Defend Digital Me Briefing (working draft)
Copyright, Children's content and AI

1. Introduction **2**
2. Children's Rights in the context of Copyright and IP **4**
3. Other matters on copyright compliance **5**
What about learners' moral rights? 5
What about teachers' work and copyright? 6
4. The overlap with UK Data Protection law compliance **6**
UK Data Protection law is moving away from the GDPR standards 7
5. The UN Convention on the Rights of the Child **7**
UNCRC Article 16: Every child has the right to privacy 8
UNCRC Article 32: Protection from Economic Exploitation 8
6. Rights under the ECHR **8**
Article 8: Privacy of communications 8
Handwriting constitutes Personal Data 8
7. The special case of state education **9**
Non-consensual environment 9
Power imbalance renders consent invalid (GDPR Sweden: DPA case) 10
Children as the rights holder and the overlap with parental rights 10
Commercial Use Licensing creates a duty [of the institution] towards the rights-holder for the lifetime of the licence 12
8. Proposed UK changes rely on flawed EU comparison **12**
The context of text and data mining in European laws 12
Article 3: TDM for Scientific Research 13
Article 4: General TDM Exception 13
The Three-Step Test and Its Implications for AI Training and TDM 14
ECJ application of the Three-Step Test 15
Case Law 16
9. Government public engagement on AI in education **17**
"There was widespread consensus that work and data should not be used without parents' and/or pupils' explicit agreement." 17
10. Summary conclusions **18**

defend|digital|me

1. Introduction

The DSIT Copyright Consultation[1] sought, "*views on how the government can ensure the UK's legal framework for AI and copyright supports the UK creative industries and AI sector together.*"

Paragraph 60 suggested that the proposals are merely bringing the UK in line with other jurisdictions. "*However, it is highly likely to make the UK significantly less competitive compared to other jurisdictions – such as the EU and US – which do not have such restrictive laws.*"[2] The comparison is flawed as there are stronger safeguards in place in the EU law, than proposed in the UK changes.

Similarly, the Lords Committee, in its February 2025 report[3], found:

> *"Matt Clifford's AI Opportunities Action Plan recommended that the UK reform its text and data mining regime so that it is "at least as competitive as the EU" 389*
>
> *"**Our report on the future of news cautioned strongly against "adopting a flawed opt-out regime comparable to the version operating in the EU".** Witnesses to that inquiry told us the EU's regime lacked transparency about illegal scraping and the use of crawlers, as well as a clear enforcement mechanism for infringements."* [176, p.49]

Only the consultation question 16 allows space to answer the consultation section C3 which is a wider subject than consultation items 99-102, **Use of AI in education.** Our response considers only this topic. While welcome that the sector is not entirely overlooked, the consultation is not well suited to respond on the Use of AI in education. This is inadequate to properly consult on IP and education in particular with regard to the special case of (a) children's rights and (b) state education. We therefore address some of the reasons.

"99. Much of the public debate about copyright and AI concentrates on creative professionals, the creative industries, and AI developers. However, a range of other individuals and bodies also create and use copyright works and may be affected by any policy changes. We want to ensure that all views are considered.

100. For example, copyright and the use of AI tools is relevant to pupils, schools, and the wider education system. We want to find approaches to ensure that:

- *high-quality AI tools that support teachers can be developed*
- *the rights of pupils as creators of intellectual property can be protected*
- *the management of pupil IP does not add unnecessary burdens to the education workforce*

[1] Copyright and Artificial Intelligence consultation seeking "views on how the government can ensure the UK's legal framework for AI and copyright supports the UK creative industries and AI sector together." Closing February 25, 2025 https://www.gov.uk/government/consultations/copyright-and-artificial-intelligence

[2] Ibid paragraph 60 (archived) https://web.archive.org/web/20250117225523/https://www.gov.uk/government/consultations/copyright-and-artificial-intelligence/copyright-and-artificial-intelligence

[3] 'House of Lords Communications and Digital Committee Report: 2nd report of session 2024-25 AI and creative technology scale-ups: less talk, more action' (2025). https://publications.parliament.uk/pa/ld5901/ldselect/ldcomm/71/71.pdf

101. The IPO will support the Department for Education to explore this with children, young people, parents, carers, and teachers, to ensure their views are considered in the design of any subsequent processes or solutions.

102. We would welcome views on issues relating to copyright and AI which affect other specific sectors, bodies, and individuals."

The House of Lords Communications and Digital Committee Report also noted that,

> "*The consultation document does not include details about enforcement or sanctions for non-compliance under such a regime. If, following its consultation, the Government decides to progress its proposals for a broad text and data mining exemption with a mechanism to allow rights holders to reserve their rights, this must be underpinned by strong transparency measures, technical enforceability, and meaningful sanctions.*"[4]

Current intentions appear to be that personal data will be made anonymous in order to remove obstacles to processing personal data for incompatible purposes (as AI development is not why a child goes to school and why their data was collected at the point it was given to an educational setting).

However, if data and content are made anonymous (which is very difficult to do it at all possible) it will be impossible to manage the moral rights obligations of copyright licence management, that mean the rights' holder (the learner) must be known to the licence manager and able to be contacted in regard to the licence.

