Ethical Issues in Learning Analytics (Μαθησιακή Αναλυτική και Ζητήματα Ηθικής)

Τμήμα Πληροφορικής - Αριστοτέλειο Πανεπιστήμιο

Keywords:

- Learning analytics, ethics, privacy, transparency

Definitions / Background

- Data analytics
- Learning Analytics (LA)
- LA ethics

TABLE 1 A sample of LA specialized lines of research and studies

LA line	Description
Social LA (Martin, Nacu, & Pinkard, 2016)	Provides methods to study, understand, and evaluate the use of social media for learning by content and network analyses of social media texts and networks.
Smart LA (Giannakos, Sampson, & Kidziński, 2016)	Enables the analysis of valuable information gathered from heterogeneous sources and ways to deploy personalized and smart learning.
Video LA (Giannakos et al., 2016)	Transforms video streaming into useful knowledge to improve learning based on videos.
Ubiquitous LA (Mouri & Ogata, 2015; Peña-Ayala, 2015)	Analyses learner traits and contextual data to depict interactions between learners and their contexts, and learners with context based learning materials.
Visual LA (Hillaire, Rappolt-Schlichtmann, & Ducharme, 2016)	Supports pedagogical decisions by interactive visualizations that claim information design to acquire, parse, filter, mine, depict, and interact with a data collection.
Multimodal LA (Andrade, Delandshere, & Danish, 2016; Ochoa & Worsley, 2016)	Gathers multimodal information in human activity through data-capturing methods and sensing technologies.
Dispositional LA (Tempelaar, Rienties, & Nguyen, 2017)	Combines learning log data with learner data (e.g., experiences, social relations, values, and attitudes that influence the engagement with learning).
Open LA (Muslim, Amine, Mahapatra, & Schroeder, 2016)	Considers diverse actors with specific goals that demand a broad range of data from several settings to elicit knowledge and gain insight into learning processes.

Research Questions (What & Why)

- What is essential in LA ethics for key educational stakeholders?

- What are the methods for covering LA ethics?

- What should a proposed checklist for LA ethics include for educational stakeholders?
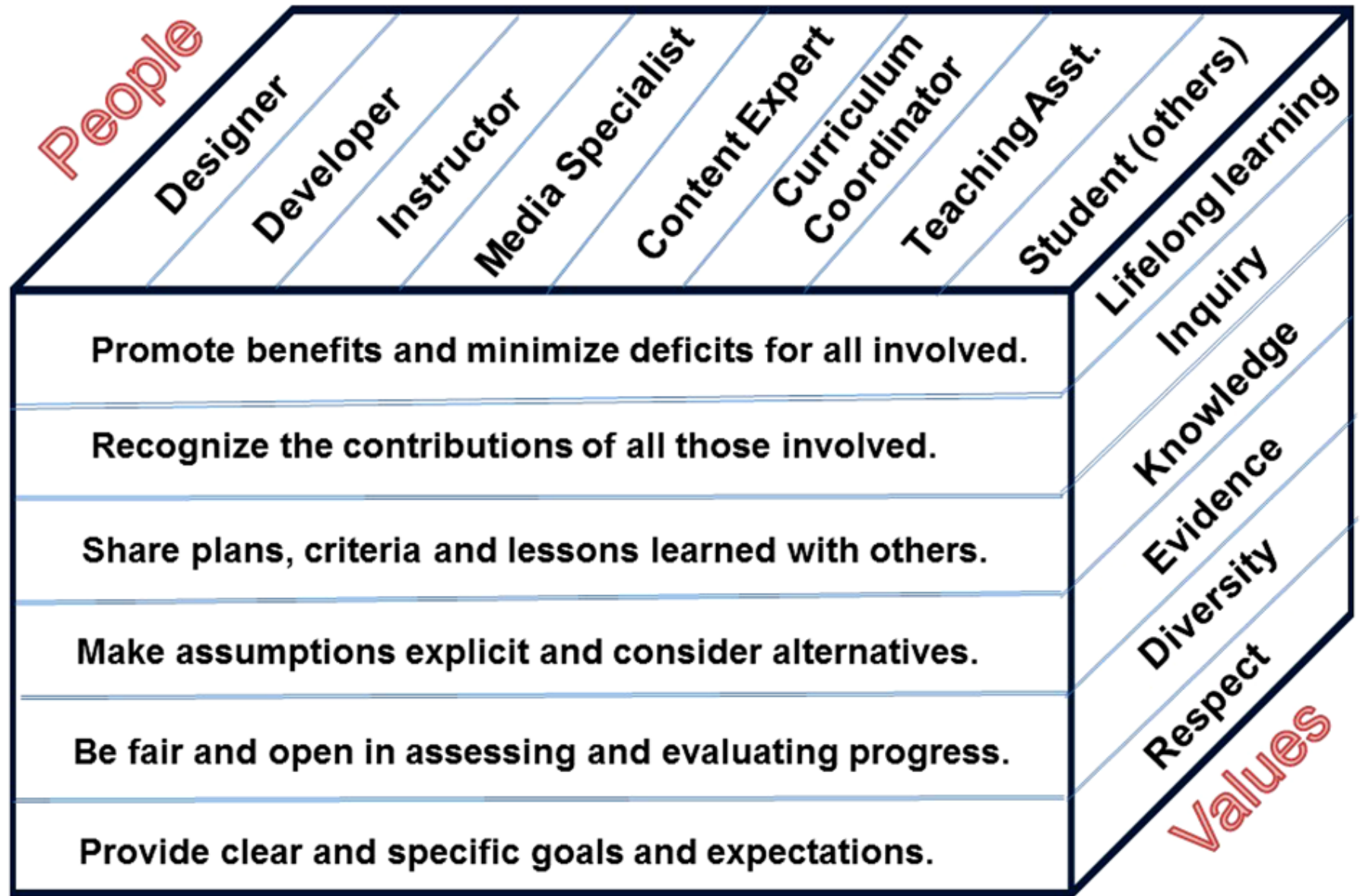
Research method

■ Review

Databases

■   IEEE Xplore Digital Library, Elsevier Digital Library through Scopus search engine, ScienceDirect, Wiley InterScience, Oxford University Press Digital Library, ACM digital library, and Springer.

