

Experimenting TDM (I, Robot?)

Journée d'étude de l'Association belge pour le droit d'auteur, La transposition de la Directive 2019/790 sur le droit d'auteur et les droits voisins dans le marché unique numérique en droit belge

Brussels, 21 April 2023

Julien Cabay (Associate Professor ULB and ULiège, Faculty of Law, JurisLab)

julien.cabay@ulb.be

Thomas Vandamme (Ph.D. candidate ULB, Faculty of Engineering, LISA)

thomas.vandamme@ulb.be



Outline

- Introduction
- Experimenting TDM in IPSAM
- TDM: Technical Aspects
- An overview of TDM exceptions in the EU
- Belgian implementation
- Some practical and legal hurdles in IPSAM
- Call for action

Introduction

- **Text and Data Mining (TDM)?**

- Art. 2(2) DSM Dir. : « 'text and data mining' means any automated analytical technique aimed at analysing text and data in digital form in order to generate information which includes but is not limited to patterns, trends and correlations »

- **Two exceptions**

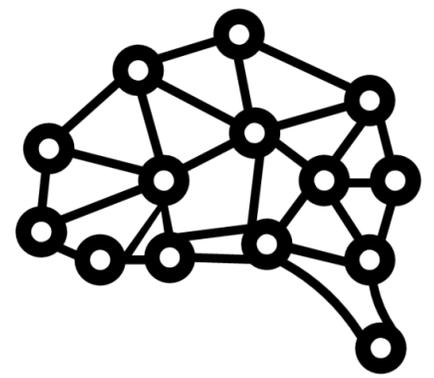
- Art. 3 (scientific research)
- Art. 4 (general)

- **Much more than copyright...**

- « (...) *It may be argued that under the misleading label Text and data mining (TDM) what has been regulated at the EU level in Arts. 3 and 4 goes far beyond a mere copyright exception. In fact, it should be reclassified as the legal regulation of AI via the allocation of property rights in its building blocks, or in other words, as a property-right approach to the regulation of AI (...)* » (Margoni/Kretschmer)

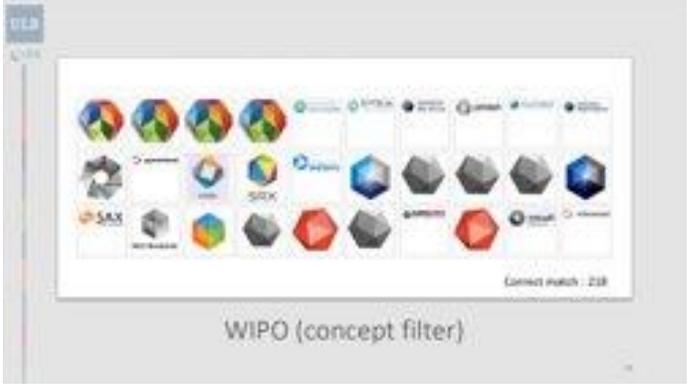
Introduction

- TDM Application
 - Search Engines
 - Consumer Statistics
 - Data-Intensive Applications :
 - Neural networks (see GPT trained on Common Crawl, 3.15B webpages)



Experimenting TDM in IPSAM

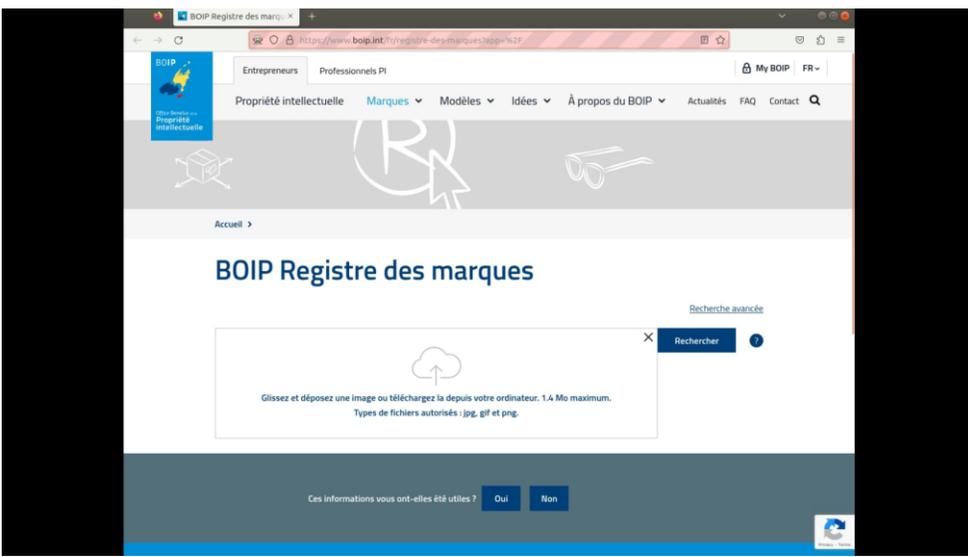
- Exploratory research : Assessing **TM search engines' performances**
- Using EUIPO **Opposition Division decisions**
- Query systems with applicant TM
- Performance -> Position of Opponent TM in output



J. Cabay and T. Vandamme, "Assessing IP Similarities Through Technology: A Trademark Exploration of Challenges and Avenues", AI Tech & Policy Talks, University of Geneva, 4 November 2021

Experimenting TDM in IPSAM

- Run experiment for 8k decisions (all EUIPO Opp. Div. From 23/03/2016 to 31/05/2022)
- => Necessity to automate

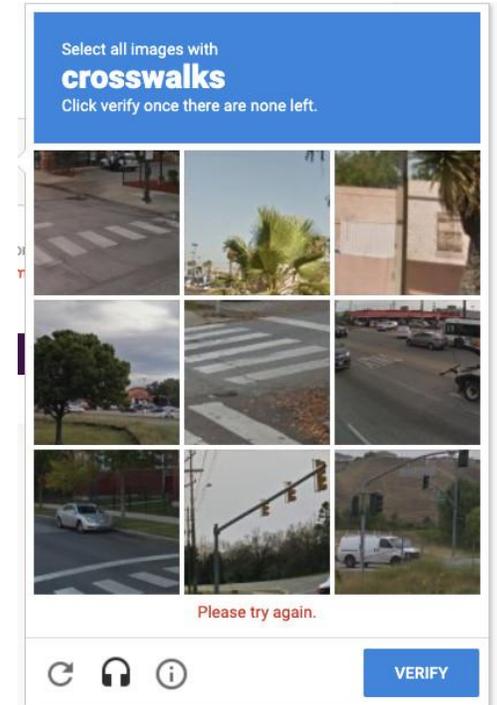
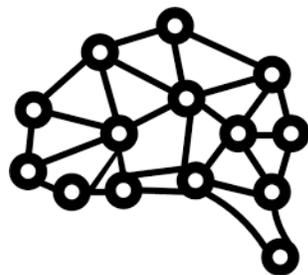