Both positions cannot be met, and would yet both be necessary in law. The first, because consent is rarely a lawful valid basis for personal data processing in educational settings today, and would unlikely in future for AI development. If an opt-out and not opt in mechanism is planned, then consent is not applicable anyway. Consent must be an active and informed decision and cannot rely on opt-out.

The latter, anonymisation, would not be valid to uphold the demands of copyright law with regard to the assertion of moral rights and obligations to be able to maintain the relationship with the author over time in order to manage the ongoing licence obligations.

In summary, we find that the intentions to reuse personal data and content from learners' education for the purposes of AI development, commercial in particular but otherwise in principle also, are incompatible with UK law, as well as deeply unethical. Even if copyright law and data protection law were changed in UK law, to seek to make what is currently unlawful lawful, it would be incompatible with the obligations of the Convention 108 and wider instruments and remain lacking in legal scrutiny, and ethical underpinnings.

---

[4] The Communications and Digital Committee Report (February 2025) AI and creative technology scale-ups: less talk, more Action HL Paper 71. Page 50, para 179. https://publications.parliament.uk/pa/ld5901/ldselect/ldcomm/71/71.pdf

2. Children's Rights in the context of Copyright and IP

The Government prefers a rights' reservation mechanism. This does not work for children and any learner in the disempowered environment of education.

Children and learners of all ages, as rights holders, whether as copyright (IP law) or personal data (data protection law) are ill-served by the status quo. The special environment of education means learners have limited agency and the power imbalance is a detriment to the promotion of their lifelong rights. Adults take decisions on their behalf and for learners of all ages there is a disincentive to prioritise individual interests over the interests of the institution.

If the government may seek to gain financially or intangibly from AI development, especially if it is promised to the same institution that may also be responsible for your assessment with lifetime outcomes, learners will be under pressure to give up their rights in return for the financial return. This is not a consent model, as the pressure affects the freely given nature of valid consent, and it is not protective of learners' rights. Furthermore, schools will not be able to manage its scale for the lifetime of pupils with whom they no longer have any relationship. Indeed, the academisation process of schools sees one legal entity end and another begin, and lack of clarity can exist over data today.

The current IP mechanism is not strong for children and has never been considered in these ways in the education system at scale before. However, the government has rejected this option, described in the consultation as Option zero:

"Option 0: Do nothing (status quo)

- *The existing legal framework remains unchanged.*
- *Legal uncertainty persists, making it difficult for right holders to enforce rights and for AI developers to operate in the UK.*
- *Risk of AI model training shifting overseas.*

*Government **rejects** this option but seeks feedback."*

While Option 1 is also rejected by the government as having restrictive impact on AI development, strengthening copyright and rights protection broadly for learners in the disempowered education environment would better protect their rights, which today are not upheld in national pupil data processing[5], but without public knowledge.[6]

"Option 1: Strengthen copyright (require licensing in all cases)

- *AI developers must obtain explicit licences for copyrighted material.*
- *Increases control and remuneration for right holders.*
- *Could discourage AI investment in the UK due to licensing complexity.*
- ***Not preferred** due to its restrictive impact on AI development."*

[5] Defend Digital Me (2025) The National Pupil Database: A Timeline https://defenddigitalme.org/national-pupil-data-the-ico-audit-and-our-work-for-change-a-timeline/

[6] Survation (2018) poll commissioned by Defend Digital Me. The majority of parents (69%) polled said they had not been informed that the Department for Education may give their child's information to third parties https://www.survation.com/1-in-4-parents-dont-know-child-signed-systems-using-personal-data/

There is no merit in Option 2 other than for those who profit from AI development, undermining rights protection for learners and placing them at further disadvantage, and economic exploitation simply by going to school.

"Option 2: Broad data mining exception

- *Allows AI training on copyrighted material without permission.*
- *Boosts AI innovation but removes control and remuneration for right holders.*
- *Favours large AI firms over smaller businesses due to legal uncertainty.*
- ***Not preferred** it fails to balance control, access, and transparency."*

While Option 3 is the government preferred option, this is imbalance in favour of presumption of use, which favours AI users not the rights holders. When it comes to children, and learners of all ages, this is not a practical nor a moral protection.

"Option 3: Data mining exception with a rights reservation mechanism (preferred option)

- *Allows TDM for any purpose, but right holders can **opt out** using a machine-readable rights reservation system.*
- *Balances access to data for AI developers with control for right holders.*
- *Supports **licensing, legal clarity, and transparency**.*
- *Government's **preferred approach** for consultation."*

This is true for both learners' data and content created in educational settings, including any computer generated content they may create for example in vocational training and further education, or any relevant classes with an element of computer generated content.

In summary, the current IP mechanism is not strong for children because it is not deigned with children in mind, and IP has never been considered in these ways in the education system at scale before, because pupils do not go to school to have their work and outputs exploited by others. This is an ethical as well as legal debate still to be had.

3. Other matters on copyright compliance

What about learners' moral rights?

Non-economic interests are also protected under copyright as '**moral rights**'.

This means the assumption is mistaken, to suggest that personal data or content could be anonymised and be made therefore 'exempt' from copyright protection. Section 77 of the Copyright, Designs and Patents Act 1988, states that anonymisation of pupils' work may be an infringement of the moral rights of the author at the AI input level. Their rights of attribution and their rights of integrity are intrinsic to the child as an author'[7] (and their steward acting on their behalf).