Results

- Evidence-based works

1. Gursoy, Inan, Nergiz, & Saygin (2017): They conclude that increasing the level of **privacy** creates reduced **accuracy** in the LA outcomes

2. Ifenthaler, & Schumacher (2016): students are **conservative** in sharing personal data and that learners would share more data if the LA task **transparently** provided **meaningful** information.

3. West, Huijser, & Heath (2016): the lack of understanding and awareness of ethical LA issues from the key players.

A preliminary educational technology ethics framework (Spector, 2016)

People

Designer
Developer
Instructor
Media Specialist
Content Expert
Curriculum Coordinator
Teaching Asst.
Student (others)

Promote benefits and minimize deficits for all involved.

Recognize the contributions of all those involved.

Share plans, criteria and lessons learned with others.

Make assumptions explicit and consider alternatives.

Be fair and open in assessing and evaluating progress.

Provide clear and specific goals and expectations.

Lifelong learning
Inquiry
Knowledge
Evidence
Diversity
Respect

Values

Ethical Principles

Results – RQ 1 The ethical concerns with learning analytics

- **Data privacy**

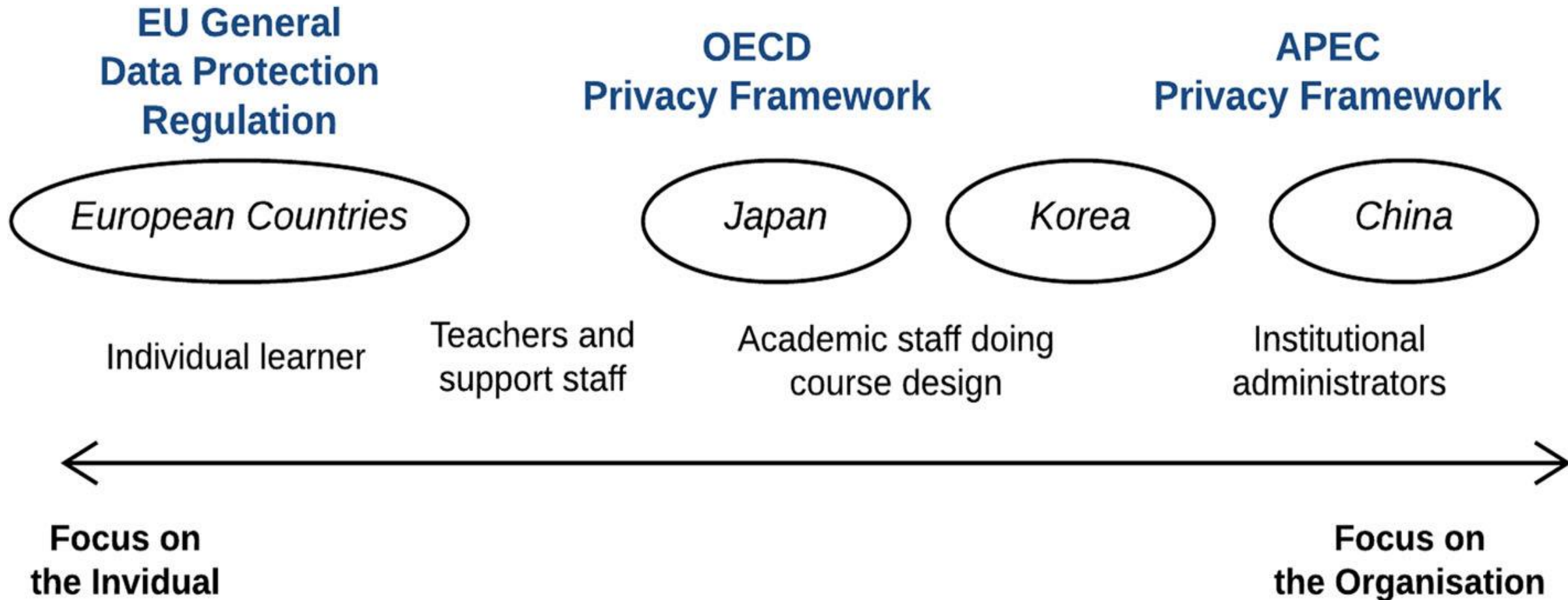- The boundaries and meaning of what is private differ among <u>cultures</u>.
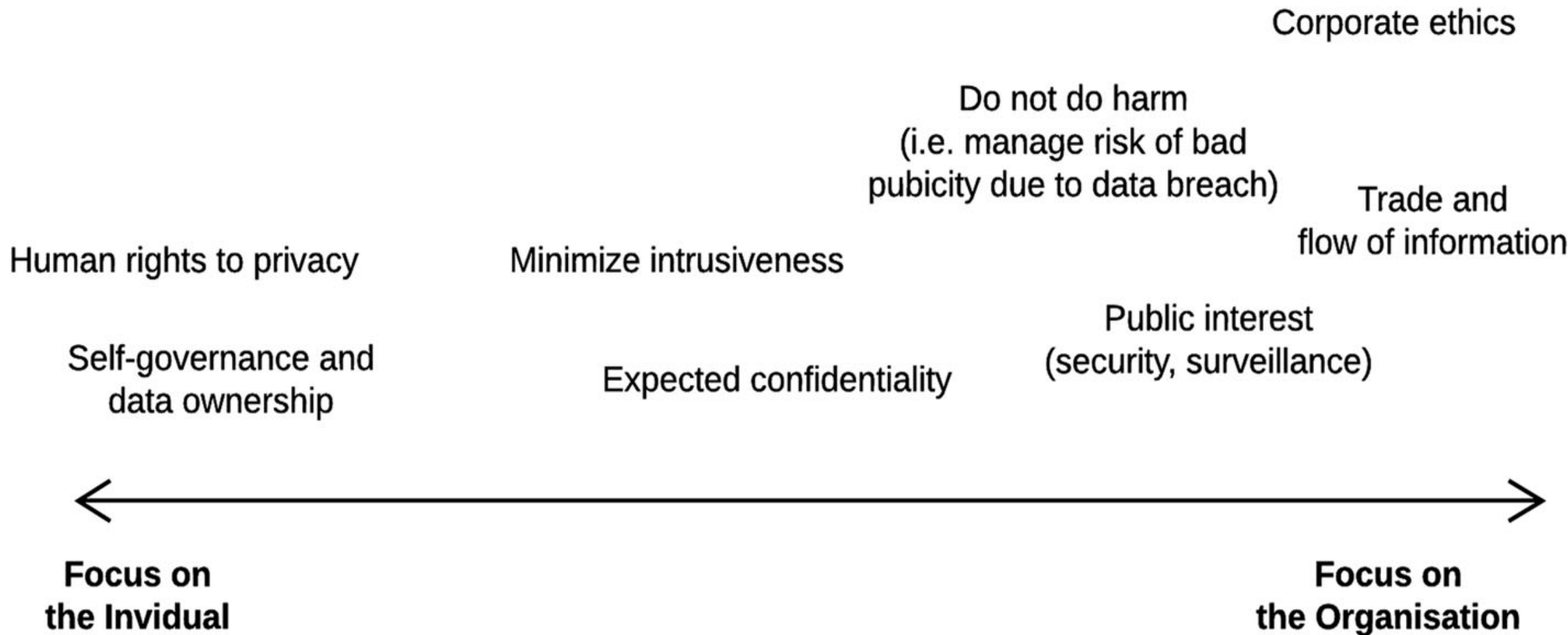- Trust

