CoSTAR
THE UK R&D NETWORK FOR CREATIVE TECHNOLOGY
NationalLAB

TIME TO ACCCT:

Providing Creative Industries and AI Developers with a Copyright Framework of Access, Control, Consent, Compensation and Transparency

NFTS NATIONAL FILM AND TELEVISION SCHOOL
Abertay University
UKRI Arts and Humanities Research Council
UNIVERSITY OF SURREY
ROYAL HOLLOWAY UNIVERSITY OF LONDON
SHERIDANS
DECaDE

Authors:

Prof. James Bennett - Associate Pro Vice Chancellor (Creative Industries) at Royal Holloway, University of London, and Director of the CoSTAR National R&D Lab for Creative Technologies.

Prof. John Collomosse - Professor of Computer Vision and AI, University of Surrey and Director of DECaDE: UKRI/EPSRC Centre for the Decentralised Digital Economy.

Rebecca Gregory-Clarke – Co-Head of Innovation, National Film & Television School and CoSTAR National R&D Lab.

Jack Jones – Partner, Interactive and Creative Media Team, Sheridans.

Dr. Lynn Love - Senior Lecturer in Computer Arts, Abertay University, and Member of the Worldbuilding Team CoSTAR National Lab.

Prof. Mark Lycett - Professor of Digital Innovation, Royal Holloway University of London, and Head of Business Futures for the CoSTAR National Lab.

Prof. Will Saunders MBE - Professor of Creative Industries, Royal Holloway University of London, and Creative Head of the CoSTAR National Lab.

About CoSTAR National R&D Lab

This research was supported by Arts & Humanities Research Council CoSTAR National Lab. Grant Ref: AH/Y001060/1. CoSTAR is a £75.6 million investment by the UKRI Infrastructure Fund, delivered by the Arts and Humanities Research Council (AHRC).

CoSTAR National Lab is the central hub of the CoSTAR Network. This is infrastructure that combines state-of-the-art facilities, technology research and an exciting series of industry and academic partnerships to fuel creative industries growth across the UK. Led by the award-winning StoryFutures team at Royal Holloway, University of London it is delivered via a world-leading industry x research partnership including Pinewood, Disguise, BT, University of Surrey, Abertay University and the National Film and Television School.

The National Lab collaborates, innovates and experiments across the UK's screen and performance sectors, supporting UK creative companies to innovate and grow ethically, sustainably, and inclusively. Their flagship CoSTAR Stage at Pinewood Studios will open in January 2026, placing R&D at the heart of UK's globally recognised home of storytelling.

About DECaDE

DECaDE is the UKRI Next Stage Centre for the Decentralized Digital Economy, a multi-disciplinary research centre led by the University of Surrey in partnership with the University of Edinburgh and the Digital Catapult. DECaDE is a £10 million research centre funded 2020-2026 by industry and the UKRI/EPSRC under grant EP/T022485/1.

DECaDE's mission is to explore how decentralised platforms and data centric technologies such as AI and Distributed Ledger Technology (DLT) can help create value and in our future digital economy in which everyone is a producer and consumer of digital goods and services. DECaDE studies these questions primarily through the lens of the creative industries, which have shifted from monolithic content producers to a decentralised model where individuals and smaller production houses also increasingly produce and consume content disseminated via online platforms. As a multidisciplinary academic research centre, DECaDE brings together technical expertise in AI, DLT and Cyber-security, with business, law, and human factors / design.

About Sheridans

Sheridans is a leading media and technology law firm whose lawyers combine in-depth legal and commercial knowledge with breadth of expertise and experience to give unparalleled advice to their clients.

Since 1956, Sheridans has represented individuals and organisations across a wide variety of media sectors, providing quality advice which is commercially focused and personally delivered. Specialist lawyers in media, entertainment, sport, leisure and technology work closely alongside the firm's corporate finance, employment, real estate, family and dispute resolution groups to provide a truly bespoke service to its clients.

First published May 2025.

Contents

Section 1:
Executive Summary and Recommendations

The brief for this report was to consider the technical and business challenges in developing a publicly available, machine-readable approach to consent that enables copyright holders to either consent or protect their work from text and data mining, whilst allowing AI companies to legally access data. The report sits outside of the TDM (Text and Data Mining) opt-in/opt out government review to provide creative thinking from stakeholders in the creative and technology sectors as to possible solutions. It does not endorse any of the options set out in the review. Instead, we considered what technologies might exist today and those that might emerge tomorrow to build a machine-readable model and examine what a model for wider implementation might involve. As we make clear, the technologies available now – or tomorrow – cannot stand alone, they will need support and change in regulation and in markets.

AI developers in need of data for model training have different options for accessing data, which include using in-house data sets and direct deals with owners of large content archives. The primary focus of this research exercise, however, was to consider the case whereby an AI developer may search for a large volume of assets available online, such as through use of a web crawler. A key assumption in this case is that the volume of assets and different rights holders is so large that manual identification of rights holders, direct negotiations, issuing of licences, contracts and payments is simply not feasible. This is therefore the primary driver behind the framework outlined – which aims to begin mapping out how a machine-readable solution may be constructed. The framework presented in this document is a starting point which could allow for permissions to be asserted, authors to be credited, and value to be created. However, not all of the technologies required are yet at maturity, and so we present some key recommendations for the further work required to realise a solution of this nature. Finally, we note that this framework is not designed to cover all situations all of the time. It is a framework that should be available to everybody, but it does not seek to intervene with direct deals, nor does it assert a default that rights holders or AI developers are compelled to use: some rights holders may deem their works too valuable or complex to ever put into a Generative AI marketplace, and that is a form of control that should be supported in parallel.

The era of Generative AI presents a complicated opportunity and challenge for the UK's current approach to copyright. Whist Generative AI as a technology is new, the conundrum itself is not: balancing freedoms and control, consent and protection, incentives and penalties have always been at the core of copyright considerations.

The threat to the decades' old balancing act is, however, existential. It is not simply how the balance is weighted but the very nature of the balancing act itself. This balancing act now needs to be reconfigured around a new set of priorities that we call ACCCT:

Access: providing AI developers with high quality data sets and all stakeholders with accessible tools to enact the framework.

Consent: the ability for rights holders to fundamentally permit or prevent use of their work and specify for what purposes, how and where these data are used.

Control: the ability for rightsholders/holder to manage the terms under which rights are granted for generative AI use at a level of granularity that reflects the complexity of licensing and ownership, such as the multiple parties that may be involved and their conditions of use.

Compensation: the need to establish fair, equitable and easily processed financial and non-financial recompense for the use of data, underpinned by a shared understanding of value in the generative AI ecosystem.

Transparency: the ability for all stakeholders to understand clearly and easily what data is available for use in generative AI, what licensing terms are fair and equitable and what it is/isn't possible to do with data once it is has been used in generative AI training, fine-tuning or adaptation.