[7] Copyright, Designs and Patents Act 1988 Part I Chapter IV Clause 77 legislation.gov.uk/ukpga/1988/48/part/I/chapter/IV

Schools will be unable to manage this, nor the DfE at scale for millions of people for lifetime plus 70 years, the duration of copyright licenses.

What about teachers' work and copyright?

This is an essential question often connected with learners' work. Although much of their work may be 'works made in the course of employment', not all works are.[8] (See case Law since Noah v Shuba [1991] FSR 14).

Rights holders are individual and therefore licences for staff would need management on an individual basis, for individual works, for the duration of the licence. Negotiating licences with companies is likely to be legally complex and burdensome and cannot be lawfully grouped into a mass agreement, on a whole-school basis 'on behalf of' all staff, for example.

4. The overlap with UK Data Protection law compliance

High level summary

- AI developers must also comply with data protection law when mining text and data that includes personal information.
- The concept of 'consent' in data protection terms, is distinct from the concept of consent or 'permission' in other regimes such as copyright.[9]
- The context in which personal data is collected matters – in particular
    i. the relationship with the individual data subject and what they would reasonably expect;
    ii. any link between your original purpose why data was collected, and the new purpose;
    iii. the risk level to the fundamental rights and freedoms of the data subject (the individual) as well as of the personal data.

Purpose limitation

- The purpose limitation principle places a duty on data controllers that: Personal data shall be: (b) collected for specified, explicit and legitimate purposes and not further processed in a manner that is incompatible with those purposes; further processing for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes shall, in accordance with Article 89(1), not be considered to be incompatible with the initial purposes.
- Developers who are reusing personal data for training AI must consider whether the purpose of training a model is compatible with their original purpose of collecting that data. This is called a compatibility assessment.[10]

[8] Schools Week (2017) Is it legal for teachers to sell teaching resources? https://schoolsweek.co.uk/is-it-legal-for-state-school-teachers-to-sell-teaching-resources/
[9] The ICO (2024) The lawful basis for web scraping to train generative AI models https://web.archive.org/web/20250216002926/https://ico.org.uk/about-the-ico/what-we-do/our-work-on-artificial-intelligence/response-to-the-consultation-series-on-generative-ai/the-lawful-basis-for-web-scraping-to-train-generative-ai-models/
[10] https://ico.org.uk/about-the-ico/what-we-do/our-work-on-artificial-intelligence/response-to-the-consultation-series-on-generative-ai/purpose-limitation-in-the-generative-ai-lifecycle/

- As a general rule, if the new purpose is either very different from the original purpose, would be unexpected, or would have an unjustified impact on the individual, it is likely to be incompatible with your original purpose. In practice, you are likely to need to ask for specific consent to use or disclose data for this type of purpose.

Right to be Informed[11]

- The data controller must provide privacy information to individuals at the time their personal data is collected from them.
- If personal data is obtained from other sources, they must provide individuals with privacy information within a reasonable period of obtaining the data and no later than one month.

UK Data Protection law is moving away from the GDPR standards

The LIBE Committee noted at the end of 2022, that the UK reform is, "*focused solely on economy, innovation and growth and did not make a link to protection of fundamental rights.*" In fact, the UK is not only moving away from the GDPR, but when it comes to education more specifically, has not yet made efforts to put into practice the Council of Europe Data Protection Guidelines for Educational Settings (2020) adopted by the Committee of the Convention 108, the drafting of which the UK was involved in, and is a signatory.

Changes to UK Data Protection law currently underway, undermine the very essence of what data protection law is for; to prioritise the protections of the person from arbitrary interference in their private and family life and ensure people have agency. However, while we may diverge from the GDPR, learners' rights remain protected under Convention 108 and the European Convention of Human Rights.

Children in Wales[12] and Scotland[13] are further protected under the rights of the child, as set out in the UNCRC, and adopted into domestic law respectively.

5. The UN Convention on the Rights of the Child

> "The realisation of children's rights is not an automatic consequence of economic growth and business enterprises can also negatively impact children's rights."[14]
>
> General comment No. 16 (2013) On State obligations regarding the impact of the business sector on children's rights

[11] The ICO. Right to be informed. https://web.archive.org/web/20230601065403/https://ico.org.uk/for-organisations/uk-gdpr-guidance-and-resources/individual-rights/individual-rights/right-to-be-informed/
[12] The UNCRC in Wales https://www.gov.wales/childrens-rights-in-wales
[13] The UNCRC in Scotland https://www.legislation.gov.uk/asp/2024/1/contents
[14] General comment No. 16 (2013) (1) On State obligations regarding the impact of the business sector on children's rights https://www.refworld.org/legal/general/crc/2013/en/102811

UNCRC Article 16: Every child has the right to privacy

The law should protect the child's private, family and home life, including protecting children from unlawful attacks that harm their reputation.

UNCRC Article 32: Protection from Economic Exploitation

The child's right to protection against economic exploitation is enshrined in Article 32[15] of the United Nations Convention on the Rights of the Child (UNCRC). Although Article 32 is generally interpreted as the child's right to protection against child labour (the right to be protected 'from economic exploitation and *from performing any work that is likely to be hazardous or to interfere with the child's education*') in this contribution we argue that, in the digital environment, children need protection against a myriad of economically exploitative practices.