Line	App. No.	Date	Status	Classification	Other
159	0031021-30-09-20-1044851	+	LOC. +	28020(28028, 4)	0 EN
160	0030455-11-03-20-204063	-	LOC. -	44534259(487811)	0 EN
161	0023379-16-01-20-199250	+	LOC. +	43733465(403074)	0 EN
162	0030839-22-10-20-10540	+	LOC. +	43781489(451124)	0 EN
163	0030887-27-01-20-282116	+	LOC. +	47999642(179482)	0 EN
164	0030476-09-07-20-304955	+	LOC. +	41413261(150828)	0 EN
165	0030476-09-07-20-304955	+	LOC. +	47564139(550400)	0 EN
166	0026183-28-02-20-315092	+	LOC. +	44748561(833945)	0 EN
167	0031326-14-01-20-537358	+	LOC. +	41541418(112501)	0 EN
168	0030941-18-07-20-710699	+	LOC. +	43762295(631738)	1 EN
169	0027466-16-08-20-647346	-	LOC. -	45275589(152367)	0 EN
170	0031040-24-09-20-232540	+	LOC. +	48086515(531266)	0 EN
171	0031099-16-04-20-46215	+	LOC. +	48174243(180232)	0 EN
172	0026265-25-10-20-368843	+	LOC. +	44536437(136090)	0 EN
173	0026279-10-11-20-115269	+	LOC. +	44599815(843348)	0 EN
174	0031011-12-05-20-1026709	+	LOC. +	48088956(322505)	0 EN
175	0030789-28-02-20-1239581	+	LOC. +	48002693(621259)	0 EN
176	0024310-10-05-20-609691	-	LOC. -	43013991(124326)	0 EN
177	0030494-09-04-20-918011	+	LOC. +	43563891(127230)	0 EN
178	0031277-26-07-20-111894	+	LOC. +	48232556(924585)	0 EN
179	0026019-20-09-20-448764	+	LOC. +	44282545(106659)	0 EN
180	0031139-11-10-20-203776	+	LOC. +	48154477(302016)	0 EN
181	0028244-03-04-20-126710	+	LOC. +	45840457(301648)	0 EN
182	0029793-18-01-20-557747	+	LOC. +	47068842(118066)	1 EN
183	0028600-23-11-20-356180	+	LOC. +	44564961(325451)	0 EN
184	0031255-16-06-20-918293	+	LOC. +	48198190(181275)	0 EN
185	0030669-22-02-20-42451	+	LOC. +	47933861(153587)	0 EN
186	0028163-10-08-20-781531	+	LOC. +	45789308(526475)	0 EN
187	0031208-11-04-20-592855	+	LOC. +	48174699(121294)	0 EN
188	0030900-18-06-20-755449	+	LOC. +	48046277(152840)	0 EN
189	0031401-08-02-20-63552	+	LOC. +	48325551(302014)	409
190	0030709-17-12-20-810470	-	LOC. -	47949202(152493)	0 EN
191	0030256-04-02-20-882836	-	LOC. -	43707372(397042)	0 EN
192	0025536-21-12-20-196097	+	LOC. +	43764481(106921)	0 EN
193	0031444-30-03-20-259325	-	LOC. -	48365489(120683)	0 EN
194	0030065-23-07-20-2624	+	LOC. +	47397779(123553)	0 EN
195	0031239-27-07-20-138827	+	LOC. +	48199602(149161)	0 EN
196	0024062-22-09-20-171878	+	LOC. +	42847571(714422)	0 EN
197	0030904-15-10-20-76176	+	LOC. +	48056034(273838)	0 EN

Forthcoming : T. Vandamme, J. Cabay, O. Debeir, "A Quantitative Evaluation of Trademark Search Engines' Performances through Large-Scale Statistical Analysis", *ICAAIL '23: Nineteenth International Conference for Artificial Intelligence and Law, Braga Portugal, June 19- 23, 2023*

TDM: Technical Aspects

- Data acquisition :
 - Methods : Applications Programming Interface (APIs, gateways made available by services), Web Scrapers (robots using human interfaces)
 - Challenges : server issues and limitations, various protections (IP blocking, Captchas)
- Data treatment :
 - Methods : Parsing (pdfs, docs, ..), cleaning and filtering, relational information
 - Challenges : Irregular data formats, broken files and data
- Data analysis : models, insights, applications



An overview of TDM exceptions in the EU

« (...) in the Union, such organisations and institutions are confronted with **legal uncertainty** as to the extent to which they can perform text and data mining of content. In certain instances, **text and data mining can involve acts protected by copyright**, by the sui generis database right or by both, in particular, the reproduction of works or other subject matter, the extraction of contents from a database or both which occur for example when the data are normalised in the process of text and data mining. **Where no exception or limitation applies, an authorisation to undertake such acts is required from rightholders** »

(Recital 8, DSM Dir.)

An overview of TDM exceptions in the EU

- Relevance of reproduction right?
 - CJEU, *Infopaq*, C-5/08 (2009) « ***An act occurring during a data capture process (...) is such as to come within the concept of reproduction in part (...) if the elements thus reproduced are the expression of the intellectual creation of their author*** »
 - Critics: no « **use of a work as a work** » (Ducato/Strowel)
- Irrelevance of previous exceptions?
 - Limited scope, in particular in relation to commercial uses
 - Private use exception for AI business? Art. 5(2)(b) InfoSoc Dir.: « *neither directly nor indirectly commercial (...)* »
 - Research exception for AI business? Art. 5(3)(a) InfoSoc Dir.: « *non-commercial purpose* »
 - Exception for transient copies for AI business? Art. 5(1) InfoSoc. Dir.: « *no independent economic significance* »