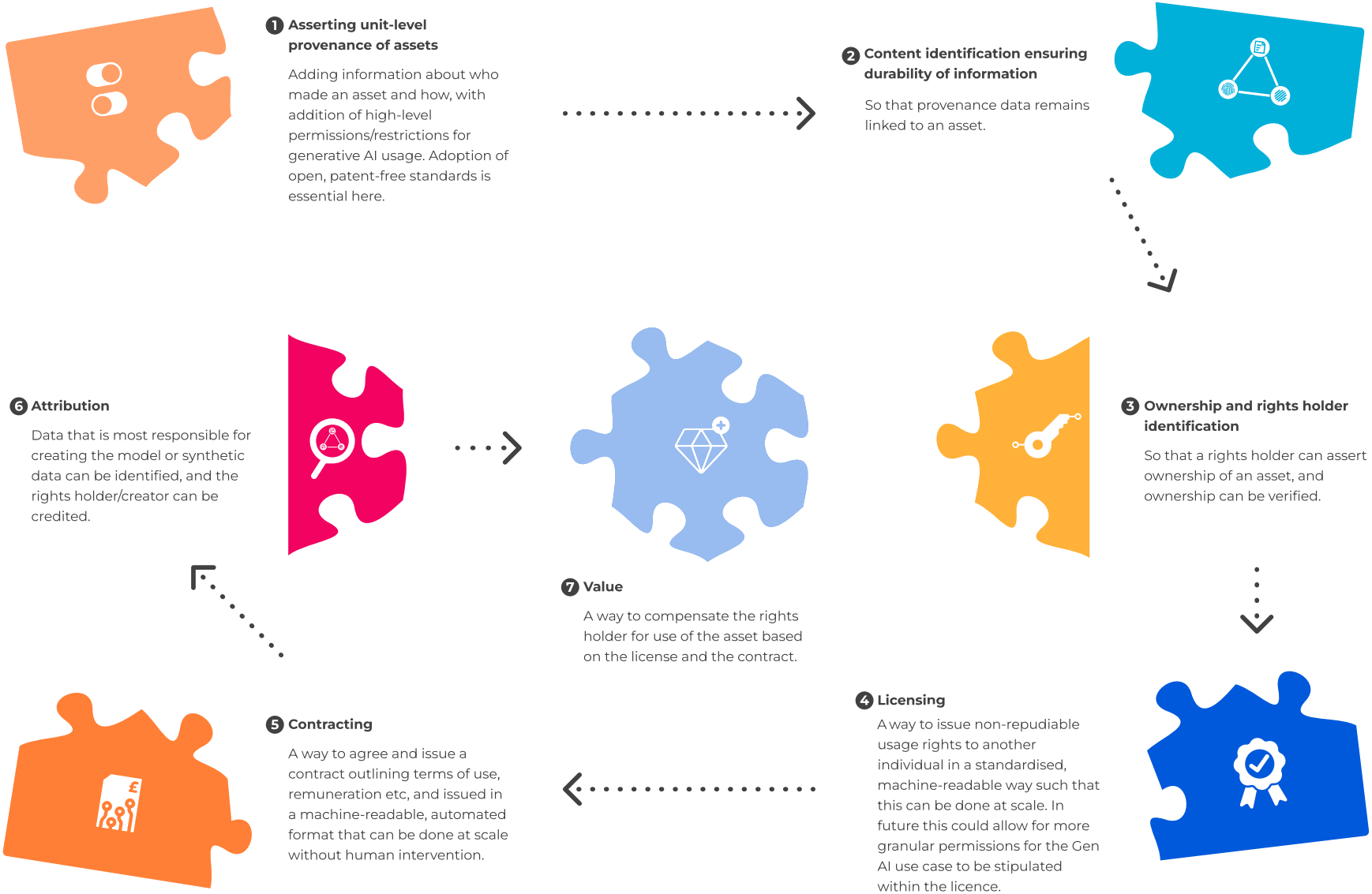
To create a new balance the UK will need to address and adopt new standards, develop technologies to implement and iterate those standards, revise and renew its legal and regulatory framework, and establish the right environment to foster both creative and technological innovation. Emerging provenance technologies exist today that can form a base for an AI and copyright framework to build on and produce a fair and equitable ecosystem [DECaDE, 2025].

The time to ACCCT is now: the framework proposed here is not a complete solution but it represents a set of first steps to a workable solution, underpinned by a set of proven and emerging standards and technologies, coupled with an innovative approach to regulation that has the potential to position the UK as a world leader in ethical generative AI. These first steps are significant: inaction at a time when international markets adopt a laissez faire approach will only harm the UK. We believe there is an essential role for government to take in adopting and promoting the ACCCT framework as it creates an incentive for technology companies to work with rights holders and build solutions that work: whilst the UK does nothing or continues to contemplate an exemption, AI developers may stand back from the UK leaving the nation behind in terms of both protection and competition. Through ACCCT we can begin the process of machine-readable solutions through the ability to assert permissions/restrictions and realise attribution. From here we can begin to tackle the more complex problems that remain ahead.

As we make clear in the limitations discussion (section 5), ACCCT is not a silver bullet, and it cannot solve the myriad of problems and concerns that stakeholders bring to the table. In the age of generative AI, the work involved in protecting and controlling rights necessarily moves further up the copyright value chain, requiring all stakeholders to play a role in this emerging ecosystem: there is much work to be done on knowledge sharing, disseminating and incentivising behaviour changes for all parties. Provenance technologies exist today that allow us to take first steps in establishing a new copyright balance, but a range of urgent and parallel work will need to be undertaken to accelerate the technologies of tomorrow. This includes awareness-raising alongside deeper R&D on unlearning in generative AI models and attribution algorithms to better identify the use of resulting work from Generative AI, with the issue of cascading and composite rights a particularly pressing area for further work. We point to ways forward on these areas whilst also proposing the ACCCT framework that allows significant and tangible steps to be taken now that that will enable this Government to continue to position the UK as a world-leading creative and innovation hub at the forefront of ethical AI development.

Informed by consultation with rights holders and developers working with generative AI and prior work on provenance for decentralised content rights and attribution [Balan et al, 2023, Balan et al, 2025], the ACCCT framework seeks to improve access, consent, control, compensation and transparency in seven steps (Figure 1):

1. Asserting unit-level provenance of assets
2. Content identification ensuring durability of information
3. Ownership
4. Licensing
5. Contracting
6. Attribution
7. Creating value



1 Asserting unit-level provenance of assets

Adding information about who made an asset and how, with addition of high-level permissions/restrictions for generative AI usage. Adoption of open, patent-free standards is essential here.

2 Content identification ensuring durability of information

So that provenance data remains linked to an asset.

3 Ownership and rights holder identification

So that a rights holder can assert ownership of an asset, and ownership can be verified.

4 Licensing

A way to issue non-repudiable usage rights to another individual in a standardised, machine-readable way such that this can be done at scale. In future this could allow for more granular permissions for the Gen AI use case to be stipulated within the licence.

5 Contracting

A way to agree and issue a contract outlining terms of use, remuneration etc, and issued in a machine-readable, automated format that can be done at scale without human intervention.

6 Attribution

Data that is most responsible for creating the model or synthetic data can be identified, and the rights holder/creator can be credited.

7 Value

A way to compensate the rights holder for use of the asset based on the license and the contract.

Figure 1: The seven jigsaw pieces of the ACCCT Framework

RECOMMENDATIONS

1. Government backing of open standards for provenance information ensures compliance by requiring consistent implementation of a standardised approach across platforms and ecosystems, with consistent support for image, video, audio and text modalities, enabling creators to seek legal remedies if their rights are violated.

2. Awareness raising and upskilling initiatives that lower the barrier to entry for content creators and rights holders to access current and upcoming technologies for 'open rights standards' that reflect the nuances of their domain and ultimately help to make licencing a more machine-readable process.

3. Support for further research activity is required to trial and scale systems for asserting, ascertaining and verifying ownership, in conjunction with provenance technologies. This should build upon existing work in the area (as outlined above) and include research into the practicality and usability of systems as well as technical solutions.

4. Trialling a non-mandatory Government-backed and verified database or 'forum' to enable rights holders to list their assets and AI developers (or others) to understand where valuable data sets may be found and who owns them.

5. Continued government backing for the 'open rights standard', which will drive adoption and ensure compliance as the standard gains maturity.

6. Create a bank of agreement templates that provide a UK-wide infrastructure for all players to access transparency of terms, similar to the Lambert and Brunswick templates made available through the IPO. This would include access to the technological forms of the contract for attribution – such as the use of DLT Smart Contracts, or more traditional e-signature systems, depending on the scale of data/agreement being executed. Such templates could form the basis for negotiations between parties; however, use or adaptation of these contract templates ultimately remains optional.

7. To provide support, such as targeted research and innovation funding, for the development and scaling of attribution algorithms which enable rights holders and AI developers to transparently link outputs from generative AI inputs.

8. Support for the research and testing of new forms of business marketplaces and payment models that are effective, scalable and incentivise the creator economy as well as AI developer investment.

9. To enable functioning marketplaces to develop, government should work with industry to co-design an approach to regulating marketplaces, such as an industry oversight body to promote best practice and enforce compliance: for example, an independent legal ombudsman, or a non-departmental public body which reports to Parliament, such as the Information Commissioner's Office (ICO).

10. Develop a ‘charter’ of best practice principles to which AI developers can sign up that enables AI developers to achieve a ‘Kitemark’ certification, including: fair and transparent attribution practice; terms of trade for compensation; use of appropriate open standards to allow content provenance, rights information, and appropriate permissions; and transparency of high-level information about what data will and will not be used for. To support this initially through incentivisation, such as mandating that only Kitemark approved AI developers can win government contracts, before adding further regulatory teeth such as through an industry oversight body.

11. To provide support to urgent and underlying socio-technological challenges, enabling emerging technologies to scale rapidly. In particular, support should be given to address the linked problems of generative AI models unlearning where copyright permissions have not been given and/or withdrawn, and attribution algorithms which enable rights holders and AI developers to transparently link outputs from generative AI inputs. Whilst prototype technologies in these spaces exist, they currently don’t scale with accuracy and raise important associated legal and social questions that need tackling hand-in-hand.

12. Undertake a series of scaling prototype and demonstrator projects that enable the technical, legal and social solutions proposed in ACCCT to rapidly progress. The need is for safe and supported trials to test technologies and legal solutions that can then be implemented simultaneously to protect all parties and enable a vital, innovative creative ecosystem.

METHODS: WORKSHOPS, PERSONAS AND BACKGROUND RESEARCH

The CoSTAR National R&D Lab for Creative Industries, with the support of over 5 years of content provenance research from the DECaDE (UKRI Centre for the Decentralised Digital Economy) team and expert legal practice from Sheridans, convened a workshop of leading creative, rights holders, legal experts and technology companies to develop pathways to address the challenge of consent and control posed by the UK Government Copyright and AI Consultation.

