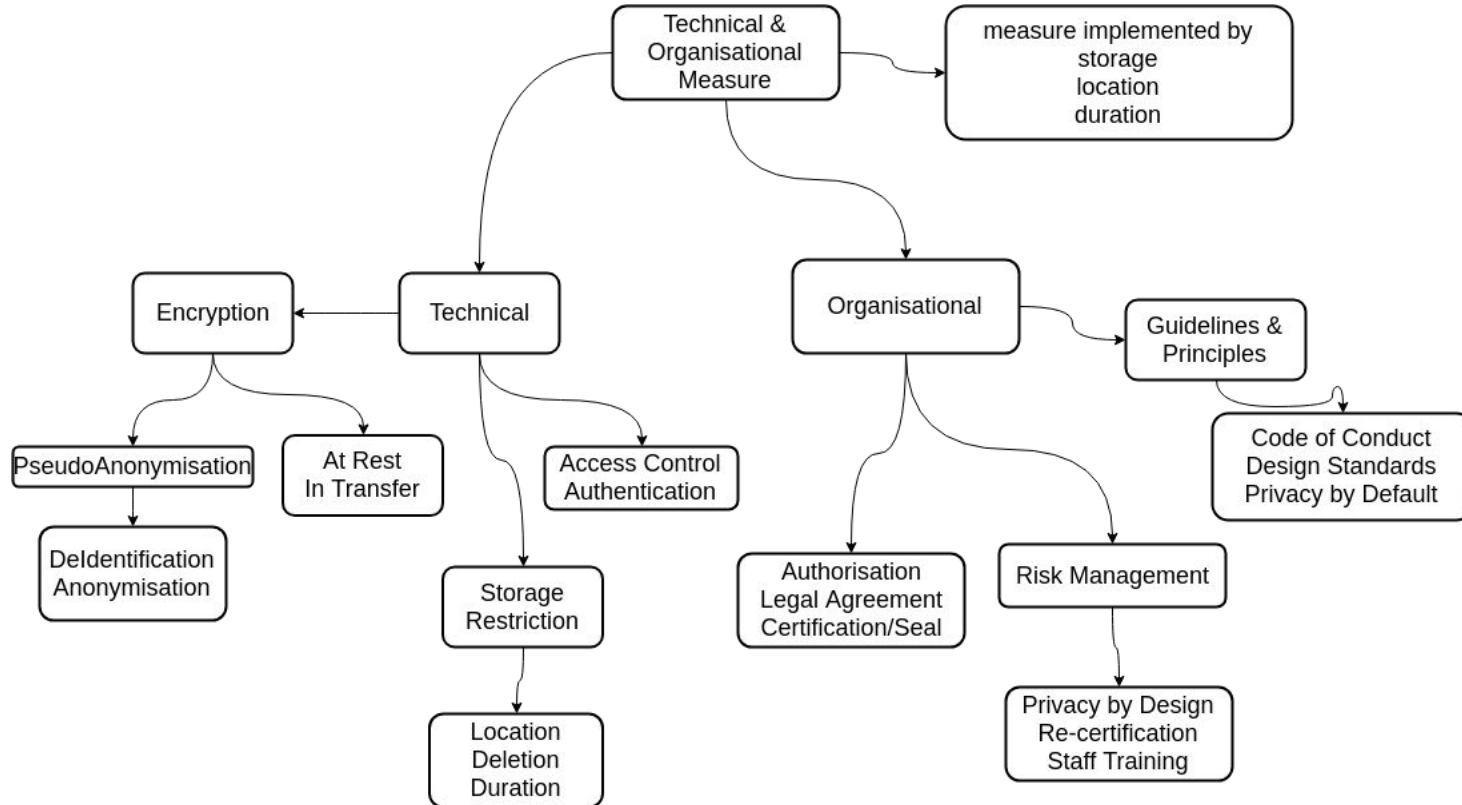
- Professional**
Information about an individual's educational or professional career
jobs/titles, salary work history, school attended, employee files, employment history, realizations, references, resumes, certifications, disciplinary actions
- Criminal**
Information about an individual's criminal activity
convictions, charges, pending
- Public Life**
Information about an individual's public life
character, general reputation, social status, marital status, religion, political affiliations, interactions, communications, media, etc.
- Family**
Information about an individual's family and relationships
family structure, siblings, offspring, marriages, divorces, relationships
- Social Network**
Information about an individual's friends or social connections
friends, connections, acquaintances, associations, group memberships
- Communication**
Information communicated from or to an individual
text messages, voicemail, email