An overview of TDM exceptions in the EU

Art. 3 (scientific research TDM)	Art. 4 (general TDM)
<p>(1) Member States shall provide for an exception to the rights provided for in Article 5(a) and Article 7(1) of Directive 96/9/EC, Article 2 of Directive 2001/29/EC, and Article 15(1) of this Directive for reproductions and extractions made by research organisations and cultural heritage institutions in order to carry out, for the purposes of scientific research, text and data mining of works or other subject matter to which they have lawful access.</p>	<p>(1) Member States shall provide for an exception or limitation to the rights provided for in Article 5(a) and Article 7(1) of Directive 96/9/EC, Article 2 of Directive 2001/29/EC, Article 4(1)(a) and (b) of Directive 2009/24/EC and Article 15(1) of this Directive for reproductions and extractions of lawfully accessible works and other subject matter for the purposes of text and data mining.</p>
<p>(2) Copies of works or other subject matter made in compliance with paragraph 1 shall be stored with an appropriate level of security and may be retained for the purposes of scientific research, including for the verification of research results.</p>	<p>(2) Reproductions and extractions made pursuant to paragraph 1 may be retained for as long as is necessary for the purposes of text and data mining.</p>
	<p>(3) The exception or limitation provided for in paragraph 1 shall apply on condition that the use of works and other subject matter referred to in that paragraph has not been expressly reserved by their rightholders in an appropriate manner, such as machine-readable means in the case of content made publicly available online.</p>
<p>Juncto Art. 7(1) : Any contractual provision contrary to the exceptions provided for in Articles 3, 5 and 6 shall be unenforceable.</p>	
<p>Juncto Art. 7(2) : The first, third and fifth subparagraphs of Article 6(4) of Directive 2001/29/EC shall apply to Articles 3 to 6 of this Directive.</p>	

An overview of TDM exceptions in the EU

Art. 3 (scientific research TDM)	Art. 4 (general TDM)
<p>(1) Member States shall provide for an exception to the rights provided for in Article 5(a) and Article 7(1) of Directive 96/9/EC, Article 2 of Directive 2001/29/EC, and Article 15(1) of this Directive for reproductions and extractions made by research organisations and cultural heritage institutions in order to carry out, for the purposes of scientific research, text and data mining of works or other subject matter to which they have lawful access.</p>	<p>(1) Member States shall provide for an exception or limitation to the rights provided for in Article 5(a) and Article 7(1) of Directive 96/9/EC, Article 2 of Directive 2001/29/EC, Article 4(1)(a) and (b) of Directive 2009/24/EC and Article 15(1) of this Directive for reproductions and extractions of lawfully accessible works and other subject matter for the purposes of text and data mining.</p>
<p>(2) Copies of works or other subject matter made in compliance with paragraph 1 shall be stored with an appropriate level of security and may be retained for the purposes of scientific research, including for the verification of research results.</p>	<p>(2) Reproductions and extractions made pursuant to paragraph 1 may be retained for as long as is necessary for the purposes of text and data mining.</p>
	<p>(3) The exception or limitation provided for in paragraph 1 shall apply on condition that the use of works and other subject matter referred to in that paragraph has not been expressly reserved by their rightholders in an appropriate manner, such as machine-readable means in the case of content made publicly available online.</p>
<p>Juncto Art. 7(1) : Any contractual provision contrary to the exceptions provided for in Articles 3, 5 and 6 shall be unenforceable.</p>	
<p>Juncto Art. 7(2) : The first, third and fifth subparagraphs of Article 6(4) of Directive 2001/29/EC shall apply to Articles 3 to 6 of this Directive.</p>	

An overview of TDM exceptions in the EU

- Many issues in relation to EU provisions, especially in relation to art. 4
 - Application to works made available prior adoption/entry into force of the directive?
 - See CJEU, *Soulier & Dore*, C-301/15 (2016) (no prior information on future use = hypothetical consent)
 - Applications to works with country of origin is not EU Member States?
 - See art. 5(2) Berne Convention: « The enjoyment and the exercise of these rights shall not be subject to any formality (...) »
 - Application to works not lawfully made available online?
 - See in particular CJEU, *Renckhoff*, C-161/17 (2018) (unauthorized making available online of work already lawfully made available online)
- Many different regimes worldwide with broader reach

An overview of TDM exceptions in the EU (and beyond)

	EU	China	USA	Japan
TDM	Art. 3 (scientific research) and 4 (general) Dir. 2019/790	Closed list, no TDM but 'open' clause (Art. 24(13) referring to 'other circumstances' possible 'fair use/fair dealing')	17 USC 107 (<i>fair use</i> , after <i>Authors Guild v. Google Inc.</i> , 804 F.3d 202 (2d Cir. 2015))	Art. 30-4 Japanese Copyright Act (exception 'not for enjoying the idea or emotions expressed in a work')
Reach	Limited	« Chinese courts have realised the disadvantage of limiting fair use to enumerated circumstances and brought in a more flexible approach for finding fair use by learning from their US counterparts » (Wang, He)	« Uses involving robotic readers are fast-tracked for fair use » (Grimmelmann) ; but doubts as to « expressive machine learning » (Sobel)	« Japan as a paradise for machine learning » (Ueno)

Belgian implementation

- Art. 2(2) (definitions)
 - l.13, 10° (text and data mining)
 - l.13, 9° (research organisation)
 - No use of wording ‘cultural heritage institution’ (and no definition), directly implemented in substantial provisions (*see Doc. Parl., Doc. parl., Chambre, 2021-2022, No 2608/1, p. 66*)

DSM Dir. / CDE	3 (scientific research TDM)	4 (general TDM)	7(1) (contradictory contractual provisions unenforceable)	7(2) (technological protection measures)
General © ; Database ©	XI.191/1, §1er, 7°	XI.190, 20°	XI.193	XI.291, § 2 *
Software ©	-	XI.299, § 5	XI.301	XI.291, § 2 *
Database <i>sui generis</i> right	XI.310, § 3, 1°	XI.310, § 3, 2°	XI. 314	XI.291, § 2 *❖
Neighbouring right, including new publisher’s right	XI.217/1, 6°	XI.217, 19°	XI.219	XI.291, § 2 *

Belgian implementation

- Explicit choices:
 - **No remuneration** for (scientific research) TDM exception
 - // recit. 17
 - *See Doc. Parl., Doc. parl., Chambre, 2021-2022, No 2608/1, p. 57 (limited prejudice)*
 - General TDM opt-out in the case of content made publicly available online **only if machine-readable means**
 - Art. XI.190, 20° CDE: « *En ce qui concerne les contenus mis à la disposition du public en ligne, la réservation n'est considérée appropriée que si elle est effectuée au moyen de procédés lisibles par machine.* »
 - *Contra art. 3(3) (« such as ») ; // recit. 18*
 - *See Doc. Parl., Doc. parl., Chambre, 2021-2022, No 2608/1, p. 49 (contra opinion Conseil d'Etat ; « cette manière de transposer favorise la sécurité juridique tant des titulaires de droits que des utilisateurs »)*
 - General TDM exception also '**imperative**', if contractual reservation not deemed 'appropriate'
 - Comp. art. 7(1), that only applies to scientific research TDM
 - *See Doc. Parl., Doc. parl., Chambre, 2021-2022, No 2608/1, p. 57 : « L'exception de l'article 4 concernant l'extraction de textes et de données permet aux titulaires de droits de faire une "réserve appropriée". Ainsi, on ne peut déroger à cette exception que si une telle réserve appropriée est faite. S'il n'y a pas de réserve appropriée, on ne peut s'écarter de cette exception, quelle que soit la stipulation contraire. L'interprétation de la réserve appropriée dépend de la jurisprudence. Toutefois, on ne peut exclure a priori que cela puisse se faire par convention. »*