As an alternative, the intention was to chart a collaborative and practical way forward to implement a workable framework for copyright and generative AI that places access, control, consent, compensation, and transparency at its core.

Prior to the workshop, 8 'user persona' interviews were conducted to provide a range of perspectives on how different content types, rights frameworks, industry practices and generative AI use cases might inform the requirements of any framework (see section 3). Within the workshop, discussion focussed on three key areas:

- Understanding user perspectives and requirements
- Mapping out the core functional requirements of a future framework
- Exploring the steps required to realise such a framework, along with potential ways of advancing and supporting its development

The participants worked in three groups whose discussions were captured by dedicated note-takers and through video recording of each group's final summary presentation. The data were coded and analysed to inform the development of what became the ACCCT framework, supported by desk research and expert consultation to determine the current state of the art of each step.

In presenting the framework, we also outline its limitations, acknowledging that there were some aspects that were important to participants but that require further research and development beyond what was possible in the workshop.

The workshop was conducted under Chatham House Rules, with 24 companies and organisations taking part, representing large and small companies, rights holder organisations, distributors, producers and AI developers from across creative sub-sectors and content types. Their participation in the workshop was without prejudice to their separate submissions to the AI copyright review itself.

We are grateful for their time, insight and willingness to engage in ongoing dialogue about how to produce a framework that enables inclusive and equitable innovation.

Section 2: Legal Context

When looking to establish a viable framework for rights management with respect to AI development, the balance to be struck is between trust and protection for rights holders and creatives on the one hand and certainty for AI developers on the other: that is, trust from rights holders that this technology accurately records and tracks the rightsholders/holder assets, and certainty for developers to know what they can and cannot do with those assets.

To provide the kind of granular level of control stakeholders from creative industries and AI developer communities desire, there needs to be a reliable and accessible set of technologies. Moreover, these technologies need to be simultaneous with changes in the law.

There is a lesson to be learnt from the EU, which some critics have described as an example of the "tail wagging the dog": legislation has been introduced without sufficient prior due diligence on the practical implications or required technologies to enact it. As a result, disagreement at national court level persists regarding what are universally accepted ways of opting in or out are. Such a situation ultimately will not build trust or certainty for rightsholders/holder or developers.

In the UK, the key point at issue stems from the exception laid out in section 29A of the *Copyright, Designs and Patents Act*, which currently only applies to lawfully accessed data used for non-commercial research purposes. This exception provides little clarity to AI developers or rights holders given the commercial uses that the data-trained models are often put to. Web crawling is not new, but what is new is its application on the broad scale that we are currently witnessing for generative AI use cases. However, the legal principles remain the same. This is not an AI-specific problem; it applies to any rights holder whose works are available online.

So, whilst s29A Copyright, Designs and Patents Act is at the core of this matter, this is not just a copyright issue. This framework and the technologies underpinning it concern all intellectual property and related rights, including trademarks, design rights, personal data and image / personality rights. What those in the workshop and more widely in the creative sector need is an approach that is neither copyright-specific nor AI-specific. It must be capable of tracking and protecting assets, their complex underlying rights, as well as the potential to protect derivative rights. A technology solution that proves attribution and ownership, not just with respect to data sets for AI-training purposes but for any use online, is something that should be welcomed by rights holders.

The workshop and our associated work have made clear that rights holders are not averse to leveraging this value, and there has recently been an uptick in licensing deals between rights holders and AI developers (Kretschmer et al 2024). But there has also been a significant increase in legal claims against AI developers, with 40+ in action globally at the time of writing. In the US, fair use and copyright exceptions present grey areas for the use of copyright material in generative AI training and are just starting to see pushback in legal settings (e.g. Thomson Reuters vs ROSS).

To reduce uncertainty and the number of legal claims, the UK needs to develop technologies hand-in-hand with changes in the law that are backed by a regulatory environment that is demonstrably independent of political influence. The ACCCT framework outlines the first steps in this respect, enabling the Government to continue to position the UK as a world-leading creative hub and innovator that is at the forefront of ethical AI development.

Section 3:
Mapping the Problem Space

Generative AI in the context of copyright presents challenges to stakeholders across the breadth of the creative industries, calling into question existing approaches to asset creation, ownership, use and protection in online spaces. These challenges, in part, are due to the diverse and complex nature of the problem space. At its simplest, it is the point where four key factors intersect: users, content forms, rights frameworks and generative AI use cases (see Figure 2). However, the variables that exist within each of these four factors mean that there are countless combinations of users, content, rights and use cases to consider when seeking routes towards a 'solution' to the problem.

To better define the problem space each of the four factors will be discussed in turn to illustrate the complexity of the landscape and begin to draw out key themes. However, it should be noted that this section cannot be viewed as an exhaustive mapping of the space, given the limited nature of this research exercise.

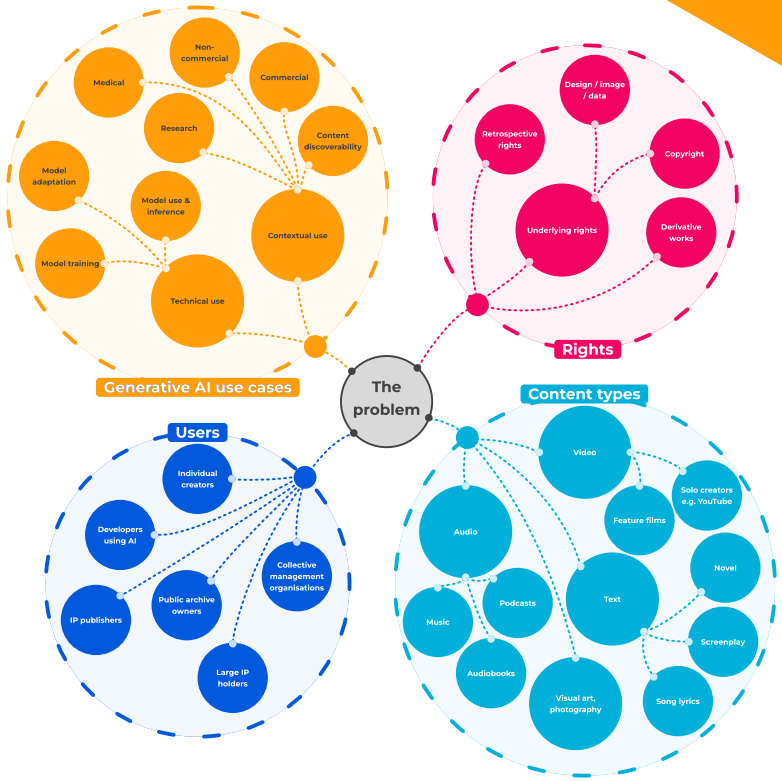


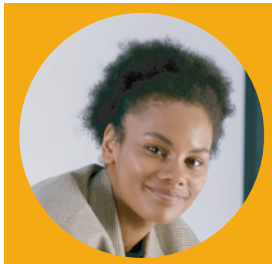
Figure 2: Mapping the generative AI x Copyright Problem Space: a galaxy of users, rights processes, content forms and use cases come together.

"In a world in which you've sort of generally had contracts with publishers, TV and film producers for known things... the book on the shelves or the TV show in the home or in the cinema. You kind of know what you're doing there, and you know what the output is. With AI changing always, presumably you don't quite know what you're signing up for."

Workshop participant

USERS

Ultimately the ACCCT framework needs to work for everybody involved in the creation, protection or promotion of copyrighted works or who may be involved in advocating for rights holders or is involved in using assets in generative AI development. Users vary in scale of operation, reputation, influence, content type and relationship to the content. With this variation comes differing perspectives and language around asset protection, licensing, attribution and royalties. This means that, within the user group, shared language and understanding of each field's existing practices is the first barrier to overcome before moving towards problem solving.



Ariya F.

Audio Visual Content Publisher

Aryia F. works for an organisation with 100 employees that publishes episodic and film content across multiple online and terrestrial platforms. They have a business law background and are responsible for managing relationships across the copyright landscape for AV rights of the content they purchase and publish.

The organisation is nervous about innovations in Generative AI and does not use this technology in its core business. It does keep up to date with developing policy and legislation and is exploring its options in relation to ownership and consent.

The Problem

The organisation distributes audio visual content for which it must cascade royalties across rights holders of a range of media. Whilst systems are in place for supporting the delivery of royalties for audio visual content, the use of content created by the organisation for training of foundation models is opaque. Ariya F. is frustrated by the lack of mechanisms for attribution where content is used for training and that there is no clear chain for royalties to transfer to rights holders.