Results

- Differences between countries / philosophies

- GDPR

- Policies: DMCA, FERPA, HIPPA, FOIA

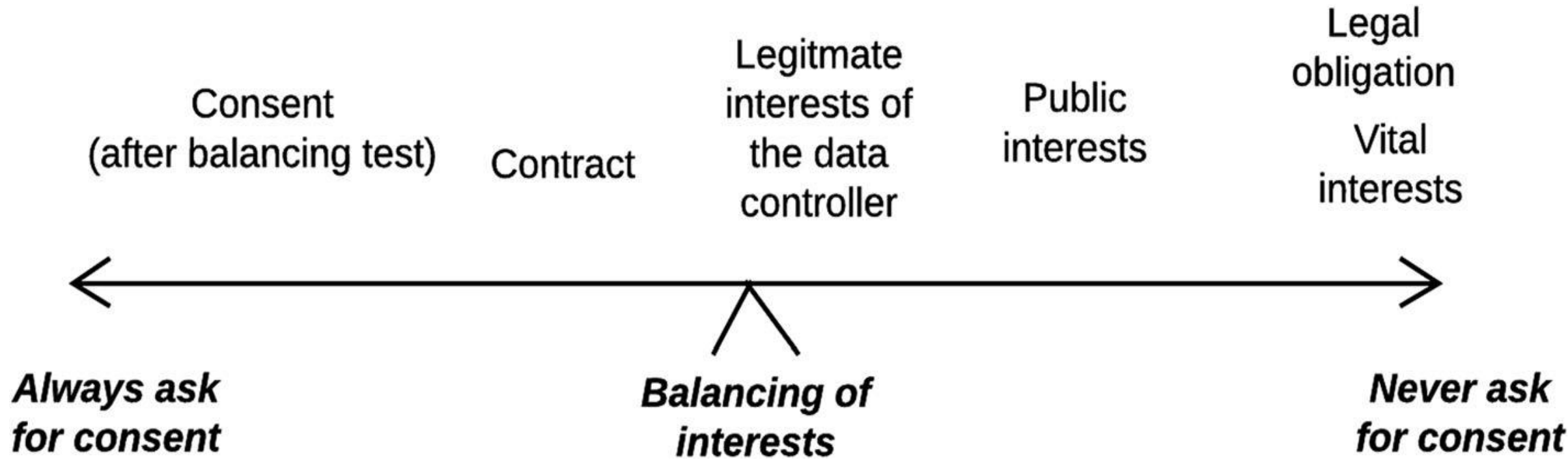- Privacy frameworks: OECD, APEC, and EU privacy frameworks.

