Legal Knowledge and Information Systems
E. Schweighofer (Ed.)

doi:10.3233/FAIA210332

A GDPR International Transfer Compliance Framework Based on an Extended Data Privacy Vocabulary (DPV)

David HICKEY[ab], Rob BRENNAN[a]
[a]*ADAPT Centre, School of Computing, Dublin City University, Dublin, Ireland*
[b]*Datanet International, Dublin, Ireland*

Abstract. This paper describes a tool using an extended Data Privacy Vocabulary (the DPV) to audit and monitor GDPR compliance of international transfers of personal data. New terms were identified which have been proposed as extensions to the DPV W3C Working Group. A prototype software tool was built based on the model plus a set of validation rules, and synthetic use-cases created to test the capabilities of the model and tool (together a compliance framework). This framework was created because the rules around international transfer compliance are complex and changing, there is an absence of a common approach to ensuring compliance, few tools exist to assist, and those that do lack interoperability. Evaluation results demonstrate that the proposed model improves compliance identification and standardisation. The tool received positive feedback from the data protection practitioners who participated in the evaluation, and an initial version of is now in use in one financial services organisation. While currently the tool only addresses international transfers, in theory the framework can be extended through further work to the broader area of compliance of other aspects of the GPDR.

Keywords. GDPR, International Transfer, Compliance, DPV, Data Protection, Privacy.

1. Introduction

The General Data Protection Regulation (GDPR) requires organisations to demonstrate compliance with data protection law. Whether needed to create standardised and reliable internal processes, or to respond to an external regulatory audit, there are few models or tools available to address this compliance requirement. Those that are available tend to focus on recording overall compliance by asking the user to confirm they are compliant with specific GDPR Articles, through guided questions, and then logging this. The shortcomings of this approach are that it generally requires legally trained staff to answer the questions effectively, there is little or no support for automated checking or validation of the answers as little data is collected on the transfer process itself, and form filling may become habitual (checking 'yes', 'yes', 'yes', repeatedly) leading to a degradation of accuracy.

In terms of available tools, some regulators such as the Irish DPC [1] have introduced checklists, others have provided templates such as the Accountability Framework [2] from the ICO in the UK, and some like the French CNIL [3] have developed open-source compliance software. And there are also commercially available compliance tools such as

OneTrust[1], PrivIQ[2] and others. These tools all seem to be limited in a number of ways: (a) rather than assisting the user in determining compliance, they focus on recording the user's view or declaration (b) they require an expert knowledge of data protection law (c) the tools are standalone and not interoperable and (d) they are not well suited for ongoing management of compliance.

For compliance of international transfers of personal data, the tools are even more limited. The tools referenced here typically record the fact of an international transfer, the legal basis stated by the user as evidence of compliance as single entries, but little else. Compliance determination for transfers of personal data is complex. There is a need for a transfer legal basis, country-specific transfer safeguards, requirements for Transfer Impact Assessments, an evaluation of whether supplementary measures need to be put in place and if so, a determination of which ones to implement. Added to this is the changing landscape where additional countries such as UK [4] can be granted GDPR Art. 45 Adequacy, and in other cases legal bases for transfers such as EU-US Safe Harbor [5] have been invalidated by the Court of Justice of the European Union (CJEU).

This paper focuses on International Transfers of personal data to third countries, and seeks to provide a framework (model and tool) to assist in ensuring such transfers are compliant with EU law. The main research question this paper asks is: *to what extent can a model based on DPV ensure regulatory-compliant international transfers of personal data ?* Related to this two additional sub-questions are asked: (i) how can we currently identify cross-border transfers and privacy issues in existing business processes ? and (ii) what are the requirements in developing a model and related tool to audit, report on, and monitor transfers ?