AI Copyright System Requirements

- Disclosure and attribution of content that was used to train foundation models
- Being able to trace secondary IP that exists within copyright material (e.g. a music track within an episodic drama), the contribution of this IP to model training and resulting output' and the ability for these cascaded rights to be recognised and attributed
- Mechanisms for remuneration for all royalties, secondary royalties etc. to all parties involved handled automatically by an independent body

Motivations

- Maintaining ethical and responsible practice by ensuring royalties are delivered across the IP chain
- Maximising content reach across online and terrestrial platforms
- Maintaining a talent ecosystem that ensures high quality content continues to be produced and can be bought for publishing

Frustrations

- It is not possible to trace the contribution of copyright material to foundation model training and the resulting output
- There is limited attribution of rights holders in the role of the organisation's content in foundation model training
- Mechanisms for royalty transfer to publisher and across the chain of rights holders do not exist

Image by Mikhail Nilov

Figure 3: Example of user personas based upon interviews with practitioners, rights holders, legal teams and industry consultants. See Appendix for full set of personas.

The ultimate opportunity for AI developers working in the UK, and for creative content owners to collaborate in order to realise further value from their work is clear.

"We have very intense competition from US and Chinese companies and others around the world. And so the opportunity is that if we are able to maintain our leadership and we can become a generational company in the UK, we want to stay in the UK, we want to grow here... that's our ambition to become the Open AI of the video space but we want to do it ina responsible and ethical manner."

Developer working with generative AI

However, all users discussed their concerns and support needed in order to achieve this. Three key themes that were surfaced during the process were around transparency, granularity and scale.

"The challenge is having no transparency about what's gone into these programmes, having no control over what or how those works have ended up in the data sets and having no transparency about opting out."

Workshop participant

In terms of transparency, there is a need to be able to give informed consent, and for clear guidance that is accessible to all. The need for granularity was discussed at length, highlighting the need for a system that can cope with the nuance of differing AI use cases as well as complex flows of rights for different forms of content. For all users, an accessible and scalable system is paramount – reflecting the quantities of assets required for this purpose. A detailed list of user requirements gathered during the exercise is given at the end of this section.

It is also important to recognise the role and needs of users who were not 'in the room' but which are an important part of this ecosystem – such as those within the Creator Economy.

"We can go to the large content marketplaces of the world and pay for data sets. But we would also find it very useful, for example, for the millions of YouTube or TikTok creators, who are uploading content every day, to have access to an open standard they can opt in to so their content is available for training, and in return they can get paid based on a licence model. That would be very useful and very powerful."

Developer working with generative AI

CONTENT FORMS

The range of media types that can be processed and inform generative AI foundation models introduce further complex variables into the problem space. Text, image, video, video game and audio-based assets all need to be considered, all of which have their own variations in form, content and layers of embedded assets that cascade the complexities of content definition and ownership.

Text based assets, as an example, might take the form of a novel, screenplay or news article, which are diverse in form and content within and of themselves, and have different stakeholders, online formats and rights frameworks attached to them. Novels and news may also have layered content, embedding images and/or video content that was produced by and owned by others, complicating the forms of media involved and the associated rights.

Photography presents another illustration of the nuance involved in making decisions about consent, which may depend on the type of photography and the context of use.

> **"[For photographers] Being able to say yes or no to usage [for generative AI] is important for moral rights into how that that photograph might be used by a third party. It's always been, in some respects a sensitive area depending on what type of photography you do. On the one end of the spectrum, very commercial means you're talking about trademarks versus the other end of the spectrum, where you're talking documentary photography where the subject and the person in the photograph may be depicted inaccurately or offensively."**
>
> *Workshop participant*

Another important angle raised by participants is the difference in existence and uptake of relevant technologies between industries and content forms. As we will explore further in Section 4, the addition of metadata to assets is a key component in many proposed approaches to this problem; however, the prevalence, accuracy and effectiveness of metadata amongst different forms of assets varies greatly.[1] As one participant from the music industry points out, "*the majority of metadata or data marking ... simply is not yet present for music that is available through digital delivery services or on the internet in general*".

The workshop made clear that one size does not fit all in this space. Creative industries stakeholders shared a concern to develop a framework that could provide control and consent for content-specific production conventions and rights (see more detail on rights considerations below). The workshop also highlighted a need to consider how tailored support may be required for different parts of the creative industries, who are working with diverse and at times interconnected forms of content, to ensure that a future solution can be fit for purpose across the breadth of the creative industries.

[1] It is widely acknowledged in the music industry that poor use and tagging of metadata is a reason that artists are not frequently credited and remunerated. In fact AI could be a solution for this if it can improve metadata tagging through Open Standards.

RIGHTS

Assets in the creative industries tend to contain multiple underlying intellectual rights either assigned or licensed by third parties. The underlying flow of rights (including IP) in creative works can considerably vary from industry to industry.

Music presents a compelling example of this complexity: one pop song released today will have, on average, contributions from a producer, multiple songwriters, composers and lyricists, one or more artists, and multiple session musicians. This leads to rights ownership and responsibilities sitting across several different entities: producers, labels, publishers, and a Collective Management Organisation to administer one or more sets of Copyright (e.g. Communication to the Public or Mechanical rights). If another production (such as a film) went on to license this song, this adds yet another layer of complexity.

To take video content as another example, there is a stark difference between a video asset made by a single creator, who might own the asset outright, and a feature film, which includes complex assets and associated rights: actors, whose image is embedded within the film, speak dialogue, written by a screenwriter, with a score containing both original compositions and licenced music. The way audio visual content is published also then needs to be considered.

> **"When we publish and distribute, we sell a package of material to platforms, so to Sky or to YouTube, and we act as a publisher. So in that sense, we have a cascaded IP that comes from the producer to us to then the platform and we just need to trace all of that through the chain... That applies to content and music."**
>
> *Audio visual content publisher*

The above examples illustrate why a binary opt in/out solution is extremely difficult to apply in many cases. The complexity of rights within the asset itself, paired with the diverse and varied conventions of differing rights processes across media forms, adds to the difficulty of asserting ownership and protecting assets in online spaces. IP, therefore, should be viewed as organised into "bundles of rights" that require a high degree of granularity. A framework that seeks to promote transparency and consent in this context would need to be able to identify, track and attribute across these bundles and their use in different creative sectors. As discussed previously, the legal landscape for protecting rights for generative AI is nascent and needs accelerating to develop in tandem with the technological solutions emerging. Ultimately all stakeholders, whether creative rights holders or AI developers, were motivated to build a solution that made copyrighted work accessible on clear and transparent terms – without need to have recourse to legal battles.

> **"If something is stolen or illegally trained, it shouldn't require a team of lawyers and a lot of software to be able to declare that's been stolen."**
>
> *Content publisher*

> **"What you'll find is that there are significantly more infringements online than you have the opportunity to manage. So with generative AI appearing on the scene, what that has done is only compounded the issue from a number of different perspectives."**
>
> *Workshop participant*

GENERATIVE AI USE CASES

In our mapping of the problem space, generative AI use cases are considered in terms of technical and contextual use.

Technical use differentiates between different ways that data are used in AI development. The phrase 'training' is often used in a generic sense; however, there are important distinctions which impact the way we might find, use and value data:

- **Model training** is the process of passing vast amounts of training data through an AI algorithm to produce an AI model, such as a foundation model.
- **Model fine tuning** (also known as adaptation) is distinct from model training, and describes the process of carefully selecting a smaller set of training data which is then used to adapt an existing AI model to train a new, more specialized model. Fine tuning thus makes a more generic model better for specific applications, or able to create personalized or targeted content.
- **Model inference** is distinct from training, and describes the process of applying the trained model to some tasks, such as generating content. Inference uses very small amounts of data (for example text prompts) to produce an output from an AI model (such as a video).

The core behaviour of a model (for example its reasoning capability) and world knowledge are determined primarily by its training data, and typically by the volume, diversity and quality of that training data. The knowledge, and to a lesser degree core capability, may also be adapted by fine tuning the model post-training. Additional knowledge may be drawn in from external sources at the time of inference. Depending on the use case, the data used to specialise a model at adaptation or inference time may have greater influence on the kind of output that an AI model can generate than any given piece of data contained within a vast data set used for training. This also further illustrates the need for a solution which can recognise this granularity of different use cases.

> **"We are interested in buying [data to train]. It's just that the pricing structure – a lot of these traditional broadcasters are trying to apply... traditional sort of licencing agreements to this kind of new space and they just don't work."**
>
> *Developer working with generative AI*

> **"We can pay you millions of pounds for fine tuning, but [we need] deep discounts on the pre-training as we need tens of thousands of hours or hundreds of thousands of hours [of content]."**
>
> *Developer working with generative AI*

Contextual use on the other hand, relates to the purpose of training – what the system is learning to be able to do. For example, workshop participants with rights responsibilities were open to their assets being used, for free, for educational purposes, research or non-commercial use. In a commercial context, rights holders may need to be able to stipulate that certain use cases are either permitted or restricted due to ethical or business conflicts. The problem is a lack of clarity in what legal consent is given, with commercial use acting as a focal point for participants' legal and ethical concerns due to the potential threat to their business models.