Table 1 Status data protection laws in some Asian countries (Primary source: DLA Piper 2017)

Country	Data protection law?	Future plans
China	No	No comprehensive data protection law. However, Cybersecurity Law (2017) first national-level law that addresses cybersecurity and data privacy protection.
India	No	Draft Personal Data Protection Bill published 2018
Indonesia	No	Draft personal data protection law published 2018.
Japan	Yes (2017)	
Malaysia	Yes (2013)	
Philippines	Yes (2012)	
Singapore	Yes, only private sector (2012)	
Thailand	No	Draft is being reviewed (as of 2016).
Taiwan	Yes (2012)	
South Korea	Yes (2011)	

Hoel & Chen, 2018

EU General Data Protection Regulation   **OECD Privacy Framework**   **APEC Privacy Framework**

European Countries   Japan   Korea   China

Individual learner   Teachers and support staff   Academic staff doing course design   Institutional administrators

Focus on the Invidual   **Focus on the Organisation**

Hoel & Chen, 2018

Corporate ethics

Do not do harm
(i.e. manage risk of bad
pubicity due to data breach)

Trade and
flow of information

Human rights to privacy

Minimize intrusiveness

Public interest
(security, surveillance)

Self-governance and
data ownership

Expected confidentiality
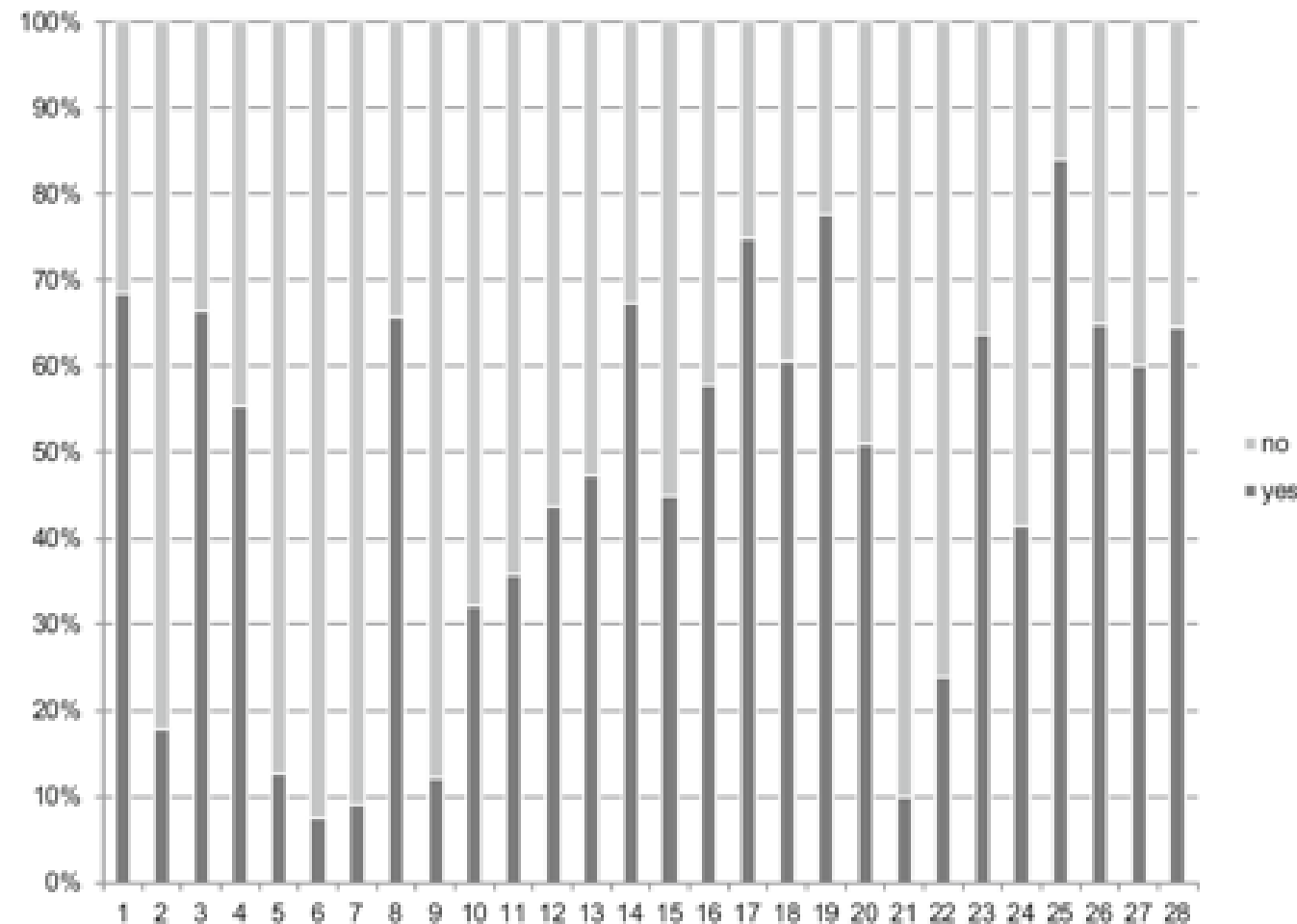
**Focus on
the Invidual**