In its General Comment on the implementation of the rights of the child, the UN Committee on the Rights of the Child found that reaching adolescence can mean exposure to a range of risks, reinforced or exacerbated by the digital environment, including economic exploitation.

6. Rights under the ECHR

Article 8: Privacy of communications

The requirement to surrender sensitive personal data for the purposes of creating AI is an unnecessary and disproportionate interference with the students' right to privacy, enshrined in Article 16 of the UN Convention on the Rights of the Child (UNCRC) and Article 8 of the European Convention on Human Rights (ECHR).

The right to privacy encompasses the right to data protection (United Nations Committee on the Rights of the Child, 2016), and requires that the collection and processing of personal data – for instance to profile a child – should comply with legal requirements. [16]

Handwriting constitutes Personal Data

A handwritten examination script capable of being ascribed to an examination candidate, including any corrections made by examiners that it may contain, constitutes personal data within the meaning of Article 2(a) of Directive 95/46/EC on the protection of individuals with regard to the processing of personal data and on the free movement of such data.[17] Examination scripts or other texts may therefore still fall under UK data protection law and Convention 108, even where names and candidates ID are removed.

---

[15] Van Der Hof, S. et al. (2020) 'The Child's Right to Protection against Economic Exploitation in the Digital World', The International Journal of Children's Rights, 28(4), pp. 833–859. Available at: https://doi.org/10.1163/15718182-28040003.
[16] https://brill.com/view/journals/chil/28/4/article-p833_833.xml#R000070
[17] CURIA Case C434/16 Documents (2017) delivered on 20 July 2017 **Peter Nowak vs Data Protection Commissioner** (Request for a preliminary ruling from the Supreme Court (Ireland)) — Directive 95/46/EC Concept of personal data in handwriting in an exam script https://curia.europa.eu/juris/document/document.jsf;jsessionid=7856630B4E358CC3DFC5DFF838B43F53?text=&docid=193042&pageIndex=0&doclang=EN&mode=lst&dir=&occ=first&part=1&cid=15834577

7. The special case of state education

This consultation inadequate to properly consult the special case of children's and parents' rights and the special environment of state education:

- Non-consensual environment
- Power imbalance that renders consent invalid (GDPR Sweden: regulatory case)[18]
- Children as the rights holder and the overlap with parental rights
- Handwriting constitutes personal data
- Commercial Use and Licensing creates a duty [of the institution] towards the rights-holder for the lifetime of the licence
- Fees systems return to the rights' holder (e.g. the learner, the teacher) not the institution.

Training data might seek to include original work produced by pupils in an education setting. For example, it could include pupils' written work, essays, creative writing, and exam papers. In the UK as in other countries, IP created by learners, including children, is treated in the same way as any other IP today, but it is rarely scrutinised in education.

Non-consensual environment

The educational environment is considered a non-consensual environment in the context of data protection law, particularly under the UK General Data Protection Regulation (UK GDPR) and the Data Protection Act 2018 (DPA 2018). This is because, in many cases, students and sometimes even staff have limited control over whether their personal data is collected, processed, or shared.

The nature of education means data processing is often mandatory, extensive, and beyond the individual's control. This raises ethical and legal concerns around privacy, surveillance, and student autonomy.

Educational institutions cannot lawfully rely on consent as their primary legal basis for processing personal data. Instead, they process data under legal bases such as public task, or legal obligation, and extend some processing to edTech companies as data processors under legitimate interests. The UK GDPR, in Article 6(1)(e), states that public bodies, including schools and universities, can process personal data if it is necessary for them to carry out their official functions. This means that students do not have a genuine choice in whether their personal data is collected, as it is often essential for their participation in education. Parental views rarely override the institution's authority to collect and process data where it is deemed necessary by the setting for routine tasks. Unlike other areas where individuals can choose whether to provide personal data, students do not have the option to withhold this information without significant consequences, such as being unable to complete their studies or access educational resources. This makes objection detrimental and therefore affects the nature of any 'freely given' choice.

[18] Sweden's first GDPR fine was for a school setting processing data on the basis of consent https://www.edpb.europa.eu/news/national-news/2019/facial-recognition-school-renders-swedens-first-gdpr-fine_sv

Power imbalance renders consent invalid (GDPR Sweden: regulatory case)[19]

The Swedish Data Protection Authority (DPA) imposed a fine of 200,000 SEK (approximately 20,000 euros) on a municipality for unlawfully using facial recognition technology to monitor students' attendance in a school in 2019. The DPA determined that the school processed sensitive biometric data without a valid legal basis. The school had relied on consent as the legal basis for processing; however, the DPA found this invalid due to the clear imbalance of power between the students and the school, rendering the consent not freely given.

Children as the rights holder and the overlap with parental rights

Children are recognised as rights holders under both European education laws and human rights law, with their rights often overlapping with those of their parents, particularly in matters concerning education and data protection. The United Nations Convention on the Rights of the Child (UNCRC), establishes that children have independent rights, including the right to privacy (Article 16), the right to express their views (Article 12), and the right to access education (Article 28). In the UDHR the overlap is expressly addressed in Article 26(3) the right to education. In the context of data protection, UK and European law acknowledge that while parents or guardians often exercise rights on behalf of younger children, the best interests of the child must always be the primary consideration in decisions that affect them, as set out in Article 3 of the UNCRC and Article 24 of the EU Charter of Fundamental Rights.