TRACKING

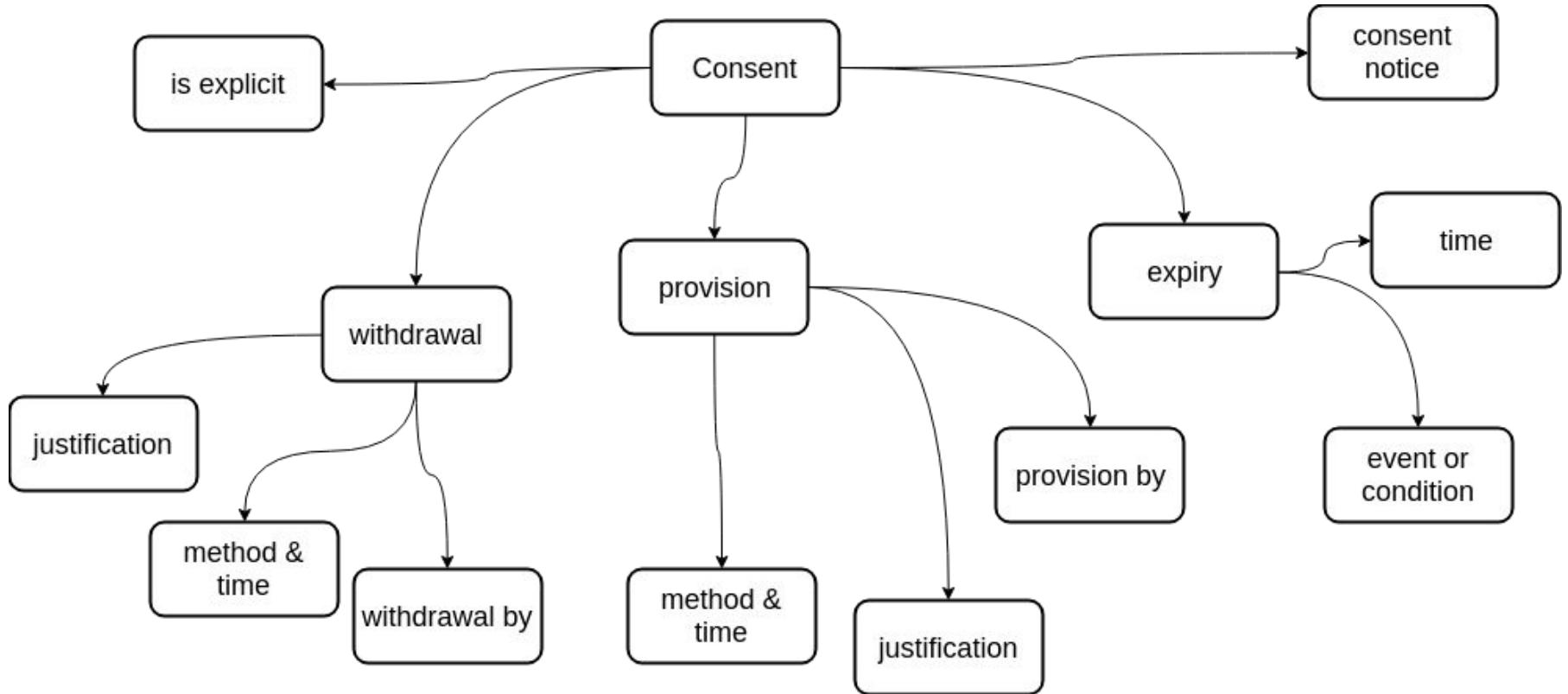
- Computer Device**
Information about a device that an individual uses for personal use (even part-time or with others)
IP address, MAC address, browser footprint
- Contact**
Information that provides a mechanism for contacting an individual
email address, phone numbers, telephone number
- Location**
Information about a individual's location
country, ZIP, coordinates, telephone number



Technical and Organisational Measure



Consent



Representation

RDF+OWL2

Semantic Web

W3C Standards

Universal Serialisation

Knowledge Graphs

Inferences + Rules

JSON-LD

Web-native JSON

Ready to integrate

Linking vocabularies

Bridge between
semantics and systems

Compatible with
Knowledge Graphs

Controlled Vocabulary

Words only (Text)

Loses semantics

Cannot be extended

Cannot be reasoned

Easy to integrate

Ongoing Work

- Examples and Use-Cases
- Expand with more categories
- Recipients, Third Parties, Roles

Contribute

- ❑ Join DPVCG
- ❑ Send suggestions/comments/feedbacks to public mailing list
- ❑ Interact on Github
- ❑ Communicate with one of the CG members

Data Protection Vocabulary

<http://w3.org/ns/dpv>

W3C Data Privacy Vocabularies and Controls CG

Harshvardhan J. Pandit

PhD Researcher

ADAPT Centre, Trinity College Dublin

me@harshp.com

HOW STANDARDS PROLIFERATE:
(SEE: A/C CHARGERS, CHARACTER ENCODINGS, INSTANT MESSAGING, ETC.)

SITUATION:
THERE ARE
14 COMPETING
STANDARDS.

14?! RIDICULOUS!
WE NEED TO DEVELOP
ONE UNIVERSAL STANDARD
THAT COVERS EVERYONE'S
USE CASES.



YEAH!

SOON:

SITUATION:
THERE ARE
15 COMPETING
STANDARDS.