**Focus on
the Organisation**
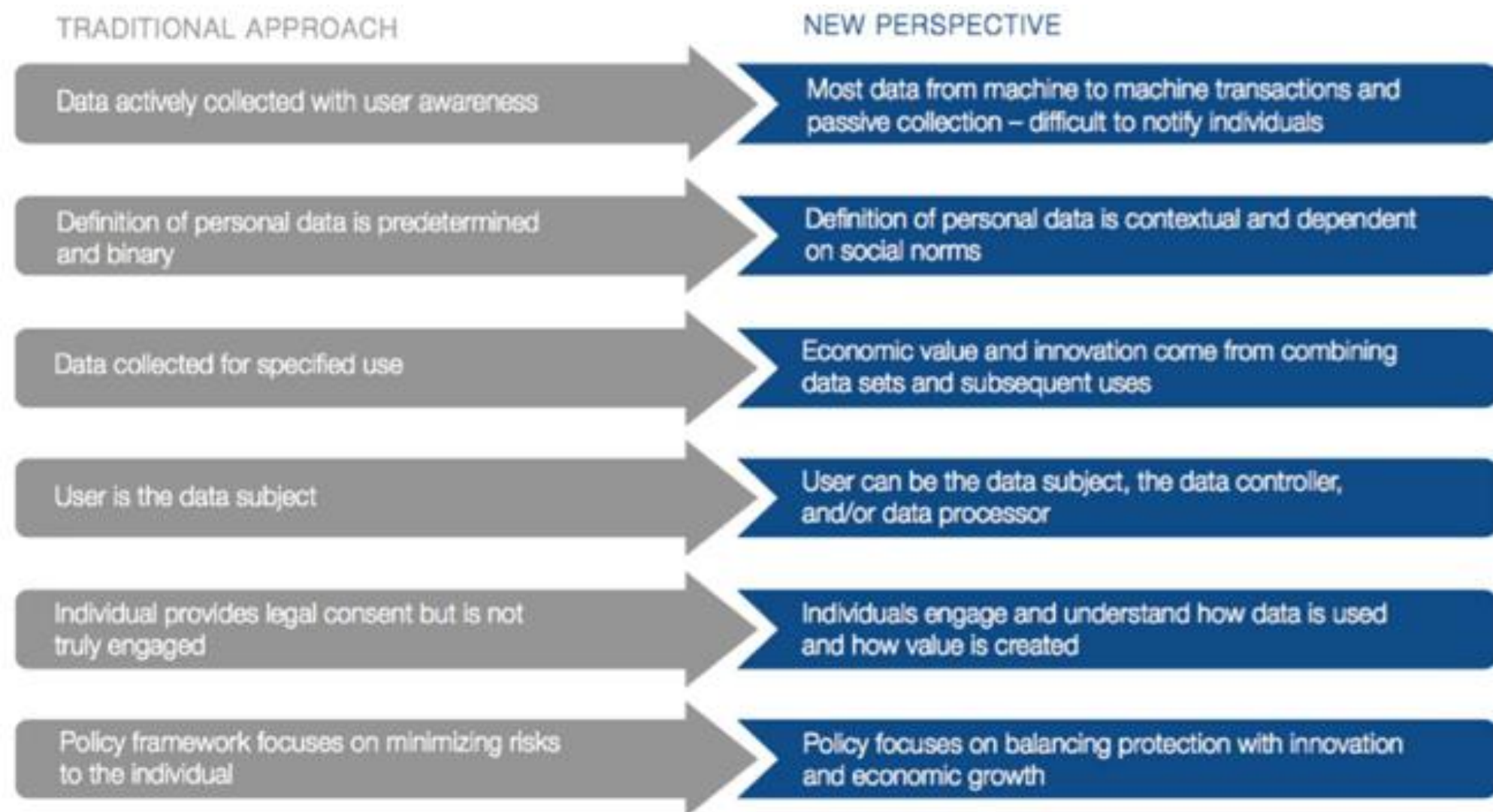
Hoel & Chen, 2018

Fig. 1 Sharing of data in learning analytics systems (1 name, 2 address, 3 email, 4 date of birth, 5 marital status, 6 medical data, 7 income, 8 prior knowledge, 9 user path, 10 online times, 11 downloads, 12 course-specific discussion activity, 13 semantic analysis of posts, 14 test scores, 15 higher education entrance qualification grade, 16 school history, 17 motivation test results, 18 interest test results, 19 learning strategies test results, 20 intelligence quotient, 21 externally produced data, e.g., social media, 22 parents' educational level, 23 academic achievements, 24 occupation other than university studies, 25 course enrollment, 26 self-test scores, 27 general discussion activity, 28 library activity statistics)

Ifenthaler & Schumacher, (2016)

TRADITIONAL APPROACH	NEW PERSPECTIVE
Data actively collected with user awareness	Most data from machine to machine transactions and passive collection – difficult to notify individuals
Definition of personal data is predetermined and binary	Definition of personal data is contextual and dependent on social norms
Data collected for specified use	Economic value and innovation come from combining data sets and subsequent uses
User is the data subject	User can be the data subject, the data controller, and/or data processor
Individual provides legal consent but is not truly engaged	Individuals engage and understand how data is used and how value is created
Policy framework focuses on minimizing risks to the individual	Policy focuses on balancing protection with innovation and economic growth

Source: World Economic Forum and The Boston Consulting Group

. (World Economic Forum 2013, p.7)

Figure 1: New perspectives on the use of data

Results - Transparency

- well-informed choice of opting in, opting out

- InBloom case
- Stichting Snappet case

Table 2 Models for handling data in educational setting

Model for data handling	Model focus	Question asked
Legal model	Justified purpose for data collection?	Are the risks to the individual balanced with the benefits to the individual and the system?
Research model	Consent, fair data handling, and safe data keeping	Have participants agreed to be part of the research?
Administrative model	Handling of personally identifiable information	Are the data de-identified and kept safe?
Pedagogical model	Learning gain	Are collected data relevant for understanding and optimising learning and the environments in which it occurs?

Results - Labeling

- Labeling, profiling, surveillance

Results – Data ownership

- Ownership refers to data collected, the analytics used, and the output of the analytics.