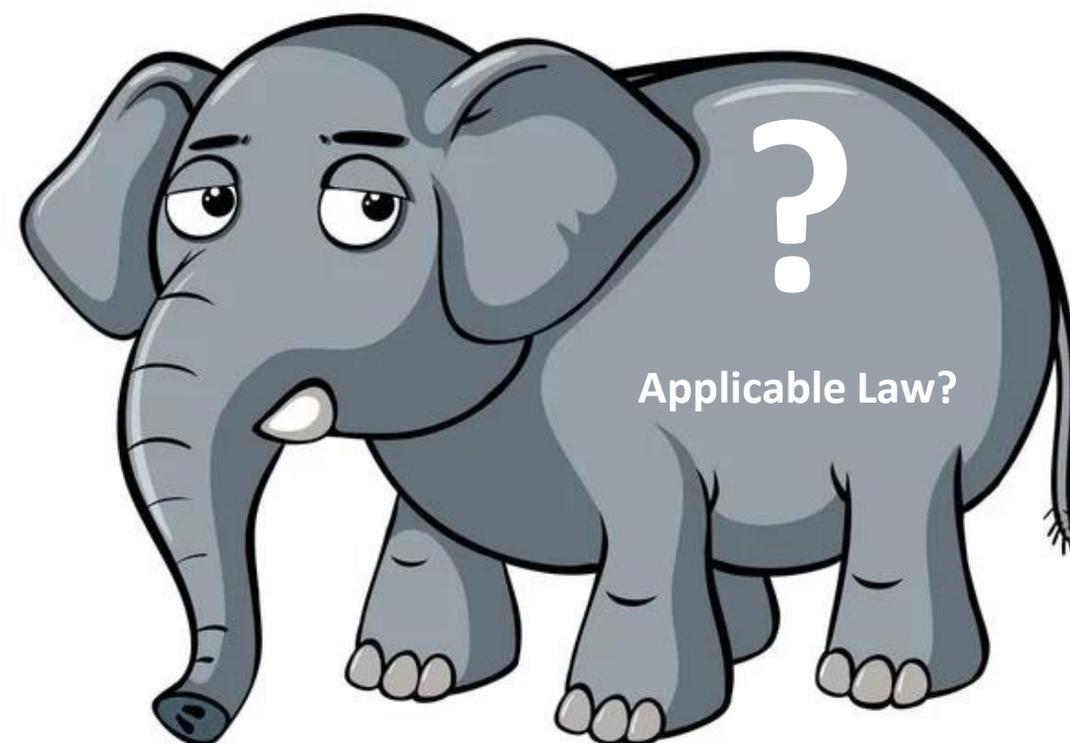
Belgian implementation

- Choices or mistakes?

- * Technological protection measures reservation for database *sui generis* right in XI.291, § 2
 - XI.291 refers to rightholders in the field of copyright ; should have been implemented in XI.316, § 2 (// in relation to the producer's database *sui generis* right)
- * XI.291, § 2 (voluntary measures by rightholders to ensure benefit of exceptions, notwithstanding technological protection measures), does not apply to works made available to the public online on agreed terms (see XI.291, § 3)
 - = all works freely available online? (maybe yes, see Dusollier)
 - If yes, contradiction with
 - Art. 7(2) DSM Dir. *juncto* art. 6(4), subparagraph 4 InfoSoc Dir. (not referred to)
 - Recit. 14 in relation to scientific research TDM : « *Lawful access should also cover access to content that is freely available online* »
 - Systematic interpretation Art. 3 and 4 DSM Dir. and Art. 6(4), subparagraph 4 InfoSoc Dir. (scientific research TDM on works freely accessible online should not be subject to contractual arrangement)

Some practical and legal hurdles in IPSAM

- Data acquisition online => ???
 - Robot.txt
 - = TDM reservation through machine-readable mean?
 - = Technological protection measure?
 - Disregard (scientific research TDM), VPN to avoid 'IP blocking'
 - Lawfull access?
 - = Circumvention?
 - Captcha
 - = TDM reservation through machine-readable mean?
 - = measures to ensure the security and integrity of the networks and database?
 - = Technological protection measure?
 - Use of Mechanical Turk (cheap outsourced labor)
 - = Lawfull access?
 - = Circumvention?
 - General/scientific research TDM overlap, yet former shall not affect the latter (art. 4(3) DSM Dir.) => solution in contractual arrangement?
 - = compatible with scientific research TDM?
 - = compatible with Academia?
 - = compatible with Open Science?
- Data treatment => OK
- Data analysis => OK



Call for action

- Need for more legal certainty & technical standardisation
- Belgium
 - Making use of Art. 3(4) DSM Dir. : « *Member States shall encourage rightholders, research organisations and cultural heritage institutions to **define commonly agreed best practices** concerning the application of the obligation and of the measures referred to in paragraphs 2 and 3 respectively* »
- EU
 - Amending the law with the introduction of **country of origin principle** (e.a. art. 5(3) DSM Dir.)
- Int'l
 - Reaching an **int'l consensus** to adress TDM (arguably outside of current int'l copyright law, see Senftleben)
 - Setting **technical standards** (TDM.txt // robot.txt?)