This was one of the most significant areas of granularity in consent required by all stakeholders beyond a binary opt-in/opt-out model. Any technological system must be able to easily define purpose of use and signal licensing options as a result. A related significant concern was to ensure that such granularity of control should be commensurate with the labour and value involved: rights holders wanted to avoid never-ending decision trees for consenting to a range of purposes whilst AI developers wanted certainty that decisions were not transient, with consent liable to be withdrawn without notice.

"I think granularity in terms of use needs to be easy to control, at scale, because I cannot give you a kind of blanket answer for [all writers] about using their data for training or not training or fine tuning or not fine tuning. So I think a system [should have] that granularity, but can be operated simply, through sliders or that sort of thing [to define varying levels of consent]"

Workshop participant

GATHERING USER REQUIREMENTS

The problem space is where the complex and varying needs of users, content forms, rights processes and generative AI use cases meet. The following requirements, derived from our persona interviews and workshop activity, are the basis on which we begin to build a framework that seeks to create a landscape for generative AI centred around access, consent, control, compensation, and transparency. However, it should be noted that the first iteration of the ACCCT framework does not aim to meet and resolve every requirement. Limitations and further work required are addressed in Section 5.

Content creators/rights holders identified the need to:

i. assert provenance (who created an asset) in an online space;

ii. assert ownership over their content in online space;

iii. understand if their assets have been used to train a model;

iv. understand the purpose of use of their data in generative AI training to allow informed consent to be given;

v. provide different forms of consent for the use of their assets for different generative AI use cases;

vi. have a scalable system that can cope with the complexities of different flows of rights that vary by content form and industry;

vii. ensure their contribution is recognised and credited (attribution);

viii. realise value from use of their assets and a clear mechanism for compensation where assets are used for a commercial purpose;

ix. have access to low cost/resource solutions that lower the barrier for copyright holders to protect their work;

x. have access to a system that is practical and can be used at scale – coping with large collections of content that exist in different forms online;

xi. have access to system that ensures compliance is quick, fair, easy and equitable to remedy.

AI developers identified the need for:

xii. clarity on the rights and permissions on which they are relying, such that they can obtain lawful access to high value data sets;

xiii. reassurance that decisions are not transient, with consent liable to be withdrawn without notice;

xiv. new business models and concepts around compensation which are reflective of the value of the data, and ensure that working with generative AI remains affordable;

xv. standardised systems that work at the scale required for working with generative AI;

xvi. incentives and support to engage in ethical legal practice beyond 'doing the right thing' which does not yet have concrete guidance.

All users shared requirements for:

xvii. an approach to ecosystem support and development that recognises the work and livelihood of creators, AI developers and rights holders at varying scale of operation, from independent through to large organisations;

xviii. clearer consensus and guidance on best practice;

xix. increased awareness and literacy around potential solutions and their use;

xx. accessible, fit for purpose legal frameworks and supporting policy.

These 20 requirements are a baseline that is generally shared across the stakeholders involved, but the shape and value place on each requirement varies from player to player and sector to sector. Whilst recognising that a tight one-size-fits-all approach is not desirable, ACCCT attempts to lay out an approach that can work for large and small players across different sectors.

As we set out in Section 5, there are important user requirements that are beyond the initial steps proposed by the ACCCT framework. Further work in all these areas is needed to better develop and test technological solutions that can feasibly facilitate better licensing arrangements. The experiments proposed in Section 6 provide a way to move some of these areas forward:

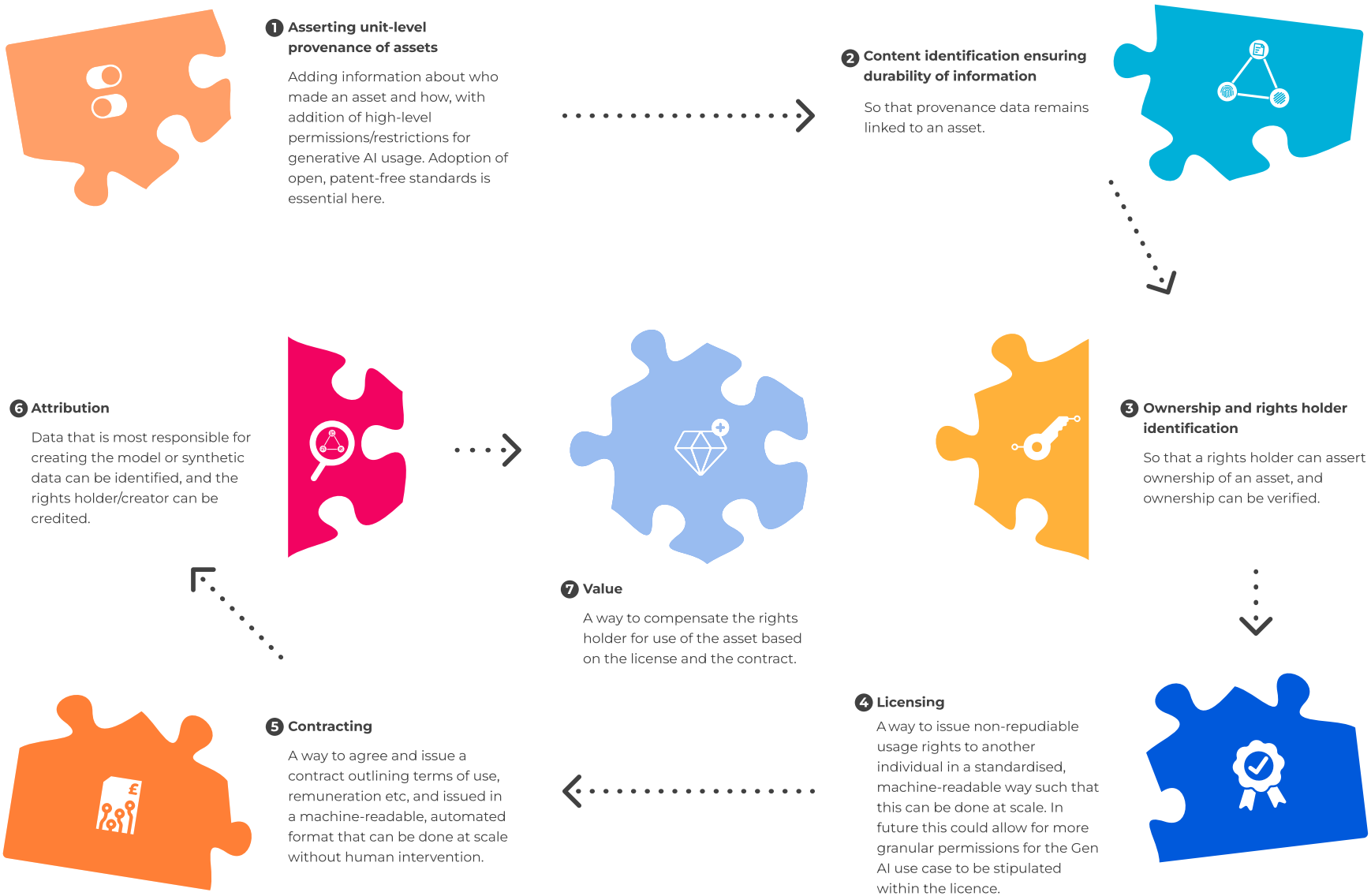
- **Composite, cascading and derivative rights:** These aspects are particularly complex with many content types made up of a patchwork of interlocking IP rights.
- **Unlearning:** There is currently no reliable process for verifiably removing the influence of distinct assets on a trained generative AI foundation model, (e.g.[Du et al, 2024, Hu et al, 2024].
- **Rights transience:** Linked to the above, there is a need to develop practical and pragmatic approaches to time-limiting any licences granted for generative AI.
- **Retrospective rights:** The ACCCT framework is limited in this regard as it focuses on resolving transparency issues going forward; current legal recourse for infringement would remain in place.
- **Notification and transparency:** Neither ACCCT requires an agreed mechanism for providing informed consent to enhance transparency where the asset owner is informed about intended and actual use of the content.
- **Digital Rights Management:** ACCCT, nor the technologies proposed, are not a digital rights management solution. Neither ACCCT nor a Digital Rights Management (DRM) solution are an absolute guarantee against fraudulent behaviour.
- **Downstream copies:** Globally accepted standards are required to ensure good actors retain this information in the assets during processing, e.g. [C2PA, 2024].
- **Invisible content:** Technical frameworks and code, which underpin much of the video game and interactive arena, present unique and more detailed challenges that need rapid testing with the technologies proposed in ACCCT.

Section 4:
Building an ACCCT Framework

A future solution that meets all the requirements of the various users and use cases, which we began to explore in Section 3, will not be achieved through a single technology, standard or legal intervention. Rather, there is likely to be a combination of technical approaches to different parts of the problem, underpinned by adoption of open standards and clear policy and legal guidance alongside best practice.