- Who owns the data?
- Right to be forgotten

Results – Algorithmic fairness

- Misinterpretation
- **Biases**

- **Algocracy**

# Results – The obligation to act

■ Obligation of knowing

http://bit.ly/2GyY7as

To what degree would you consider the following elements to be <u>barriers to the success of learning analytics at your institution</u>?   *(Sheila project, 2018)*

	Moderately-size to critical barrier	% (n=45)
Analytics expertise	34	0.76
A data-driven culture at the institution	30	0.67
Teaching staff/tutor buy-in	29	0.64
The affordances of current LA technology	29	0.64
Current infrastructure for data storage and management	27	0.60
<u>Legal framework</u>	27	0.60
<u>Privacy protection</u>	26	0.59
The capabilities of staff and students to understand LA results	26	0.58
Investment in research related to LA	26	0.58
<u>Ethics guidelines</u>	26	0.58
Institutional strategy	25	0.57
Student buy-in	25	0.56
Senior manager buy-in	24	0.53

Concept and relations mapping of key ethical issues
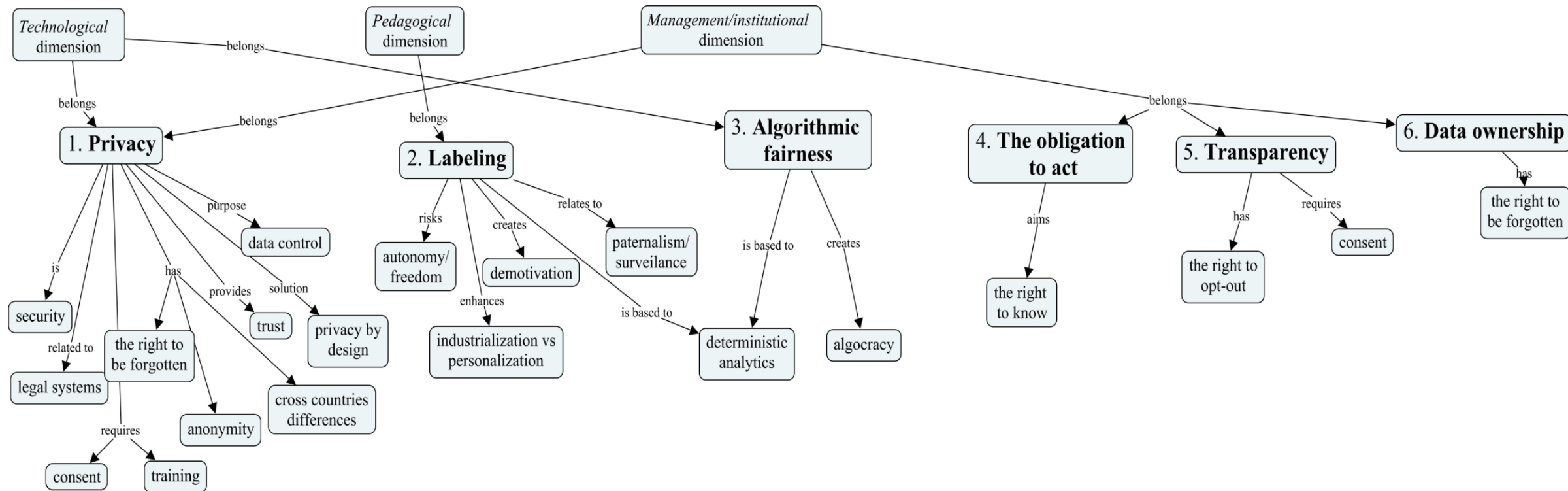
Venn diagram: The ethical issues overlap each other in the literature
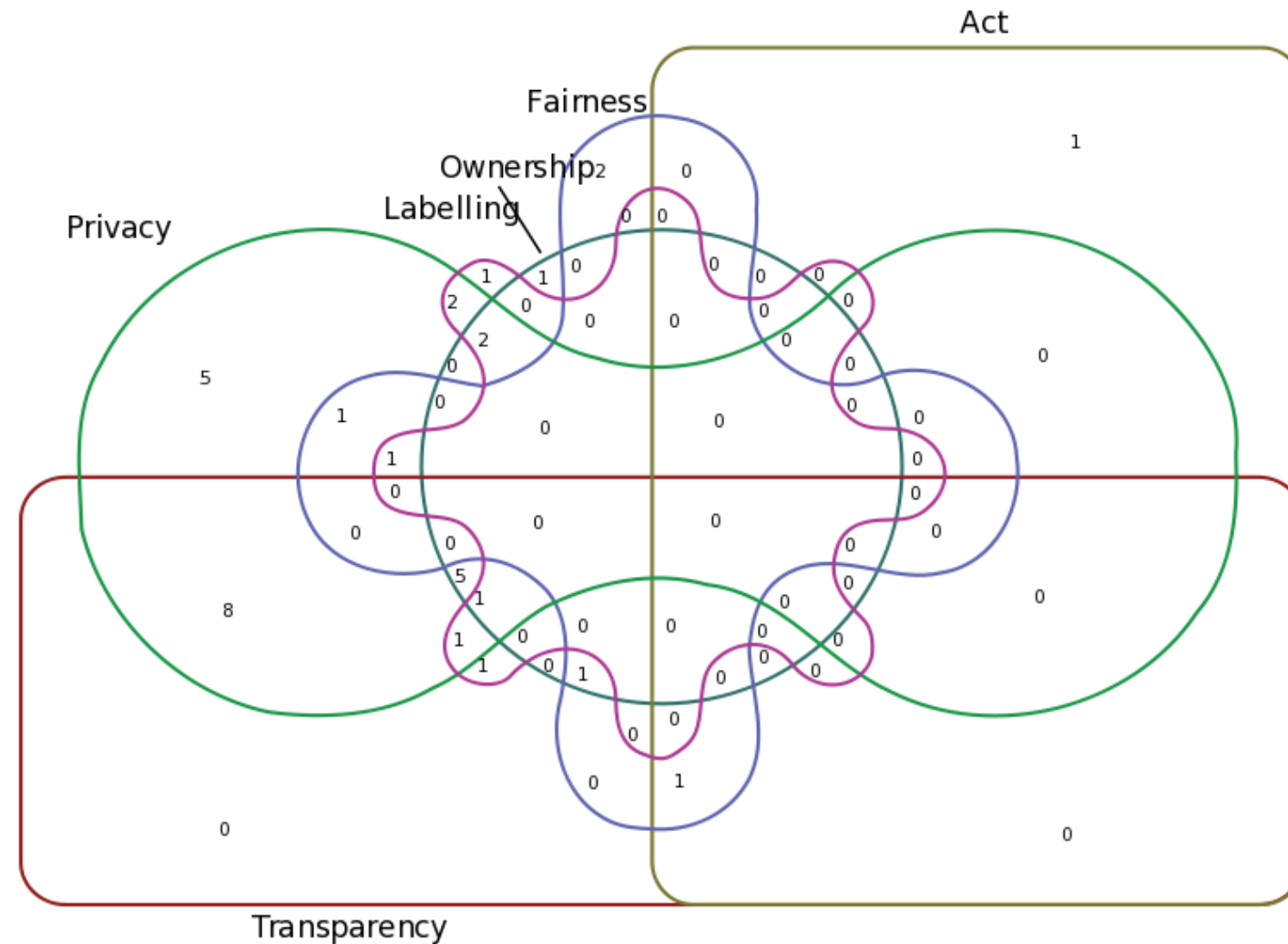
Research/open-ended questions extracted from the literature.