The intersection of children's rights and parental rights is particularly complex in the educational environment, where institutions process large amounts of student data for administrative, safeguarding, and pedagogical purposes not on the basis of consent but that are only intended for their direct care and not for indirect or ‘secondary’ reuses, such as commercial reuse in EdTech apps or AI product development. Schools have legal obligations to protect children's data while also respecting parental involvement, particularly in younger years and schools cannot assume consent or any parental rights on behalf of the child in any routine, everyday circumstances.

Under the UK GDPR and the Data Protection Act 2018, children’s personal data is afforded special protection due to their vulnerability and limited capacity to understand data risks. While parental consent is required for children under 13 in certain contexts, particularly for online services, this does not mean that a child’s data rights are completely overridden by parental authority or that from age 13 a child can “consent” which relies on further conditions such as being fully informed and having capacity to understand what they are giving consent to. For instance, once a child demonstrates sufficient maturity and understanding, they may be able to exercise their own data protection rights independently of their parents, including the right to access, rectify, or erase their data. This principle aligns with the Gillick competence test, which is applied in the UK to assess whether a child under 16 can make decisions without parental involvement. This is known as the principle of evolving capacities, which is recognised in human rights law and acknowledges that children's ability to exercise rights increases as they grow older.

[19] Sweden’s first GDPR fine was for a school setting processing data on the basis of consent https://www.edpb.europa.eu/news/national-news/2019/facial-recognition-school-renders-swedens-first-gdpr-fine_sv

In UK law, the concept of the best interests of the child remains central to balancing these overlapping rights. This means that when conflicts arise between parental authority, institutional policies, and a child's right to privacy, decisions should be made in favour of the child's well-being, autonomy, and long-term interests. The Information Commissioner's Office (ICO) has also highlighted the importance of safeguarding children's privacy in its Children's Code, (the Age Appropriate Design Code)which sets out standards for protecting children's personal data in digital environments, including educational technologies.

Ultimately, the recognition of children as independent rights holders, alongside the continuing role of parental responsibility, creates a legal and ethical framework in which educational institutions must carefully balance privacy, safeguarding, and the rights of both children and parents, always ensuring that the child's best interests take precedence.

It is fundamental for the parent / guardian / carer to be able to understand that they might put their child in a position that might be manifestly disadvantageous. This must also apply to schools in their duties loco parentis and with regard to safeguarding duties. Furthermore, as copyright law recognises the child as an author in the same way an adult is, this should be reflected in their way to exercise their rights. This is not necessarily the case when a young person is represented by a steward. This is established case law as per John and Others v James & Others (the 'Elton John' case) [1991] F.S.R. 397.

Under the DSM Directive, the requirement is that "reproductions and extractions of lawfully accessible works and other subject matter for the purposes of text and data mining" (Art 4(3)). However, case law in the creative industries suggests that release of copyright work ***on behalf of a minor*** may not be a way of making content lawfully accessible. This puts the relevant parties at risk of litigation. [our emphasis]

Exceptions to copyright law are based on international treaties. (see e.g. TRIPs 'Members shall confine limitations or exceptions to exclusive rights to certain special cases which do not conflict with a normal exploitation of the work and do not unreasonably prejudice the legitimate interests of the right holder', Article 13).[20] The approach in international treaties is apparent, e.g. measure and the Marrakesh Treaty for blind, visually impaired and print disabled people to access works protected by copyright.

If the proposed legislation is implemented, would the fact / justification that, 'it will be very challenging to train or develop AI models in the UK using data from the open internet, or source UK-based data from the open internet to train or develop models elsewhere" be compliant with international copyright treaties / TRIPs?

Article 7 demands that, "The protection and enforcement of intellectual property rights should contribute to the promotion of technological innovation and to the transfer and dissemination of technology, to the mutual advantage of producers and users of technological knowledge and in a manner conducive to social and economic welfare, and to **a balance of rights and obligations**." [our emphasis]

Since there must be no unfair contractual obligation that forces students to sign away their rights without proper consideration or negotiation, it would be likely impossible to

[20] https://www.wipo.int/wipolex/en/text/305907#part2

create this in a non-consensual environment and for vulnerable authors who are unlikely to be fully informed of what the risks are for the lifetime of their data and content reuse.

Commercial Use Licensing creates a duty [of the institution] towards the rights-holder for the lifetime of the licence

Under the Copyright, Designs and Patents Act 1988 (CDPA), copyright automatically belongs to the person who creates the work, unless there is a contractual agreement transferring those rights. This applies to essays, research papers, artwork, video projects, software, and other forms of student-generated content.

Ultimately, an institution's licensing duty is a continuous responsibility, not a one-time contractual agreement. The duty extends throughout the entire duration of the licence, requiring active monitoring, renewal management, and enforcement of proper use. Institutions must implement robust copyright policies, ensure clear documentation of licensing terms, and establish compliance mechanisms to prevent unauthorised use, ensuring that content creators' rights are respected and upheld throughout the educational ecosystem.