In this section, we present a framework that we believe has the potential to benefit content creators, rights owners and developers working with AI equally. This framework is not intended as a system architecture diagram; rather, it identifies a set of essential building blocks for a process allowing for:

1. **Asserting unit-level provenance of assets –** adding information about who made an asset and how, with addition of permissions/restrictions for generative AI usage. Adoption of open, patent-free standards is essential here.
2. **Content identification ensuring durability of information –** so that provenance data remains linked to an asset.
3. **Ownership and rights holder identification –** so that a rights holder can assert ownership of an asset, and ownership can be verified.
4. **Licensing –** a way to issue usage rights to another individual in a standardised, machine-readable way, such that this can be done at scale. This could allow for more granular permissions for the Generative AI use case to be stipulated within the licence at the same time as ensuring usage rights are not withdrawn at a moment's notice.
5. **Contracting –** a way to agree and issue a contract outlining terms of use, remuneration etc., issued in a machine-readable, automated format that can be done at scale without human intervention.
6. **Attribution –** data that are most responsible for creating the model or synthetic data can be identified, and the rights holder/creator can be credited.
7. **Value creation –** a way to compensate the rights holder for use of the asset based on the license and the contract.



1 Asserting unit-level provenance of assets

Adding information about who made an asset and how, with addition of high-level permissions/restrictions for generative AI usage. Adoption of open, patent-free standards is essential here.

2 Content identification ensuring durability of information

So that provenance data remains linked to an asset.

3 Ownership and rights holder identification

So that a rights holder can assert ownership of an asset, and ownership can be verified.

4 Licensing

A way to issue non-repudiable usage rights to another individual in a standardised, machine-readable way such that this can be done at scale. In future this could allow for more granular permissions for the Gen AI use case to be stipulated within the licence.

5 Contracting

A way to agree and issue a contract outlining terms of use, remuneration etc, and issued in a machine-readable, automated format that can be done at scale without human intervention.

6 Attribution

Data that is most responsible for creating the model or synthetic data can be identified, and the rights holder/creator can be credited.

7 Value

A way to compensate the rights holder for use of the asset based on the license and the contract.

Figure 4: The seven building blocks of the ACCCT Framework

The framework abstracts a process from our stakeholder interactions and the design patterns of early technical systems seeking to manage the authenticity, rights and monetisation of data. ACCCT builds directly upon prior works exploring data rights and monetisation. Early work from the ODI explored the stewardship and monetization of data in Data Trusts [ODI, 2019]. Frameworks for decentralised content rights and monetization derived from content provenance have been proposed in Content ARCs [Balan et al, 2025], and the Ownership-Rights-Attribution (ORA) model [Balan et al, 2023a]. Draft standards are exploring layered models for the same e.g. the Media Tokenization group of JPEG Trust [JPEG Trust, 2025].

In ACCCT we are inspired by these past systems and frameworks, to frame the process as a sequence of technology families and explicitly seek to explore the state of the art, barriers to adoption, and calls to action to help advance the technology and rate of adoption at each stage. The intention is to reflect on where we are now, and what it may take to drive up the technology readiness level (TRL) of implementations to create effective solutions for the creative economy.

A UNIT-LEVEL APPROACH

Creators today have several technical means available to express preferences about whether their content can be used for AI training or processing. These preferences can be communicated at two distinct levels: site-based (location – such as a website) or unit-based (embedded in the asset itself). Each approach has strengths and limitations in terms of scope, and persistence across the content supply chain.

Site-based methods allow a creator or content host to apply a blanket permission or restriction to assets located on all, or part, of a website. For example, the widely used robots.txt file can signal to web crawlers which parts of a website are permitted for indexing or scraping and has been used to opt-out of search engine scraping for decades. A similar mechanism, TDMRep (Text and Data Mining Reservation Protocol) [TDMRep, 2024], has been proposed by the W3C for AI-specific opt-out, allowing a website to specify whether its content can be mined for AI training. Site-based approaches are efficient for expressing opt-out in bulk; however the signal is not embedded in the assets themselves and so does not persist when content is copied, shared, or aggregated downstream. This severely limits ability to enforce creator consent in the content supply chain.

Unit-based methods, in contrast, allow creators to attach permissions and restrictions directly to individual assets. This is achieved by embedding metadata within the file itself using open standards such as C2PA (Coalition for Content Provenance and Authenticity) that can be applied to assets of most modalities (e.g. image, video, audio, text) across several commonly used file formats [C2PA, 2024]. Given the scale of data reuse in AI, any opt-in or opt-out mechanism must be designed to operate at a similarly large scale to be effective. Natural Language Processing (NLP) is not yet capable of accurately parsing and acting upon human-readable text at scale, underscoring the need for a machine-readable method that can consistently signal consent or objection in a structured and interpretable way.

A key assumption is therefore that a unit-based approach is required, and the remainder of this section will examine the essential building blocks of a unit-based approach.

DESCRIPTION OF THE ACCCT FRAMEWORK:

Asserting unit-level provenance of assets

To be able to trace who created an asset, provenance information (i.e. the identity of who made the asset and how) needs to be asserted at unit-level.

To be truly effective here, a machine-readable approach is required which must be interoperable across the diverse range of platforms that will need to read and respect these signals. A proprietary or fragmented approach would create friction and reduce compliance, making it essential to adopt open, patent-free standards that can be consistently implemented across different platforms and ecosystems. Open standards in this area are already being developed by industry, and are discussed further below. These standards describe an interoperable way to record provenance information into metadata embedded in the asset. By 'open' we refer to standards that are community-developed and free to use.

Content identification ensuring durability of information

In order for relevant provenance data to remain linked to an asset, an open standard would describe both the provenance information that may appear, as well as a way for the information to be attached in a 'durable' and machine-readable way such that it persists even when the content is distributed or modified. For example, metadata alone can be easily stripped by content platforms such as social media sites that do not implement provenance standards. Commonly provenance metadata is therefore backed by technologies to make it 'sticky', such as watermarking (where an invisible identifier is actively injected into an asset) or fingerprinting (where a digital signature called a 'hash' is passively obtained from the asset in order to identify it) [Collomosse and Parsons, 2024].

Case study

Coalition for Content Provenance and Authenticity (C2PA) Standard

C2PA [C2PA, 2024] is an example of a standard which is particularly useful in this context. It provides a framework for embedding provenance metadata (referred to as Content Credentials) directly into media files, which describe the origin, authorship, and modifications made to a piece of content. The standard has also been helpfully extended to include basic AI training opt-in/out preference – so that these preferences may also be asserted at asset level. C2PA is not owned by any company and has been developed as a free, open standard for asset (unit) level provenance by collaborators across industry and academia.

C2PA describes how Content Credentials metadata can be made durable through its combination with fingerprinting and watermarking at asset-level. Watermarking embeds an identification signal invisibly into the asset. Fingerprinting is used to improve the security of the watermark. This combination has been shown to be far more robust than any one technology alone. This can then be coupled with a registry that contains a record of the metadata, such that it may be recovered in the event that it is stripped from the asset [Collomosse and Parsons, 2024; NCSC, 2025].

The C2PA standard is already gaining widespread adoption across major platforms, including Meta, Google, Adobe, Microsoft, Amazon, TikTok, LinkedIn, and Samsung, and is being fast-tracked as an ISO standard under ISO/IEC 21617-1:2025. It has been incorporated into the Joint Photographic Experts Group (JPEG) Trust standard. Several camera manufacturers (Nikon, Canon, Sony) ship cameras that write C2PA metadata, as do many digital tools such as Adobe Photoshop, and Microsoft Designer. Microsoft Office and LinkedIn are examples of platforms that display the metadata, and several browser extensions are available to make metadata visible on the web. Adobe ship a specific tool ‘Adobe Content Authenticity’ that allows users to apply AI opt-in/out preferences in bulk to images using C2PA, and their Generative AI model (Adobe Firefly) will respect those preferences.

The development of the C2PA standard illustrates how industry is already beginning to solve a key challenge within this framework; however, despite this uptake there is still inconsistent implementation across platforms, with many platforms, devices and tools now writing this metadata, but fewer platforms currently reading it [NCSC, 2025]. The approach to durable Content Credentials outlined by C2PA can be applied across audio, video, text and image modalities to name but a few, however implementation in the wild has so far focussed primarily on images.

Ownership and rights holder identification

The next step in the framework requires the owner and/or rights holder (who may or may not be the same person) to be able to assert their ownership of an asset in a way which can be linked to provenance, and for the AI developer to be able to verify that they are the owner.

Asserting and verifying ownership of an asset is not straightforward, due to the requirement to create an interoperable way of recording this information at the asset level (again, motivating open standards). Due to the potential for platforms or adversaries to strip metadata containing ownership information, it is desirable from a technological perspective to record this information in a registry. For a registry approach to work on an international scale, a centralised registry is likely to present scalability and governance challenges, making a decentralised ledger attractive. Technologies to support a decentralised ledger of this nature (such as Distributed Ledger Technologies or 'DLT') have been used in this way.