Paper	Questions / Key perspective
Avella et al. (2016)	What are the challenges of using LA in education?
Pardos et al. (2016)	<u>Transparency</u>: what data is being collected, how is it being represented?
Greller & Drachsler (2012)	<u>Privacy</u>: is the analysis in accordance with privacy arrangements, are the students properly informed?
Pardo & Siemens (2014)	How are privacy and ethics addressed in other contexts? <u>Who owns the data</u>: the institutions, the students, the companies using them?
Scholes (2016)	Should a decision-maker <u>sort</u> <u>students</u> on the basis of group-risk statistics?
Slade & Prinsloo (2013)	Are there some <u>labels</u> which should be prohibited? Are there circumstances in which other principles override the need for informed consent? Is it ethical to <u>ignore</u> the predictive value of research evidence?
West et al. (2016)	What ethical principles should guide the use of LA?
Willis et al. (2016)	If LA is regarded as research, then what is the response to informed consent?
Siemens & Long (2011)	If we confine analytics to behavioral data, how can we account for more than behavioral data?
Sclater (2016)	In which situations should students be asked for <u>consent</u> to the collection of their data for analytics?
Beattie et al. (2008)	There are questions in the analytics about who owns individual learners' data?
Rubel & Jones (2016)	What does it mean to have privacy?
Siemens (2013)	Who has access to analytics? Should a student be able to see what an institution sees? How long does a university keep those data?
Hoel et al. (2017)	How will the school make sure that <u>information is used for learning and not for other purposes</u>?
Arnold & Sclater (2017)	Would you be happy for data on your learning activities to be used if it kept you from dropping out?
Prinsloo & Slade (2017)	How do we then respond to the moral and legal necessity to act, when responding in appropriate and effective ways becomes impossible?
Drachsler & Greller (2016)	If there is a computational model developed from a collection of data traces in a system, can a student still <u>opt-out</u> of such a data model?

Stakeholders' ethical issues and responsibility.

Stakeholder	Issues & Outcomes (what must we do)
Learners	Performance-related stress for learners. Students' psychological and physical well-being (Reidenberg & Schaub, 2018); Profiling based on ML (Peña-Ayala, 2018); Focusing on consent means bringing the learner into the center of the discussion (Hoel & Chen, 2018); Spoon-feeding learning approach and risks of demotivation (Tsai et al., 2018); Learners being looked under a microscope. The choice of opt-out or not opt-in could affect those who choose to opt in regarding the quality of data and services provided. Rights of students to remain individuals. Learning analytics may be harmful if the predictive category is wrong. Discrimination such as bias, labeling, and profiling (Papa & Armfield, 2018); The kind of information being collected about students and anonymization of educational data. Students ask to be in a safe environment where they can make mistakes (Drachsler et al., 2015).
Teaching staff	Teachers accept classification systems as fact even though this process is subject to data entry errors, and data cleaning. Quality of data capabilities of staff to understand LA results and ethics (Papa & Armfield, 2018); Pedagogical expertise needs to be involved in making sense of data and supporting learners to take a meaningful action based on the data (Tsai et al., 2018). Confidentiality and design of interventions.
Institutional actors (designers, administrators)	There is a shortage of leadership to ensure that implementation of LA is strategically planned and a limited availability of policies for LA-specific practice. Establish a data-driven culture at the institution (Tsai et al., 2018); The institution does not give student an opportunity to correct data used in predictive model. Use informed consent to obtain permission for data. Make predictions without understanding the model. (Papa & Armfield, 2018). Who has access to the student data and who has accountability for the overall LA procedure?
Vendors (external stakeholder)	Third-party learning environments that track student behaviors present intellectual freedom issues. Digital content vendors collect and use data for a variety of reasons, including digital rights management, and consumer analytics (Jones & Salo, 2018); Vendors rely on data sharing rather than confidentiality (Reidenberg & Schaub, 2018); Lack of effective technological solutions to ensure opt-in/-out options without affecting the quality of data (Tsai et al., 2018). Libraries and vendors must work to ensure that the contracts governing the use of digital information reflect library ethics, policies, and legal obligations. We recommend that librarians work closely with policymakers to design policies in ways that consider their professional ethics (Intellectual freedom, privacy, and confidentiality of library data and analytics) (Jones &