This would mean that when an institution wishes to reproduce, distribute, adapt, or use a student's work beyond personal assessment purposes—such as for AI production or research—it must ensure that a valid licence agreement is in place and that the learner's rights are respected for the lifetime of that licence on an individual basis towards the rights holder, the learner.

8. Proposed UK changes rely on flawed EU comparison

The Lords Committee in its February 2025 report[21], found:

"Matt Clifford's AI Opportunities Action Plan recommended that the UK reform its text and data mining regime so that it is "at least as competitive as the EU".389 **Our report on the future of news cautioned strongly against "adopting a flawed optout regime comparable to the version operating in the EU".** Witnesses to that inquiry told us the EU's regime lacked transparency about illegal scraping and the use of crawlers, as well as a clear enforcement mechanism for infringements." [176, p.49]

The context of text and data mining in European laws

The UK consultation that closed on February 25, 2025 claimed that the preferred government option brought UK law into alignment with the European Union's legal framework for data and text mining (TDM) which aims to balance a desire for greater use and access to data and content for AI development with copyright protection, with rights holders rights to privacy and protection of IP. We must therefore consider what the EU model is.

The primary regulation governing TDM is the Directive on Copyright in the Digital Single Market (CDSM) (Directive (EU) 2019/790), which introduces key exceptions for text and

[21] House of Lords Communications and Digital Committee Report: 2nd report of session 2024-25 AI and creative technology scale-ups: less talk, more action (2025). https://publications.parliament.uk/pa/ld5901/ldselect/ldcomm/71/71.pdf

data mining under Articles 3 and 4. In EU law these require safeguards not mentioned in the UK proposals.

Article 3: TDM for Scientific Research

Scope and Beneficiaries:

Article 3 establishes a mandatory exception for text and data mining by research organisations and cultural heritage institutions. These beneficiaries are defined as:

- Research organisations: Non-profit entities or institutions designated by a Member State with a public service research mission (Art. 2(1)).
- Cultural heritage institutions: Publicly accessible libraries, museums, archives, and film or audio heritage institutions (Art. 2(3)).

Key Provisions:

- This exception allows acts of reproduction and extraction of copyrighted works and databases for scientific research purposes.
- TDM activities are permitted only if the organisation has lawful access to the material. Lawful access includes:
    - Content accessed through contractual agreements (e.g., subscriptions, open-access licenses).
    - Content freely available online (Recital 14).
- Research organisations can store and retain copies of mined content to ensure reproducibility and verification of research (Art. 3(2)).

Restrictions:

- Rights holders cannot contractually override this exception (Art. 7).
- Rights holders may apply security and integrity measures to protect networks and databases, but these must be strictly necessary for security purposes and not for commercial motives (Art. 3(3), Recital 16).
- No financial compensation is required for rights holders (Recital 17).

Article 4: General TDM Exception

Scope and Beneficiaries: Unlike Article 3, Article 4 provides a broader optional exception that applies to all users, including commercial entities such as AI developers and businesses.

Key Provisions:

- Allows reproduction and extraction of legally accessible works for text and data mining purposes, regardless of commercial intent (Art. 4(1)).
- Mined content can be retained for as long as necessary for mining activities.
- The exception applies to any content that is lawfully accessible, such as:
    - Open-access repositories.
    - Publicly available web pages.
    - Content made available under licenses allowing TDM.

Opt-Out Mechanism:

- Rights holders can opt out by explicitly reserving their rights through machine-readable metadata, such as robots.txt files or embedded digital content signals (Art. 4(3), Recital 18).
- Opt-outs can also be enforced through contractual agreements or unilateral declarations.
- Even where TDM is permitted under Article 4, licensing agreements still play a role. Many content providers impose contractual TDM restrictions, requiring AI developers to negotiate licensing terms.
- The reliance on metadata-based opt-outs (robots.txt, content signals) raises enforceability concerns, as AI developers may still scrape [openly available] content where opt-out signals are absent.

The Three-Step Test and Its Implications for AI Training and Text and Data Mining

The Three-Step Test is a fundamental principle in EU copyright law, ensuring that exceptions and limitations to copyright remain narrowly defined and do not unreasonably harm rights holders.

Understanding the Three-Step Test

The Three-Step Test originates from international copyright agreements, including:

- The **Berne Convention** (Article 9(2))
- The **TRIPS Agreement** (Article 13)
- The **WIPO Copyright Treaty** (Article 10)

While the EU is not a party to the Berne Convention, the CJEU has confirmed that the Union is bound by Articles 1 to 21 of the Berne Convention under Article 1(4) of the WCT, which forms part of the EU legal order[22].

The EU has incorporated this test into multiple legal instruments, including:

- The **InfoSoc Directive** (2001/29/EC), Article 5(5)
- The Computer Programs Directive (2009/24/EC), Article 6(3)
- The Database Directive **(**96/9/EC), Article 6(3)
- The **CDSM Directive** (2019/790), Article 7(2), refers to the test set out in Article 5(5) of the InfoSec Directive, **thereby limiting the introduced**

[22]CJEU, Joined Cases C-403/08 and C-429/08 Football Association Premier League/QC Leisure ECLI:EU:C:2011:631, para 189; CJEU, Case C-277/10 Luksan/van der Let ECLI:EU:C:2012:65, para 59; CJEU, Case C-510/10 DR and TV2 Danmark/NCB ECLI:EU:C:2012:244, para 29.

exceptions, since it is explicitly mentioned that **exceptions and limitations shall only be *applied* when the conditions of the test are fulfilled**.