Cited sources

- Ducato, R., Strowel, A., “Ensuring text and data mining: remaining issues with the EU copyright exceptions and possible ways out”, *European Intellectual Property Review*, 2021, 43(5), 322-337
- Dusollier, S., *Droit d'auteur et protection des œuvres dans l'univers numérique : droits et exceptions à la lumière des dispositifs de verrouillage des œuvres*, Bruxelles, Larcier, 2007
- Grimmelmann, J., “Copyright for Literate Robots”, *101 Iowa Law Review* 657 (2016)
- Margoni, T., Kretschmer, M., “A Deeper Look into the EU Text and Data Mining Exceptions: Harmonisation, Data Ownership, and the Future of Technology” (July 14, 2021). Available at SSRN: <https://ssrn.com/abstract=3886695> or <http://dx.doi.org/10.2139/ssrn.3886695>
- Sobel, B. L. W., “Artificial Intelligence’s Fair Use Crisis”, *41 Columbia Journal of Law & the Arts* 45 (2017)
- Senftleben, M., “Compliance of National TDM Rules with International Copyright Law: An Overrated Nonissue?”, *IIC - International Review of Intellectual Property and Competition Law*, 53 (10), 2022, 1477-1505
- Ueno, T., “The Flexible Copyright Exception for ‘Non-Enjoyment’ Purposes – Recent Amendment in Japan and Its Implication”, *GRUR International*, 70(2), 2021, 145–152
- Vandamme, T., Cabay, J., Debeir, O., "A Quantitative Evaluation of Trademark Search Engines' Performances through Large-Scale Statistical Analysis ", *ICAIL '23: Nineteenth International Conference for Artificial Intelligence and Law, Braga Portugal, June 19- 23, 2023*
- Wang, J., He, T., “To share is fair: The changing face of China’s fair use doctrine in the sharing economy and beyond”, *Computer Law & Security Review*, 35(1), 2019, 15-28

Many thanks for your attention, comments and questions !

julien.cabay@ulb.be
thomas.vandamme@ulb.be

Find out more about IPSAM :

<https://droit-prive.ulb.be/ipsam-adressing-intellectual-property-relevant-similarities-in-images-through-algorithmic-decision-systems/>

Fundamental Rights, Copyright and Content Recognition (The Good, the Bad and the Ugly?)

Journée d'étude de l'Association belge pour le droit d'auteur, La transposition de la Directive 2019/790 sur le droit d'auteur et les droits voisins dans le marché unique numérique en droit belge

Brussels, 21 April 2023

Julien Cabay (Associate Professor ULB and ULiège, Faculty of Law, JurisLab)

julien.cabay@ulb.be

Thomas Vandamme (Ph.D. candidate ULB, Faculty of Engineering, LISA)

thomas.vandamme@ulb.be



Outline

- Article 17 and the 'Filtering' issue
- Safeguarding Fundamental Rights: From the EU, to the MS, to the service providers
- Technical Aspects of Filtering
- Insights from IPSAM
- Call for action

Article 17 and the 'Filtering' issue

- Art. 13 DSM Dir. Proposal
 - « *Information society service providers that store and provide to the public access to large amounts of works or other subject-matter uploaded by their users shall, in cooperation with rightholders, take measures to ensure the functioning of agreements concluded with rightholders for the use of their works or other subject-matter or to prevent the availability on their services of works or other subject-matter identified by rightholders through the cooperation with the service providers. Those measures, such as the **use of effective content recognition technologies**, shall be appropriate and proportionate. (...) »*
- Scraped in Art. 17 DSM Dir., but...
- CJEU, *Poland*, C-401/19 (2022)
 - « (...) *in order to be able to carry out such a prior review, online content-sharing service providers are, depending on the number of files uploaded and the type of protected subject matter in question, and within the limits set out in Article 17(5) of Directive 2019/790, required to use **automatic recognition and filtering tools**. In particular, neither the defendant institutions nor the interveners were able, at the hearing before the Court, to designate possible alternatives to such tools.* » (§ 54)

Article 17 and the 'Filtering' issue

- The main Fundamental Rights issue :
 - « (...) **a filtering system which might not distinguish adequately between unlawful content and lawful content**, with the result that its introduction could lead to the blocking of lawful communications, **would be incompatible** with the right to freedom of expression and information, guaranteed in Article 11 of the Charter, and **would not respect the fair balance** between that right and the right to intellectual property » (CJEU, Poland, C-401/19, (2022), § 86)
 - See also CJEU, *Scarlet Extended*, C-70/10 (2011), § 52; CJEU, *Sabam*, C-360/10 (2012), § 50

Article 17 and the 'Filtering' issue

- Fundamental Rights in the EU
 - Charter of Fundamental Rights
 - *'respect for those rights [recognised by the Charter] being a **condition of the lawfulness of EU acts**, so that measures incompatible with those rights are not acceptable in the EU'* (CJEU, Opinion 2/13, § 169)
 - *'situations cannot exist which are covered in that way by European Union law without those fundamental rights being applicable. The **applicability of European Union law entails applicability of the fundamental rights guaranteed by the Charter**'* (CJEU, Fransson, C-617/10, § 21)
 - Limitations?
 - Art. 52(1) : *'Any limitation on the exercise of the rights and freedoms recognised by this Charter must be **provided for by law and respect the essence** of those rights and freedoms. Subject to the principle of **proportionality**, limitations may be made only if they are necessary and genuinely meet objectives of general interest recognised by the Union or the need to protect the rights and freedoms of others'*
 - Conflict between Fundamental Rights?
 - *'reconcile the requirements of the protection of different fundamental rights' => 'fair balance'* (CJEU, Promusicae, §§ 65-68)
 - Addressees
 - EU institutions when adopting EU law (CJUE, Deutsches Weintor, C-544/10, § 47)
 - MS legislator when implementing EU law (CJEU, Promusicae, § 68)
 - MS jurisdictions when applying implemented EU law (CJUE, Deckmyn, C-201/13, § 32)
 - What about private companies? *See infra*

Article 17 and the ‘Filtering’ issue

- Art. 17 as a topical exemple of ‘fair balance’ analysis
 - See CJEU reasoning in *Poland*, C-401/19, §§ 59-100

Charter Fundamental Rights EU	Article 17 DSM Directive
Protection of personal data (8)	<ul style="list-style-type: none"> - No general monitoring obligation (8) - No identification of individual users / processing of personal data, except in accordance with GDPR <i>e.a.</i> (9)
Freedom of expression and information (11) Freedom of the arts and sciences (13)	<ul style="list-style-type: none"> - Autorisation OCSSP shall also cover acts carried out by users (2) - No prevention of legitimate uses, including copyrihgt exceptions (7)
Freedom to conduct a business (16)	<ul style="list-style-type: none"> - Taking into account features of service, availability/costs means (5) - Taking into account market position (6)
Intellectual property (17(2))	<ul style="list-style-type: none"> - Extension communication to the public right (1)
Right to an effective remedy (47)	<ul style="list-style-type: none"> - Complaint and redress mechanism (9)