The approach taken in this paper to solving this problem is to:

- gather requirements for international transfers from domain experts
- identify extensions to DPV needed to model transfers
- develop a set of validation rules
- develop a prototype using Flowfinity Actions, the extended DPV and rules
- evaluate the model and this tool using synthetic use cases and external test users

The contributions of this paper are:

- identifying gaps, proposing 5 new DPV concepts and 6 properties for transfers
- creating a software tool based on the model (extended DPV and rules)
- developing a new framework for transfers based on the model and tool
- evaluating the tool and model against sample use cases

The rest of this paper is structured as follows: section 2 describes related work in this area; section 3 explores the DPV as a basis for the model; section 4 provides a model design leading to a tool design in section 5; section 6 and 7 describe the testing of the framework and results respectively; section 8 provides an analysis and discussion; and section 9 draws conclusions from the work.

[1] https://www.onetrust.com

[2] https://priviq.com/solutions-gdpr-software/

2. Related Work

Reasons for examining this area are three-fold: (a) limited work done to date on modelling compliance of transborder data transfers; (b) such transfers have been a focus of European regulators in the recent past – particularly since the introduction of the GDPR; (c) concerns about a need to evaluate the adequacy of data protection in the country in which an organisation processes personal data.

GDPR Art. 5(2) introduced new accountability obligations on Controllers and Processors to demonstrate compliance, with some specific examples being Art. 30 Record of Processing Activities (ROPA) and Art. 35 Data Protection Impact Assessment (DPIA). Since then, organisations have approached compliance in non-standard ways, typically using paper records, spreadsheets or a range of commercial tools. Even in 2018, there was a recognition of the gap in this approach. Ujcich et al. [6] comment on how increasingly complex personal data workflows create challenges for GDPR compliance. Research since has further highlighted the need for a standardised model for compliance. In [7], a year after the introduction of the GDPR, Torre et al. comment on the lack of an automated solution and the need for manual audits.

In 2020, Ryan et al. [8] reviewed tools in use for GDPR compliance and determined that there are gaps – particularly in standardization and interoperability. They suggest a RegTech approach to explore a prototype for GDPR compliance. At the same time, Pandit et al. [9] recommended the creation and adoption of standards and a common language for the exchange of GDPR compliance data. This theme aligns with the creation and development of the Data Privacy Vocabulary [10] or DPV. Early work by Pandit et al. [9,11] on interoperability and consent gives us a rich set of concepts and an extended (common) vocabulary relating to personal data processing. Their research references related work by the W3C Community Group for Data Protection Vocabularies and Controls (DPVCG) who have published the Data Privacy Vocabulary (DPV) specification. Ryan et al. [12] look at how the DPV might be used to model compliance with the GDPR requirement for Records of Processing Activities (ROPA), and proposes extensions to the vocabulary. Research on OWL2/DPV in the context of privacy policy language by Bonati et al. [13] shows how it can encode consent, business processes, and regulatory obligations. It also highlights the need for automated compliance checking. In [14], Leoni and Di Caro use the DPV to look at natural language processing of privacy policies.

The research work to date has identified a gap and a need for an automated compliance tool. GDPR requirements and new European Data Protection Board (EDPB) guidance [15] and [16] are creating an even stronger demand for a suitable standardised compliance framework.

3. Extending the DPV for International Transfers

To address the research question, it was necessary to understand (a) what the scope of such a model would need to be and (b) if the DPV could address the necessary scope. A baseline survey of domain experts was carried out in March 2021. This survey posed 20 questions, covering two areas: (a) how the organisation identifies and monitors compliance and (b) the nature of those transfers (in terms of the data subjects, personal data transferred, purposes and legal basis). Questions were deliberately open, to identify as many free-text vocabulary terms as possible in relation to transfers.

The survey was sent to students in the DCU Masters Programme in Data Protection and Privacy Law (with expert knowledge). 13 responses were received from respondents in various industry sectors helping to gather a representative view. Participants in this survey stated that spreadsheets are the most common way of tracking compliance, there are few other tools available. They also stated that identification of transfers typically take place through consultation with the business, or discovered from the outputs of impact assessments, ROPAs or breaches. Analysis of the responses (see Table 1) showed a great variation in the terminology used, without an agreed and defined vocabulary.