To be legally valid, any exception to copyright protection must satisfy the following three conditions:

1. Certain Special Cases
    - Exceptions must be well-defined and specific in scope.
    - They must address a clearly identified purpose and not create overly broad or ambiguous rights for users.
2. No Conflict with Normal Exploitation of the Work
    - Exceptions must not interfere with the marketability or commercial value of the protected work.
    - Rights holders should retain the ability to monetize their works through licensing and other means.
3. No Unreasonable Prejudice to Rights Holders
    - Exceptions must not cause excessive harm to rights holders.
    - In cases where rights holders suffer economic harm, compensation mechanisms may be required.

ECJ application of the Three-Step Test

The CJEU has consistently applied the Three-Step Test to ensure that exceptions and limitations to copyright remain strictly confined. The Court first referenced the test in *Laserdisken*, emphasizing that Article 5(5) of the InfoSoc Directive places **strict boundaries** on the system of exceptions[23]. This restrictive approach was reiterated by Advocates General in later cases[24].

In *Infopaq I*, the Court confirmed that exceptions must be interpreted narrowly, particularly when they derogate from the general principle that rightholder authorization is required for reproduction[25]. This strict interpretation was also applied in *Ulmer* and *Stichting Brein*, where the Court ruled that any limitation must be read in light of Article 5(5)[26]. Similarly, in *ACI Adam*, the Court held that copyright exceptions cannot be extended beyond what is explicitly permitted[27].

---

[23] CJEU, Case C-479/04 Laserdisken/Kulturministeriet ECLI:EU:C:2006:549, paras 78-80.
[24] CJEU, Joined Cases C-457/11 to C-460/11 VG Wort/Kyocera ECLI:EU:C:2013:34, Opinion of AG Sharpston, para 68. CJEU, Case C-301/15 Soulier and Doke/Premier ministre ECLI:EU:C:2016:536, Opinion of AG Wathelet, para 29.
[25] CJEU, Case C-5/08 Infopaq International/Danske Dagblades Forening ECLI:EU:C:2009:465, para 56.
[26] CJEU, Case C-117/13 Technische Universität Darmstadt/Ulmer ECLI:EU:C:2014:2196, paras 47-49. CJEU, Case C-527/15 Stichting Brein/Wullems ECLI:EU:C:2017:300, para 62.
[27] CJEU, Case C-435/12 ACI Adam /Stichting de Thuiskopie ECLI:EU:C:2014:254, paras 22-23.

The Court has further clarified that for an exception to apply, it must comply with both its specific legal conditions and the Three-Step Test[28]. In *Infopaq I*, the Advocate General stated that the exception for temporary acts of reproduction under Article 5(1) of the InfoSoc Directive must also meet the additional requirements of the Three-Step Test[29]. This reasoning was reinforced in *Premier League*, where the Court confirmed that fulfilling the conditions of Article 5(1) automatically implies compliance with the Three-Step Test[30]. The same principle was affirmed in *Public Relations*, *Ulmer*, and *VCAST*[31].

Thus, the CJEU has established that all copyright exceptions in the EU legal order are constrained by the Three-Step Test, ensuring that they do not undermine the normal exploitation of works or unreasonably prejudice rightholders.

Case Law

On September 27, 2024, the Regional Court (Landgericht) of Hamburg issued the first ruling on the transposed rules of the directive. The court of first instance dismissed a cease-and-desist claim filed by photographer Robert Kneschke against LAION e. V. regarding the scraping of his photos from a stock photo website for AI training purposes.

The Court found that LAION could rely on the statutory copyright exception under Section 60d of the German Copyright Act—transposing Article 3 of the CDSM Directive—which permits the reproduction of copyrighted content for TDM for **non-commercial** scientific research purposes without the rights holder's consent[32]. While the Court did not rule on the applicability of the TDM exception for commercial purposes under Section 44b—transposing Article 4 of the CDSM Directive—it expressed doubts about whether LAION could invoke this exception for commercial use.

A key issue in the Court's analysis was whether the opt-out in the stock photo website's terms of service constituted a valid reservation of rights under Article 4(3). The Court suggested that, depending on the state of technology at the time of data scraping, an opt-out expressed in natural language could be considered machine-readable if modern natural language processing (NLP) tools were capable of recognizing and interpreting such reservations[33]. This interpretation makes it easier for rights holders to opt out, thereby further restricting the limitations set out in Article 4 of the CDSM Directive.

While this decision can still be contested in higher courts and ultimately reviewed by the CJEU for uniform interpretation, it nevertheless reflects the EU courts' restrictive approach to copyright limitations, consistent with existing CJEU jurisprudence.