Article 17 and the 'Filtering' issue

- Filters shall « *respect the **essence** of the right to freedom of expression and information* » (Poland, § 76):
 - « *In that regard, it must be noted that the first subparagraph of Article 17(7) of Directive 2019/790 expressly states that the 'cooperation between online content-sharing service providers and rightholders **shall not result in the prevention of the availability of works** or other subject matter uploaded by users, **which do not infringe copyright** and related rights, including where such works or other subject matter are covered by an exception or limitation' of those rights* » (§ 77)
 - « *According to its unambiguous wording, the first subparagraph of Article 17(7) of Directive 2019/790, unlike point (b) and point (c), in fine, of Article 17(4) of that directive, is **not limited** to requiring online content-sharing service providers **to make their 'best efforts'** to that end, **but prescribes a specific result to be achieved*** » (§ 78)

Article 17 and the 'Filtering' issue

- CJEU in *Poland* did not rule out that filters might strike a **'fair balance'**, **provided** (among others) that:
 - **'strictly targeted'** (§ 81)
 - **'by excluding, in particular, measures which filter and block lawful content** when uploading' (§ 85)
 - **'only on condition that the rightholders concerned provide them with the relevant and necessary information** with regard to that content' (§ 89)
 - **'services cannot be required to prevent the uploading and making available to the public of content which, in order to be found unlawful, would require an independent assessment of the content** by them in the light of the information provided by the rightholders and of any exceptions and limitations to copyright' (§ 90)
 - **'procedural safeguards'** (§ 93)
- See also
 - CJEU, *YouTube and Cyando*, C-682/18 and C-683/18, § 102 (« *appropriate technological measures that can be expected from a reasonably diligent operator in its situation in order to counter credibly and effectively copyright infringements on that platform* »)
 - Art. 7 Digital Services Act [Regulation 2022/2065] ('Voluntary own-initiative investigations')

Safeguarding Fundamental Rights: From the EU, to the MS, to the service providers

- CJEU, *Poland*, § 71:
 - « *In addition, the present examination, in the light of the requirements laid down in Article 52(1) of the Charter, concerns the specific liability regime in respect of online content-sharing service providers, as established by Article 17(4) of Directive 2019/790, which **does not prejudge any examination which may subsequently be carried out in relation to the provisions adopted by the Member States for the purposes of transposing that directive or of the measures determined by those providers in order to comply with that regime.** »*

Safeguarding Fundamental Rights: From the EU, to the MS, to the service providers

- Stakeholders dialogue (EU Commission Communication, *Guidance on Article 17 of Directive 2019/790 on Copyright in the Digital Single Market*, COM(2021) 288 final)
 - **Technology neutrality**
 - « Article 17(4)(b) should be implemented in a technologically neutral and future proof manner. Therefore, the Member States should not in their implementing laws mandate either the use of a technological solution nor impose any specific technological solutions on service providers in order to demonstrate best efforts. »
 - « (...) online content-sharing service providers should remain free to choose the technology or the solution to comply with the best efforts obligation in their specific situation.
 - **Authorised blocking of ‘manifestly infringing upload’**
 - « (...) preventing the upload by the use of technology, should in principle be limited to manifestly infringing uploads »
 - **Ex post human review for noticed ‘non manifestly infringing uploads’**
 - « (...) other uploads, which are not manifestly infringing, should in principle go online and may be subject to an ex post human review when rightholders oppose by sending a notice »
 - **Information**
 - « Such information could for example include a description of the **type of technologies** (if any) or other means used by the service providers, information on third party technology providers whose services they may use, the **average level of efficiency** of these tools, any changes to the tools/services used (such as possible updates or changes in the use of third party services). Service providers should **not** be obliged to give specific information which would go against their **business secrets**, such as detailed characteristics of the software used, which may be proprietary. »

Safeguarding Fundamental Rights: From the EU, to the MS, to the service providers

- Belgian implementation, referral to the King (taking into account the stakeholders dialogue)
 - **‘relevant and necessary information’** (*Poland, § 89*)
 - XI.228/5, § 4 CDE: « (...) modalités en lien avec les conditions fixées aux paragraphes 1er à 3, notamment en ce qui concerne la notification et les informations pertinentes et nécessaires »
 - **‘procedural safeguards’** (*Poland, § 93*)
 - XI.228/8, § 4 CDE: « (...) modalités en lien avec les dispositifs de traitement des plaintes et de recours visés aux paragraphes 1er à 3, notamment en ce qui concerne le délai dans lequel ces plaintes doivent être traitées, la procédure à suivre et la situation du contenu faisant l'objet de la plainte au cours du traitement de celle-ci »
 - **Information on the functioning of the technology**
 - XI.228/7, § 3 CDE : « (...) modalités en lien avec l'obligation d'information visée aux paragraphes 1er et 2 [obligation d'information] »

Safeguarding Fundamental Rights: From the EU, to the MS, to the service providers

- Assumption: it is possible to build technologies that will strike a 'fair balance', yet:
 - Stakeholders dialogue: « (...) **In the present state of the art, no technology can assess to the standard required in law whether content, which a user wishes to upload, is infringing or a legitimate use (...)** »
 - Adv. Général Saugmandsgarrd Øe in *Poland*, § 212 : « (...) the obligation laid down in Article 17(7) of Directive 2019/790 does not mean that the mechanisms which lead to a negligible number of cases of 'false positives' are automatically contrary to that provision. Nevertheless, the error rate should be as low as possible. It follows that, **in situations in which it is not possible, in the current state of technology**, for example as regards certain types of works and protected subject matter, **to use an automatic filtering tool without resulting in a 'false positive' rate that is significant, the use of such a tool should, in my view, be precluded** under paragraph 7.
- => SOTA?

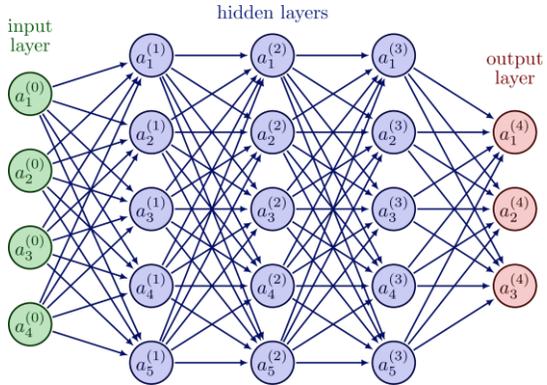
Technical aspects of filtering

- Watermarking
 - Secret signature embedded into works
 - Robust to various transformations
 - Invisible, yet retrievable by owner
 - No false positive possible