Table 1. Variation in Responses to Survey 1

Category	Different answers	Common answers
Legal basis for transfer	8	11
Data subject types	22	13
Purposes	21	2
Overall Common Responses		33%

A dictionary of terms commonly used when describing international data transfers was then built from the survey results combined with a review of transfer-related terms in the GDPR, to determine whether the DPV could usefully provide a potential basis for standardising compliance of international transfers. A total of 114 relevant concepts were listed, and compared to the DPV.

The results were promising, as 44% of the terms needed were fully provided by the DPV, and a further 40% partially matched. In building the initial model it was determined that these partial matches were usable. Of the remaining 18 terms not in the DPV, 4 related to Purposes/Measures and were not critical, and 4 were safeguards that had been replaced, leaving 10 which were proposed to the W3C-DPVCG in Aug 2021. In October 2021, 7 of these terms were adopted into the DPV[3]. These terms are shown in **Table 2**.

Table 2. Transfer compliance rule requirements requiring additions to DPV

Validation Rule Requirement	DPV Concept Proposed	Adopted in DPV v0.3
Transfer Start and Date	**dpv:hasStartDate, dpv:hasEndDate**	(alternative = dpv:Duration)
Data Exporter	**dpv:DataExporter**	**YES**
Data Importer	**dpv:DataImporter**	**YES**
Transfer to Country	**dpv:TransferCountry**	(Under discussion)
Legal basis for transfer	**dpv:DataTransferLegalBasis** **dpv:DataTransferTool,**	**YES** **YES**
Safeguard in use	**dpv:Safeguard** **dpv:SafeguardForDataTransfer**	**YES** **YES**
Supplementary Measures	**dpv:SupplementaryMeasure.**	**YES**

[3] DPV and DPV-GDPR v0.3 release can be found at https://w3.org/ns/dpv and https://w3.org/ns/dpv-gdpr

4. Framework Design

This now facilitated the initial design of the model, which seeks to identify international transfers, their legal basis, the levels of safeguards in use, and to show whether the transfer is compliant. For consistency and future interoperability, the model was designed to draw from the terms and concepts in the (extended) DPV. The final stage of design was to map the information capture requirements, the workflow and the associated DPV concepts and properties as shown in Figure 1.

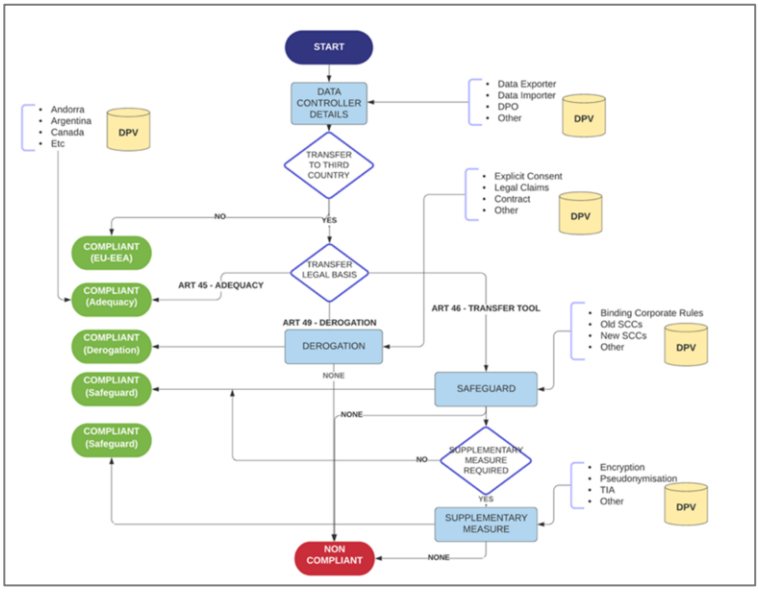

Figure 1. International transfer framework design and workflow

5. Prototype Design

A prototype software tool was designed with a user interface, guided workflow, a database (DPV extract), validation rules and an output (compliance report). The tool (Figure 2) takes user input, with intelligent workflow and conditional logic and is built on a schema based on a snapshot of the extended DPV. Natively using the DPV Linked Data rather than real-time queries (for no real gain) avoids possible versioning problems.