RECOMMENDATION:

Government backing of open standards for provenance information ensures compliance by requiring consistent implementation of a standardised approach across platforms and ecosystems, with consistent support for image, video, audio and text modalities, enabling creators to seek legal remedies if their rights are violated.

RECOMMENDATION:

Awareness raising and upskilling initiatives that lower the barrier to entry for content creators and rights holders to access current and upcoming technologies for 'open rights standards' that reflects the nuances of their domain and ultimately helps to make licencing a more machine-readable process.

During the global pandemic Non Fungible Tokens [NFTs] [ERC 721, 2018] become a popular way for creators to establish and transfer (i.e. sell) ownership of assets using blockchain technology. Blockchain is a form of decentralised database, and although most commonly used to track cryptocurrency ownership, it can be used for other purposes. Despite suffering a boom-and-bust cycle, NFTs showed for the first time that secure global decentralised registries of asset ownership capable of supporting billions of pounds of transactions was practically achievable. Although interest in NFTs has waned, the core technology (i.e. DLT) underpinning them retains value in asserting ownership in a decentralised way. Recently efforts have been made to build upon DLT to create a decentralised database for opt-in/out signals [Balan et al, 2023b]. A further complication is the need to link the unique identifier of the asset (see Section on provenance) with an identifier of the owner using an agreed identification technology such as Self-Sovereign Identity or Verifiable Credentials) [VC, 2008], in a transparent way that both rights holders and AI developers can trust. This has been trialled in the ORA (Ownership-Rights-Attribution) framework, linking ownership provenance with creation provenance to create a licensing platform for generative AI data [Balan et al, 2023a] that facilitated micropayments to creators. Other blockchain based registries for asset ownership include the Ocean [Ocean] and Story [Story] Protocols, both of which are pivoting to AI data commodification business models.

GLOSSARY

Registries The term 'registry' can mean different things from a technical and copyright/legal perspective. In this context, in a technical sense it is used to describe the use of some form of database in order to, for example, store and recover lost metadata or record ownership data. From a copyright perspective, this can refer to 'rights registry' such as is used in the US-style system, which may be used for policy rather than technical reasons.
In practice there may be overlap – but ascertaining clarity of purpose and intended methods of use between stakeholders is key in exploring this area.

RECOMMENDATION:

Support for further research activity to trial and scale systems for asserting, ascertaining and verifying ownership, in conjunction with provenance technologies. This should build upon existing work in the area (as outlined above) and include research into the practicality and usability of systems as well as technical solutions.

Case study

Data Discoverability and Innovative Marketplaces

As well as the case where assets are 'found' online through web crawling, for example, there is also a different scenario whereby AI developers perform a more manual search for owners of large data sets that may be relevant to their work, and make direct approaches to licence their content. While this was not the focus of the workshop, some discussion of this surfaced, which may offer further opportunities for innovation. Challenges include AI developers struggling to identify who might own large data sets (that may or may not be accessible online), which rights holders may be open to negotiating use for AI development purposes, and the need for expensive pre-processing of data to understand the form, quality, and potential value of data within a data set. Beyond this, the deals themselves are challenging due to little precedent for licencing and valuing data sets for this use. Nonetheless, some AI developers describe exploring innovative new ways of working with rights holders, whereby they agree a research licence to evaluate data sets and explore which data are of most value, before deciding which data they wish to licence for training and use for production software.

Another option that was raised was the idea of a voluntary, government-backed 'forum' that could act as a marketplace for those interesting in granting and obtaining rights for these purposes, which would aid with data discoverability. In this way, rights holders could choose to list their works and appropriate licence information, or else signal that direct deals are possible. This could potentially encourage innovation and collaboration, particularly if coupled with the idea of a charter and/or kitemark discussed in Section 4 that helped drive best practice.

RECOMMENDATION:

Trialling a non-mandatory Government-backed and verified database or 'forum' to enable rights holders to list their assets and AI developers (or others) to understand where valuable data sets may be found and who owns them.

Case study

Granularity of Permissions for AI Training

A clear user requirement identified is the need for a more granular approach to opting in or out of different generative AI use cases, under different conditions.

The C2PA [C2PA, 2024] and IPTC [IPTC, 2023] standards, for example, do provide for the ability to stipulate basic preferences about opting an asset in or out of AI training. If this information has been asserted by the creator, the information would then be available to any AI developer also implementing the standard. Assuming a commercial use case, the remaining steps (identifying and verifying ownership and rights holders, licencing, contracting, attribution and value creation) still apply, but the essential permissions information has been established by looking at the provenance information described above.

However, by design, as this standard focuses on data provenance, it does not contain further granularity about how and when an asset can be used in the context of AI or generative AI (such as allowing for content discoverability or medical research, for example). It also does not contain any information about the owner, rights holder or licences. Detailed information about ownership and rights is generally considered to be outside the scope of an open provenance standard. This information therefore needs to be standardised and obtained for an asset in a different way.

In future, more granular AI permissions could be contained within the licence itself, alongside other standard licencing information describing how an asset may be used. The licence would then be linked to the provenance information of the asset. However, for this to work at scale, a licence would need to be held in a standardised, machine-readable format – for which there is not currently an agreed upon industry standard (see section on Licencing, p. 29).

Licencing

If the previous steps of the framework have been followed, the AI developer can now trace and verify the identity of the creator, owner and rights holders of the asset they wish to use. The next step is then to establish the terms under which they may use the asset including, in future, nuanced permissions/restrictions for generative AI use. This is done by requesting a licence from the rights holder, who then needs to be able to issue usage rights to the AI developer (Figure 5).

Some existing technologies may be used here to issue this licence, such as a signed digital file, or emerging technologies have been explored such as representing the licence as a token on a blockchain [Balan et al, 2025]. However, the main limitation here is that the information contained within a licence is not currently in a standardised, machine-readable format which can be processed with the consistency and scale required.

The innovation required is therefore the development of some form of 'open rights standard' to help automate the process at scale. For the avoidance of doubt, we are not making recommendations here about the nature or specific terms and conditions of the licence itself (granting permissions with associated warranties remain down to rights holders themselves, as it is now). Here we are referring only to the format that the licence is presented in – such that it may be machine-readable and therefore issued in an automated fashion, at scale.

RECOMMENDATION:

As an 'open rights standard' gains maturity, government backing for this form of standard will drive adoption and ensure compliance.

GLOSSARY

Open Standards 'Open standards' refer to collaboratively developed, freely available technical specifications, which promote interoperability and data exchange. Different standards have different scopes – two examples given in this document are "open provenance standards" and the suggestion of an "open rights standard". This should not be confused with the notion of "open access" in a rights context, which describes forms of freely accessible content. For the avoidance of doubt, the use of an interoperable, open standard to express the rights of an asset does not imply the granting of open access to re-use an asset.

In contrast to the maturity and adoption of provenance and identity standards, today, the development of open standards for rights and licensing remains an area of active development. Some examples include early work on machine readable usage rights by the Open Digital Rights Language (ODRL) [ODRL, 2018]; however, at present there is still not an agreed industry standard. There is nascent work in the Joint Picture Experts Group (JPEG) community to explore the use of such standards in the context of content licensing [JPEG Trust, 2025] and the wider landscape is surveyed in [Balan et al, 2025]. It is also important to distinguish that this is a suggestion for a form of open rights expression rather than a mechanism for technical access control. Technical access controls – commonly known as Digital Rights Management (DRM) – have historically failed to prevent unauthorized content reuse. Notable examples include DeCSS (which bypassed DVD encryption) and AACS (hacked for Blu-ray), which demonstrated the limitations of purely technical enforcement mechanisms. In contrast to the technical enforcement offered by DRM, the enforcement of an open rights standard for licensing could be through future regulation or legal recourse.