---

[28] Infopaq I (n 109) paras 57-58. Premier League (n 94) para 162. Public Relations (n 111) para 23.
[29] CJEU Case C-5/08 Infopaq International/Danske Dagblades Forening ECLI:EU:C:2009:89 Opinion AG Trstenjak, para132
[30] Premier League (n 94) para 181.
[31] CJEU, Case C-360/13 Public Relations Consultants/Newspaper Licensing Agency ECLI:EU:C:2014:1195, para 23. CJEU, Case C-117/13 Technische Universität Darmstadt/Ulmer ECLI:EU:C:2014:2196, paras 56-57. CJEU, Case C-265/16 VCAST/RTI ECLI:EU:C:2017:913, paras 32, 53-54.
[32] European Union Intellectual Property Office summary of the case
[33] Ibid.

9. Government public engagement on AI in education

In August 2024 the Responsible Technology Adoption Unit (RTA) within the Department for Science, Innovation and Technology (DSIT) published research commissioned in partnership with the Department for Education (DfE) to understand how parents and pupils feel about the use of AI tools in education.[34] Parents wanted the opportunity to give explicit consent on whether pupil work can be used. They also wanted pupil data to be de-identified or anonymised.

However, these positions cannot both be met and would yet both be necessary in law. The first, consent, is rarely a lawful valid basis for personal data processing in educational settings today, and would unlikely in future for AI development, If an opt-out and not opt in mechanism is planned then consent is not applicable anyway. Consent must be an active and informed decision and cannot rely on opt-out. The latter, anonymisation, would not be valid to uphold the demands of copyright law with regard to the assertion of moral rights and obligations to be able to maintain the relationship with the author over time in order to manage the ongoing licence obligations.

"5.3.1 There was widespread consensus that work and data should not be used without parents' and/or pupils' explicit agreement."

The use of personal data in relation to AI was also a concern for both parents and children. In particular, concerns involved the sale of data to third parties by companies developing AI tools and misuse of data by other humans (for example, in the creation of deepfakes). These quotes reflect parental concerns about data and privacy, security, and how their children's work and personal information might be used in AI-driven educational tools.

1. **Parent of post-GCSE pupil, Newcastle:** *"Where does it go, where does it stop? Will it always be tagged to you? What about applying to university?"*
2. **Parent of GCSE pupil, Birmingham:** *"There is a sense of big brother about it all. Infant school, they've got your whole life in a data bank, how is that information going to be utilised."*
3. **Parent of post-GCSE pupil, Birmingham:** *"It wasn't very clear about the copyright situation, I think that's a huge thing to know...."*
4. **Parent of post-GCSE pupil, Birmingham:** *"Inaccurate information being fed through the software could be really concerning."*
5. **Parent of GCSE pupil, Birmingham:** *"Data should only be shared with schools, parents and education department."*
6. **Parent of pre-GCSE pupil, Bristol:** *"Thinking about the work... How long will it be kept there - who will it be shared with and how much of my child's personal info is attached to it?"*

[34] DSIT/ DfE research on public attitudes towards the use of AI in education (August 2024) GOV.UK. https://www.gov.uk/government/publications/research-on-parent-and-pupil-attitudes-towards-the-use-of-ai-in-education/research-on-public-a Attitudes-towards-the-use-of-ai-in-education

10. Summary conclusions

Schools are currently trusted. Central to this trust was the widely held perception that schools are not primarily profit-motivated and this would extend to exam boards, although the legal status of the interlinked organisations is opaque to school children, parents and learners of all ages.

The status quo of text and data mining for AI development is not known in the public domain. Half of UK Exam boards refused to say whether they already used children's data and content for training AI, when asked in 2018 about re-using exam papers for machine learning purposes / training AI or similar; or other product development, claiming exemption from Freedom of Information law.[35]

Were the approach in the 2025 UK government copyright consultation to change the law, to go ahead, the proposal is for an opt-***out based licensing.*** Opt-out does not serve the best interests of the child which should default to the highest level of protection, as is the case of the Protection of Freedoms Act 2012 and pupil biometric data in schools. Children without an advocate, or an adult who fully understands the risks, and is able to be the license 'middle-man' to grant or refuse permission would be disadvantaged. An opt-***out*** model is not what parents or children say they want from this engagement. They wanted to be asked for active permission, opt***-in***. However as we addressed above, whether even this 'choice' to opt-*in* could be considered valid on behalf of a child and for how long, in a disempowered nonconsensual environment remains untested in this context though it has been found to be unlawful in the context of data protection law.

Changes to this through the decision to extend commercial exploitation of learners' content and data for the purposes of AI development would damage trust in schools not only in connection with data and content, but the wider relationship between families and institutions. This would not be in the public interest or the best interests of the child, not only for children in school today, but learners who have left in the past and whose data and content is retained in state education or at national level. It may also further jeopardise trust in schools of those parents who may prefer to remove their child from the state system, or indeed in-school education at all.

In our view, the Department should "*publish the evidence of today's data reality before getting ahead of itself with imagined futures. The 2020 DfE ICO audit must be published in full, with a timeline for what remains to be done*.

"And the DfE must commit to giving families control over the current commercial re-uses of their own and their children's information from the millions of named records in the national pupil database, that few know exists."[36]

[35] Exam Boards: FOI including question on AI training https://www.whatdotheyknow.com/info_request_batch/317
[36] Whittaker, F. (2023) *Minister wants education providers to benefit from AI*. FE Week https://feweek.co.uk/minister-wants-education-providers-to-benefit-from-ai-revolution/