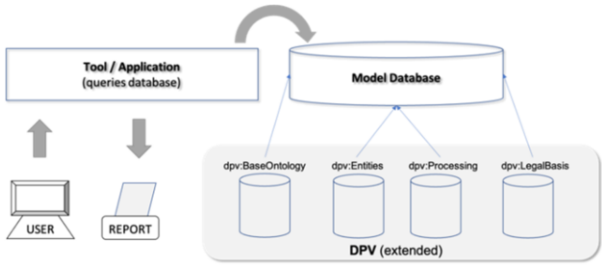

Figure 2. Transfer compliance prototype tool design

In building the tool, the primary purpose was to allow the framework to be evaluated. Rather than coding and development of a feature-rich tool, the focus became delivery of an MVP (Minimum Viable Product) to put a usable tool in the hands of test users (external participants) quickly[4]. The application development platform chosen to develop the prototype was Flowfinity Actions[5]. This provided a web-based portal, security infrastructure, workflow, a customisable user-interface, and integration capability.

The model's data capture requirements, logical rules and DPV datasets were then developed in the application. To assist in building the tool, approx. 200 test cases for international transfers were created based on the authors' own experience. The prototype Transfer Compliance tool was built and deployed in about six weeks, rather than potentially many months of development.

Two types of user or personas were envisaged while building the tool: (i) business user (without any particular domain knowledge); (ii) practitioner (someone with data protection knowledge). The tool therefore allows for separate (or common) data entry and review. In Figure 3, the user is presented with workflow driven conditional questions. At the back-end, relevant concepts from the DPV act as an internal data model schema for the tool. These DPV terms are then presented to the end user.

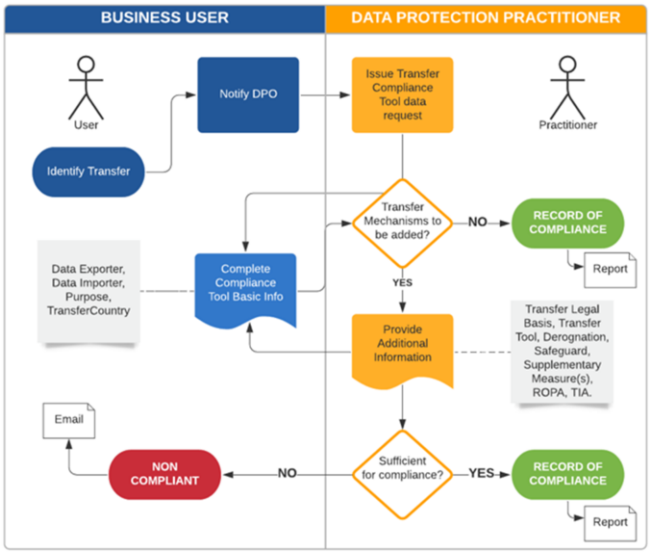

Figure 3. Tool user personas

The Transfer Compliance Tool collects a minimum of 34 data fields, of which only 3 are free-text input. The remaining fields are conditional or calculated which all combine to provide significant automation in the tool. Selection of a transfer country (see Figure 4) auto-populates compliance fields thus ensuring a high degree of consistency, and making inter-operability with other (DPV-based) tools easier.

[4] An initial version of the tool has now been deployed in a financial services organisation in Ireland.