Stakeholders' ethical issues and responsibility (Australian case study – n=181)

Fig. 1 Level of concern about selected ethical issues (responses not mutually exclusive) (West et al., 2016)

Table 2 Participant views on ethical principles that should guide the use of learning analytics (n = 112)

Ethical principles to guide use of learning analytics	Relative frequency
Autonomy	**42 % (65)**
Privacy and confidentiality	29 % (44)
Informed consent	12 % (19)
Ability to use own data	1 % (2)
Beneficence	**23 % (36)**
Non-maleficence	**6 % (10)**
Justice	**17 % (27)**
Transparency	7 % (11)
Duty of care	2 % (3)
Integrity in using learning analytics (e.g. proper use)	8 % (13)
Total responses	**153[a]**

[a] 112 participants answered this open-ended question, providing 153 discrete responses

(West et al., 2016)

Table 3 Exploration of issues around the use of social media for learning analytics

Risks of harm or potential costs	Potential benefits
• In terms of informed consent, requiring students to use a particular platform is not necessarily allowing for consent—though this depends a little on how clear this is at enrolment/commitment point	• Social media platforms—particularly those like Facebook or Twitter—which are often most convenient for many students
• Personal/professional boundaries—including safety—for example students may have reasons for not wanting to use particular social media platforms	• Students may find using social media easier and more intuitive than an LMS
• If there are no alternative options provided—does opting out mean that they cannot complete their study?	• Social media are open so access to discussions after graduation is supported
• Even if alternative options are provided does this lead to segregation or isolation from other students in the class—and thus inequitable learning and collaboration opportunities	• Social media engagement skills are a key part of many industries and work roles in contemporary context, so this may support authentic learning
• Potential erosion of trust between the institution and the student, particularly if this was not made clear at the outset	• The integration of social media with other learning analytics data (see Kitto et al. 2015 for an example), may create opportunities for institutions and students to use data about learning to improve it
• Third party providers like Facebook do not have negotiated relationships with institutions so terms of use are subject to change and may conflict with the values or wishes of students or the institution itself	• Social media use may allow easy integration of additional stakeholders—for example you could invite industry experts to contribute to discussion
• Particular cohorts of students may be more adept at using social media platforms than others	• Using social media may be financially cost-efficient for the university and students
	• The social media platform may have superior mechanisms for supporting student learning than the LMS e.g. ability to translate content into different languages

Word list sorted by weight/frequency in parentheses.

Stakeholder	Issue	Other
Learners (127)	Privacy (100)	Policy (33)
HE Institutions (88)	Obligation to act (11)	Legal (11)
Teachers (21)	Profiling (10)	IoT (10)
Researchers (7)	Transparency (8)	Moral (5)
Administrators (7)	Data ownership (8)	GDPR (5)
Librarians (2)	Surveillance (7)	
Parents (2)		

Discussion

- *PANDORA checklist* (RQ3)

Privacy: For institutions - Refers to security, data management, lawfulness, data minimisation, and control.

For learners - consent, opt-out without adverse consequences, purposeful LA.

For all stakeholders - trust and appropriate data sharing.

Autonomy: For learners and teachers - Refers to intellectual freedom, individuality, no labelling and no paternalism/surveillance.

Not probabilistic algorithms: For institutions - Quality of data and models, no interventionism, learner-oriented. Student's performance has a temporal and dynamic character.

For teachers - there is a possibility of error, so misdirected interventions should be considered.

For learners - learning analytics should not be the only source for decision-making and learning is not a snapshot.

Duty to act: For learners - obligation of knowing and doing the best.

For teachers - provide accurately, and timely interventions; encourage self-interventions for learners.

Openness & transparency: For learners - well-informed, voluntary, and complete consent; the right to withdraw; involve students.

For institutions - purpose limitation; student data must not be sold; awareness of data use and algorithms; flexibility and not complexity.

Resolve the data ownership: For learners - refers to respect and the right to be forgotten.

For institutions - responsibility, and control of data and processes.

All stakeholders

Antagonistic Viewpoints

Issue	Description
Stakeholders	
1.1 Instructors	**Ethical responsibilities vs. interventionism**
1.2 Learners	**Need support vs. skepticism**
1.3 Institutions (Academic Analytics)	**Learning analytics vs. Student perspective**
1.4 Decision-makers & data-controllers	**Data-driven algorithms: deterministic vs probabilistic**
1.5 Governance	**Different laws vs. good communication**
Benefits - Drawbacks	
2.1 Support vs. bias, privacy	**Positive vs. ineffective interventions & minimalism vs quality**
2.2 Intellectual freedom vs. surveillance	**Autonomy vs. paternalism**
2.3 Learning's innovation vs. Analytics' evaluation of what exists in data	**Educational viewpoint vs. data mining perspective**
Rights vs. Obligations	
3.1 Right to	**Be forgotten, know, restrict processing, opt-out**
3.2 Obligation to	**Act, do the best**
Technology vs. Regulations	**Dynamic vs. static**
Ethics vs. Law	**Moral conventions vs. Legal Norms**
Student-oriented vs. intervention oriented	**Active agents vs. Passive recipients**

MOOCs

http://bit.ly/2I8okPD

http://bit.ly/2IaHSmy

http://bit.ly/2IcMw3c

Ερωτηματολόγιο

http://bit.ly/2vK1W5T

Sources

http://bit.ly/2QqYWno

Σας ευχαριστώ για την προσοχή σας!