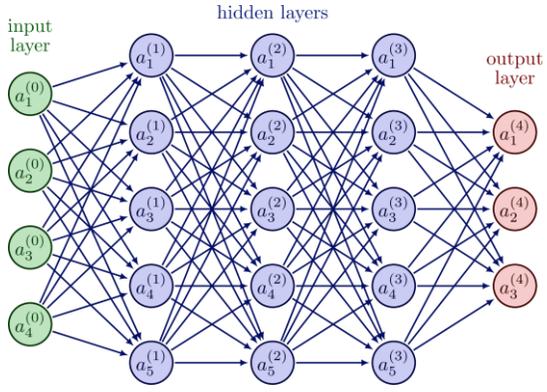
Technical aspects of filtering

- Fingerprinting
 - Fingerprints (lists of numbers) represent the work
 - Comparing fingerprints // comparing the works (ideally..)
 - Recent techniques: Deep Learning (Artificial Neural Networks)



Technical aspects of filtering

- Deep Learning
 - Complex mathematical models (millions of parameters)
 - Require massive amounts of examples for training (also millions)
 - Quality of model depends on quality/relevance of data (wrt task):
 « **Garbage in, Garbage out** »



Insights from IPSAM

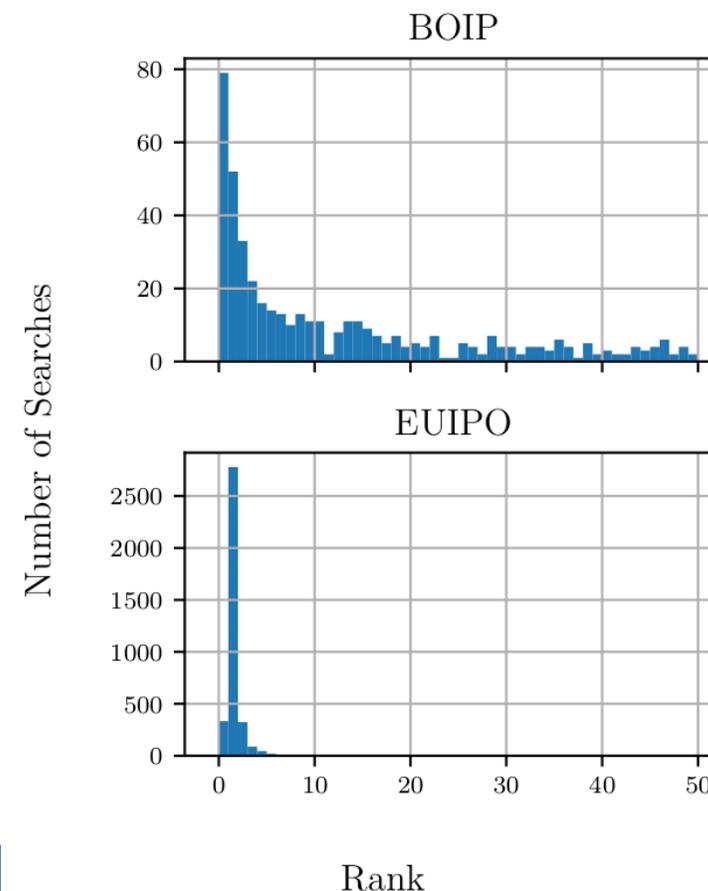
- Assessment of TM Search Engines (BOIP + EUIPO)
- Using EUIPO Opp. Div. Decisions (**task-relevant**)

=> **Very poor results !**

TM Law -> numerous task-relevant labeled data and standard objects.

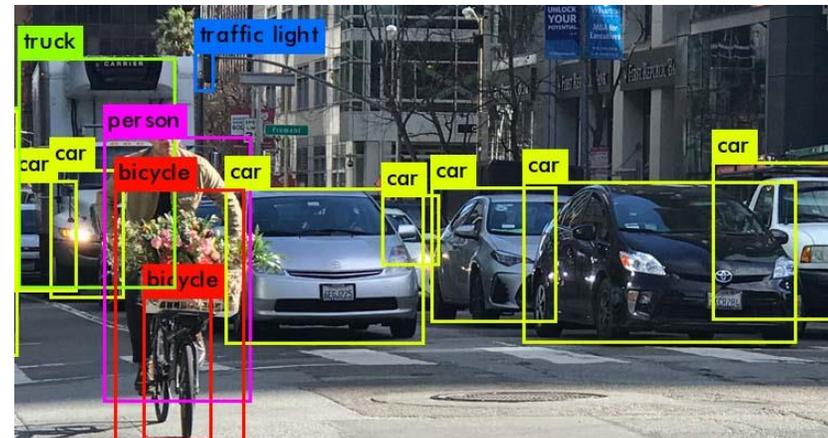
What about Copyright Law ?

System	Decisions	Matches	Ratio
BOIP	5 761	445	7.72 %
EUIPO	5 852	3 647	62.32 %



Insights from IPSAM

- Based on SOTA, is it **possible to build a « filtering system which might distinguish adequately between unlawful content and lawful content »**?
- It supposes:
 - (1) Possibility to identify what is **‘manifestly infringing’**
 - (2) Possibility of a **‘human review’** of **‘non manifestly infringing upload’** that would **not** amount to an **‘independent assessment’**
- (1) and (2) are function of **Quality** and **Quantity** of data as to infringement?
 - Case law (issue left to judges) = Data
 - Legal assessment = Labels
 - (A) Requirements for protection are met?
 - (B) Requirements for infringement are met?
 - (C) Requirements for exception are not met?

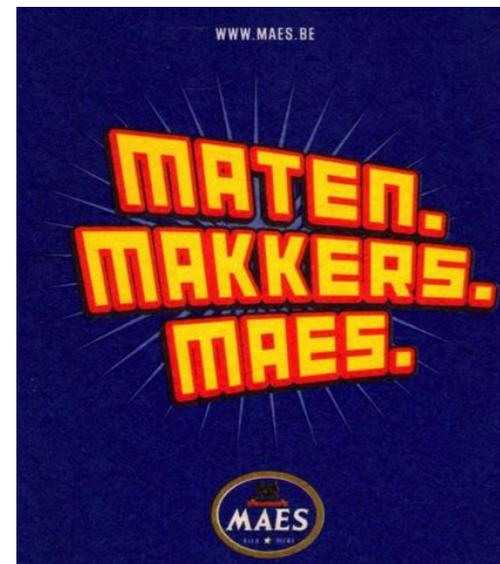


Insights from IPSAM

- Quality of data?
 - (A) Originality?



Civ. Mons (cess.), 18 Nov. 2005 (Yes)
 Mons, 3 Feb. 2014 (No)



Comm. Anvers (cess.), 17 June 2008 (No)
 Anvers, 29 June 2009 (Yes)

Insights from IPSAM

- Quality of data?
 - (B) Similarity?