[5] https://www.flowfinity.com/apps/

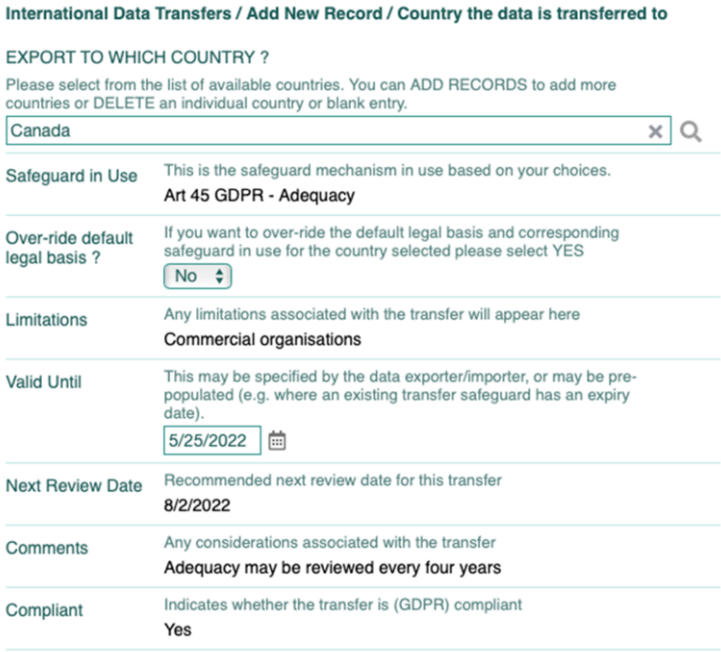

Figure 4. User interface from the prototype tool

When data entry is complete, compliance is automatically determined and a report is available within the tool (typically for the Practitioner) and an email report with details of compliance or deficiencies sent (typically to the Business user).

6. Evaluation

The hypotheses under test were whether a framework can be developed based on an extended DPV to model international transfers and ensure compliance; whether using a prototype tool based on this model will result in greater consistency in information gathered, conditions for compliance, and more informed decisions being made; and whether such a tool would achieve greater usability than current compliance methods.

A 'Gold Standard' of expected answers was created and externally validated by a domain expert. To measure accuracy, responses with expected answers were assigned a score of 1 and unexpected answers a score of 0. Tasks completed were evaluated against the Gold Standard. Consistency was measured as a percentage of expected out of 59.

Experimentation revolved around two synthetic use cases requiring the most common types of safeguards, which participants had to analyse for transfer compliance. The first (Green Bank) was a relatively simple, transfer of data to just one country (UK). The second (Childcare International) was more complex involving transfers to multiple countries (UK, Luxembourg, Singapore, USA) and some non-compliance.

A total of 13 participants was recruited from Ireland, UK and Germany based on personal contacts. Participants were classified as: Practitioners (10) with at least 3 years' experience or Business Users (3) with little or no previous data protection experience.

Participants used the Transfer Compliance Tool to complete the two use cases. Following this, a SUS questionnaire with additional open questions was completed by participants to gather feedback on the tool and model. In the evaluation, participants were expected to complete 59 fields in the Transfer Compliance tool.

7. Results

Thirteen sets of results were collected, each set covering the five international transfers. For each participant, 59 fields were evaluated against the Gold Standard answers for consistency. The results gave a high degree of correlation between actual and expected.

Table 3 Percentage of correct tasks vs Gold Standard

USE CASE	Business User Correct Tasks	Practitioner Correct Tasks	OVERALL Correct Tasks
Green Bank	93%	96%	95%
Childcare International	87%	92%	91%
OVERALL	89%	93%	**92%**

As seen in Table 3, the first use case gave an average of 95% consistency when compared to expected results. The second use case gave a slightly lower average of 91% consistency against expected results. And business users performed almost as well as practitioners. This is perhaps an indication that the tool helps more informed decision making by non-experts. Greatest discrepancies between actual and expected results were seen (Table 4) in open questions or when there were a large number of choices.

Table 4. Most common failed tasks vs Gold Standard

		Respondents	Failed %
1	Failure to correctly identify all processing purposes (C)	13	100%
2	Failure to correctly identify supplementary measures (C)	12	92%
3	Lack of standard naming for processing activity (O)	8	62%

(O) = open question (C) = over 40 choices

To evaluate the usability of the tool, a survey questionnaire based on the System Usability Scale (SUS) was used, with ten standard questions and an additional four optional open questions to gather less structured feedback about the potential value of the tool. All participants completed the SUS survey after using the tool for the two test cases. The resulting mean **SUS score was 82**.

One quote from a respondent that is representative of the responses received to the open questions reads: *"I have not found any specific tools that deal with international transfers. I think this tool could really help me with compliance."*

8. Discussion

Results show that a DPV-based framework can improve identification and compliance of international transfers. There was a clear improvement in standardisation of compliance responses (33% at the outset of this work to 92% using the framework). This resulted (i) from a linkage between the model and DPV terms (ii) from constraining data input in the tool to those terms available in the DPV and (iii) automation.

Modelling based on the DPV is not without its challenges, in particular as the DPV is still evolving. The failures shown in Table 4 mainly arise as the terms are not an exact match. Improving on the residual error rate may be possible with further development of the DPV and extensions to its concepts. Relating the results to the research question:

- The tool provides automated measurement of compliance of transfers
- Business and expert users can use the tool to identify and improve compliance
- Using the framework (DPV-based model and prototype tool) led to over 90% adherence to a gold standard for compliance validated by a domain expert.
- The prototype achieved a mean SUS score of 82, which is an 'A' grade score and where respondents are "*likely to recommend the product to a friend*" [17]

The DPV is a suitable vocabulary for such modelling, but extensions are needed. In developing the model, new concepts were identified, proposed as additions, and most adopted into the DPV. Compliance rules and guidance around international transfers change regularly. The model addresses this by importing the latest information from the DPV, and the tool allows for this to be extended to dynamic queries in the future.

9. Conclusions

By achieving a high consistency of standardised results, the framework developed has helped to answer the research question and demonstrate that the DPV, once extended and complemented with validation rules, can be used as a basis for ensuring compliance of international data transfers. The prototype tool received positive feedback on identification of transfers and helping improve compliance and accountability.

A limitation of the research, particularly for the quantitative results, was the limited dataset of respondents and use cases. Further research might revisit this and further extensions to the DPV to fully describe international transfers, while also looking at integration with other DPV tools currently being developed e.g. CSM-ROPA [12].

Acknowledgments

A word of thanks to Dr Julio Hernandez-Torres and Paul Ryan for their encouragement and sharing their own experiences.

This research has received funding from the ADAPT Centre for Digital Content Technology, funded under the SFI Research Centres Programme (Grant 13/RC/2106_P2), co-funded by the European Regional Development Fund. For the purpose of Open Access, the author has applied a CC BY public copyright licence to any Author Accepted Manuscript version arising from this submission.

References

[1] Ireland DPC. GDPR Readiness Checklist Tools [Internet]. Self-Assess. Checkl. 2021 [cited 2021 Feb 2]. Available from: https://www.dataprotection.ie/en/organisations/resources-organisations/self-assessment-checklist.

[2] ICO UK. Introduction to the Accountability Framework [Internet]. ICO; 2020 [cited 2020 Nov 15]. Available from: https://ico.org.uk/for-organisations/accountability-framework/introduction-to-the-accountability-framework/.

[3] France C. The open source PIA software helps to carry out data protection impact assesment [Internet]. PIA Softw. 2021 [cited 2021 Jan 15]. Available from: https://www.cnil.fr/en/open-source-pia-software-helps-carry-out-data-protection-impact-assesment.

[4] Commission adopts adequacy decisions for the UK [Internet]. Eur. Comm. - Eur. Comm. [cited 2021 Aug 14]. Available from: https://ec.europa.eu/commission/presscorner/detail/en/ip_21_3183.

[5] Court of Justice declares that the Commission's US Safe Harbour Decision is invalid [Internet]. Off. J. 215 25082000 P 0007 - 0047. OPOCE; [cited 2021 Aug 14]. Available from: https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:32000D0520:EN:HTML.