Civ. Bruxelles (cess.), 11 Mar. 2005 (Yes)
Bruxelles, 6 Dec. 2007 (No)



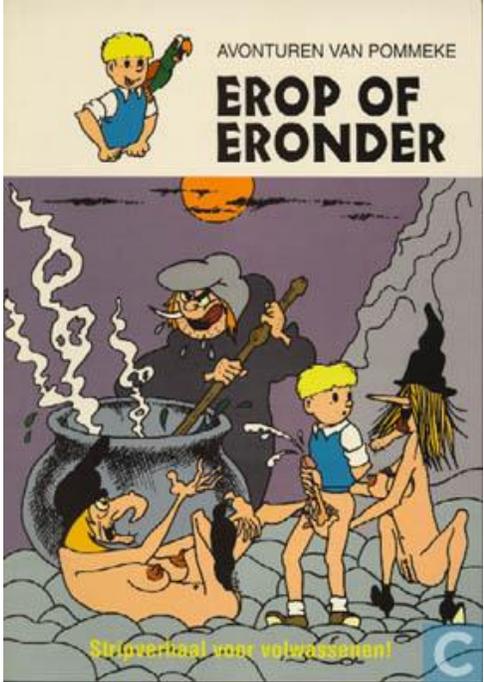
Comm. Bruxelles (cess.), 17 Sp. 2008 (No)
Bruxelles, 12 Apr. 2011 (Yes)

Insights from IPSAM

- Quality of data?
 - (C) Parody?



Bruxelles, 14 juin 2007 (Yes)



Anvers, 11 octobre 2000 (No)

Insights from IPSAM

- Quality of data? The example of relevant similarities
 - **No set test in Belgian** case law (Cabay 2012)
 - **No harmonised test at EU** level (Cabay 2016)
 - Ex. 'overall impression'
 - Netherlands: YES (Hoge Raad, 12 Apr. 2013)
 - France: NO (Cass. (com.) (France), 8 Apr. 2014)
 - Belgium : 'YES' (Cass. 25 Sep. 2003)
 - Where test is more stable (USA), **no consistency** (Asay)
 - **Biases** in similarity assessment (Balaganesh, Manta, Wilkinson-Ryan)

Insights from IPSAM

- Quantity of data? The example of relevant similarities

	© (Cabay 2012)	TM (Cabay/Vandamme/Debeir)
Period sample	1886-2011 (= 125 years)	2016-2022 (= 7 years)
Number of decisions	87 (72 infringement + 15 parody)	5.852
Type of objects	Several (graphical, applied arts, musical, literary, architecture, cinematographic, theater)	One (2D images)
Infringement test	Not set	LoC (CJEU case law)

Insights from IPSAM

- **Poor performances of SOTA in TM (IPSAM) despite 'Quantitative/Qualitative' Data :**
 - 7,72% (BOIP)
 - 62,32% (EUIPO)*
 - Probably much less (*supra*)
 - = irrelevance of “false positive’ rate’ (Adv. Gen.) or ‘average level of efficiency’ (EC)
- **Based on SOTA, it is certainly not possible to build in © a « filtering system which might distinguish adequately between unlawful content and lawful content » since no/less 'Quantitative/Qualitative' Data**
 - It will not be possible to identify what is ‘manifestly infringing’ in many cases
 - Any ‘human review’ of ‘non manifestly infringing upload’ would amount to an ‘independent assessment’

Insights from IPSAM

- The Good = identical content;
- The Bad = equivalent content
- The Ugly = similar content
- => **Any technology that would match 'similar' contents (other than 'identical/equivalent') is incompatible with Fundamental Rights and shall be prohibited (Cabay 2020)**



Call for action

- Need for more legal certainty & technical transparency
- Belgium/EU (taking into account non-stakeholders expertise)
 - **Prohibit** explicitly technology targeting 'similar' contents
 - Specify 'relevant and necessary **information**' to be provided by right holders
 - Adopt **transparency** obligation as to the 'Information on the functioning of the technology'
- Join us at ULiège on May 12!
 - J. Cabay, E. Rosati, "Regulating Technology Through Copyright Law: The Way Forward for Building a Digital Single Market?", in P. Van Cleynenbreugel, J. Wildemeersch (eds), *Questions Choieses de Droit Européen des Affaires/Selected Issues in European Business Law*, Bruxelles, Bruylant, 2023, pp. 81-114



60 years of European law at ULiège

Echoes from the past,
prospects for the future



11 & 12 MAY
2023

7 points OBFG

Journées internationales David-Constant
Évènement soutenu par le Fonds David-Constant

INFO & REGISTRATION

www.eulegalstudies.uliege.be/60-ans-droit-europeen

Cited sources

- Cabay, J., “‘Qui veut gagner son procès?’ – L’avis du public dans l’appréciation de la contrefaçon en droit d’auteur”, *A&M*, 2012, p. 13-29
- Cabay, J., *L’objet de la protection du droit d’auteur – Contribution à l’étude de la liberté de création*, Ph.D. Thesis, Université libre de Bruxelles, 2016
- Cabay, J., “Lecture prospective de l’article 17 de la directive sur le droit d’auteur dans le marché unique numérique : Vers une obligation de filtrage limitée par la CJUE, garante du « juste équilibre »”, in J. de Werra (éd.), *Propriété intellectuelle à l’ère du Big Data et de la Blockchain – Intellectual Property in the era of Big Data and Blockchain*, Genève - Zürich, Schulthess Editions Romandes, 2020, pp. 169-273
- Asay, C.D., “An Empirical Study of Copyright’s Substantial Similarity Test”, *UC Irvine Law Review*, 2022, Vol. 13, No. 1, pp. 35-102
- Balganesch, S., Manta, I.D., Wilkinson-Ryan, T., “Judging Similarity”, *Iowa Law Review*, 2014, Vol. 100, No. 1, pp. 267-290
- Vandamme, T., Cabay, J., Debeir, O., “A Quantitative Evaluation of Trademark Search Engines’ Performances through Large-Scale Statistical Analysis”, *ICAIL '23: Nineteenth International Conference for Artificial Intelligence and Law, Braga Portugal, June 19- 23, 2023*

Many thanks (again) for your attention, comments and questions !

julien.cabay@ulb.be

thomas.vandamme@ulb.be

Find out more about IPSAM :

<https://droit-prive.ulb.be/ipsam-adressing-intellectual-property-relevant-similarities-in-images-through-algorithmic-decision-systems/>