[6] Ujcich BE, Bates A, Sanders WH. A Provenance Model for the European Union General Data Protection Regulation. In: Belhajjame K, Gehani A, Alper P, editors. Proven Annot Data Process. Cham: Springer International Publishing; 2018. p. 45–57.

[7] D. Torre, G. Soltana, M. Sabetzadeh, et al. Using Models to Enable Compliance Checking Against the GDPR: An Experience Report. 2019 ACMIEEE 22nd Int Conf Model Driven Eng Lang Syst MODELS. 2019. p. 1–11.

[8] Ryan P, Crane M, Brennan R. Design Challenges for GDPR RegTech: Proc 22nd Int Conf Enterp Inf Syst [Internet]. Prague, Czech Republic: SCITEPRESS - Science and Technology Publications; 2020 [cited 2021 Mar 7]. p. 787–795. Available from: http://www.scitepress.org/DigitalLibrary/Link.aspx?doi=10.5220/0009464507870795.

[9] Pandit HJ, O'Sullivan D, Lewis D. GDPR Data Interoperability Model. 2018 [cited 2021 Mar 7]; Available from: https://zenodo.org/record/3246438.

[10] Pandit HJ, Polleres A, Bos B, et al. Creating a Vocabulary for Data Privacy: The First-Year Report of Data Privacy Vocabularies and Controls Community Group (DPVCG). In: Panetto H, Debruyne C, Hepp M, et al., editors. Move Meaningful Internet Syst OTM 2019 Conf [Internet]. Cham: Springer International Publishing; 2019 [cited 2020 Nov 23]. p. 714–730. Available from: http://link.springer.com/10.1007/978-3-030-33246-4_44.

[11] Pandit HJ, Debruyne C, O'Sullivan D, et al. GConsent - A Consent Ontology Based on the GDPR. In: Hitzler P, Fernández M, Janowicz K, et al., editors. Semantic Web [Internet]. Cham: Springer International Publishing; 2019 [cited 2021 Mar 7]. p. 270–282. Available from: http://link.springer.com/10.1007/978-3-030-21348-0_18.

[12] Ryan P, Pandit HJ, Brennan R. A Common Semantic Model of the GDPR Register of Processing Activities. In: Villata S, Harašta J, Křemen P, editors. Front Artif Intell Appl [Internet]. IOS Press; 2020 [cited 2021 Mar 7]. Available from: http://ebooks.iospress.nl/doi/10.3233/FAIA200876.

[13] Bonatti P, Kirrane S, Petrova I, et al. Machine Understandable Policies and GDPR Compliance Checking. KI - Künstl Intell. 2020;34.

[14] Leone V, Di Caro L. The Role of Vocabulary Mediation to Discover and Represent Relevant Information in Privacy Policies. In: Villata S, Harašta J, Křemen P, editors. Front Artif Intell Appl [Internet]. IOS Press; 2020 [cited 2021 Mar 7]. Available from: http://ebooks.iospress.nl/doi/10.3233/FAIA200851.

[15] Olbrechts A. Recommendations 01/2020 on measures that supplement transfer tools to ensure compliance with the EU level of protection of personal data [Internet]. Eur. Data Prot. Board - Eur. Data Prot. Board. 2020 [cited 2021 Mar 7]. Available from: https://edpb.europa.eu/our-work-tools/public-consultations-art-704/2020/recommendations-012020-measures-supplement-transfer_en.

[16] Olbrechts A. Recommendations 02/2020 on the European Essential Guarantees for surveillance measures [Internet]. Eur. Data Prot. Board - Eur. Data Prot. Board. 2020 [cited 2021 Mar 7]. Available from: https://edpb.europa.eu/our-work-tools/our-documents/preporki/recommendations-022020-european-essential-guarantees_en.

[17] Sauro PhD J. Does Better Usability Increase Customer Loyalty? – MeasuringU [Internet]. 2010 [cited 2021 Aug 16]. Available from: https://measuringu.com/usability-loyalty/.