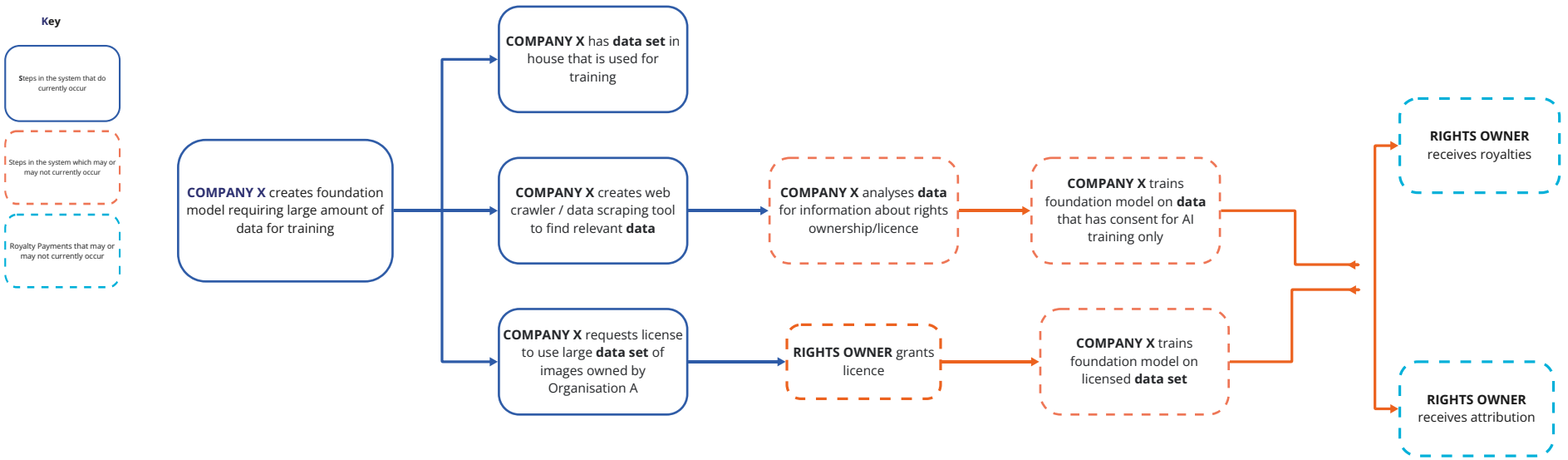


Figure 5: Developers or companies working with generative AI within the standards of the ACCCT framework, in its simplest form, could be encouraged to integrate additional checks of rights information (orange dashed boxes), the inclusion of attribution and where appropriate royalties (blue dashed boxes) into their processes.

Contracting

Once a licence is issued and relevant permissions/restrictions have been ascertained, a machine-readable contract needs to be issued automatically between the AI developer and rights holder, outlining terms of use, remuneration, etc.

Existing technologies exist for issuing electronic contracts (such as DocuSign) and for issuing contracts in the form of executable computer code (referred to as Smart Contracts) that run on a blockchain. This could facilitate real-time automated royalty payments to one party, when another performs a computationally verifiable action such as incorporating an asset into an AI training data set. Any payment model could either be determined on an upfront fee or royalty (pay per use) basis, similar to streaming platforms. The latter would need to account for fundamental terms such as audit and others.

In the context of digital marketplaces, smart contracts act as trustless intermediaries that enforce terms automatically and transparently – ensuring that content creators retain control over their intellectual property while providing AI companies with verifiable, legally compliant terms to high-quality data for training and calibration.

The nature and terms and conditions of the contracts themselves is ultimately down to the relevant parties, and for the avoidance of doubt we are not suggesting 'standardised contracts' here. However, to lower the barrier for many smaller content owners, it is suggested that development of a set of accessible contract templates that describe broader terms for the context of use, for creators and developers, would be helpful here. This could potentially reflect work done on developing best practice (as described in the Section on Value Creation below). Ultimately, use or adaptation of these contract templates remains optional.

RECOMMENDATION:

Create a bank of agreement templates that provide a UK-wide infrastructure for all players to access transparency of terms, similar to the Lambert and Brunswick templates made available through the IPO. This would include access to the technological forms of the contract for attribution – such as the use of DLT (Distributed Ledger Technologies) Smart Contracts, or more traditional e-signature systems, depending on the scale of data/agreement being executed. Such templates could form the basis for negotiations between parties; however, use or adaptation of these contract templates ultimately remains optional.

RECOMMENDATION:
To provide support, such as targeted research and innovation funding, for the development and scaling of attribution algorithms which enable rights holders and AI developers to transparently link outputs from generative AI inputs.

Attribution

In Copyright law, attribution refers to the Moral Right of a copyright holder to be identified as the author or creator of that work (i.e. credited). This is not currently happening in a consistent or transparent way in the context of usage for AI development.

Different approaches could be considered here. Potentially, a licence could stipulate that if the data is used within a training set, the author must be credited regardless of how much it has influenced the output. However, given the large volume of data and relatively small amount of attribution likely to be due to any one asset, other methods may be more appropriate. At present, rightsholders and AI developers are left with a predicament where the value of any particular asset within a dataset is difficult to calculate. Attribution algorithms can be used to determine the subset of data that is most responsible for a synthetic data or model [Wang, 2024], however these technologies are still in their early stages and are not proven at the scale of billions of assets commonly used for generative AI training. After which, the appropriate rights owner/ author may be credited in an agreed way (which may be outlined in a smart contract [Balan et al, 2023a]). This would then need to be reflected in the set of contract templates mentioned previously.

Value creation

The ACCCT framework requires both micro and macro support to develop a thriving business ecosystem. At the micro level, a method is required to compensate the rights holder for use of the asset based on the license and the contract. If a financial transaction is required, the mechanism can follow established routes (such as a Visa or Mastercard transaction, cryptocurrency micropayment etc.). However, it is vital to think beyond cash transactions to support a sustainable and inclusive ecosystem where creative contributions and the value of different forms of data and their use by developers can be recognised. Alternative forms of recompense must be supported and piloted such as digital tokens, access privileges, reputation systems and in-kind services.

At the macro level, the UK needs to support an ecosystem of interconnected organisations that collectively co-create value. This necessitates building marketplaces that enable operational efficiency and tackle the legal complexity inherent in providing compensation around content licensing. Marketplaces scale and enable different stakeholders to operate using different business models – this is important in that there is significant uncertainty as to what models might be effective and offer scale. Workshop participants were particularly keen to underscore that these business models needed to not simply suit the global big players, but also value and incentivise the creator economy, as well as provide flexibility and affordability for new entrants and SME AI developers who cannot afford to buy data in huge bulks.

The innovation here is in developing best practice for business models and payment models, which may be linked to establishing attribution as described above.

RECOMMENDATION:
Further research and testing of new forms of business and payment models that are effective, scalable and incentivise the creator economy and AI developers.

Marketplace Type	Description of Potential Use
Decentralised Licensing	Content creators list works along with detailed licensing terms (e.g., usage rights, royalty structures, permitted applications). AI companies purchase or subscribe to these licenses, with smart contracts automatically enforcing the agreed-upon terms. A distributed ledger creates an immutable record of the licensing agreement, ensuring transparency for both parties.
Royalty and Revenue Sharing Marketplaces	Smart contracts dynamically distribute royalties or revenue shares to content creators based on metrics like the frequency of usage or milestones in model development. Payments could be based on real-time or batch processing of usage data provided by trusted third parties (oracles), ensuring that content creators are compensated fairly when their work contributes to improved model performance.
Tokenised Content Ownership and Access	Content creators tokenize their intellectual property, issuing a unique digital token (or NFT) that represents ownership or usage rights. AI companies purchase or lease these tokens to obtain non-exclusive or exclusive rights for training purposes. Smart contracts manage access control, usage limits, and any resale or royalty provisions, reducing disputes about ownership and rights.
Auction and Negotiation-Based Licensing	For high-demand or high-quality content, an auction mechanism can be implemented. AI companies bid for exclusive or time-limited licenses to use certain data sets. As an alternative, or in addition, negotiation protocols encoded in smart contracts could allow both parties to fine tune parameters such as training usage limits, update frequencies, and premium fees tied to performance benchmarks.
Subscription and Access Control	AI companies subscribe to curated data feeds or data sets, with smart contracts automating recurring payments and ensuring that usage aligns with subscription parameters (e.g., data access frequency, expiration dates, and cancellation conditions). This model would enable continuous, reliable content provision as models are constantly updated.
Community Sharing and Content Sponsoring (Open Access)	Content creators make some of their works (and/or parts thereof) freely available for certain use cases, such as for non-commercial research purposes. A dedicated 'community' marketplace provides content creators with additional visibility for their work and opens the opportunity for commercial organisations to sponsor certain content to be made freely available to non-profit institutions. (Similar to: Open access of academic publications.)

Table 1: Example marketplaces that could be trialled to bring content creators and AI companies together.

For any such marketplace to exist, there will need to be a body supporting compliance and best practice.

RECOMMENDATION:

To enable functioning marketplaces to develop, government should work with industry to co-design an industry oversight body to promote best practice and enforce compliance. For example, an independent legal ombudsman, or a non-departmental public body which reports to Parliament, such as the Information Commissioner's Office (ICO).

Whilst technologies, standards and legislation are being developed, a practical step to consider to begin establishing trust is the development of a 'charter' of best practice principles which AI developers can sign up to.

RECOMMENDATION:

Explore the development of a 'charter' of best practice principles including: fair and transparent attribution practice; terms of trade for compensation; use of appropriate open standards to allow content provenance, rights information, and appropriate permissions; and transparency of high-level information about what data will and will not be used for. To provide incentive to the Charter, AI companies working to the Charter could achieve a 'Kitemark' that was mandated as essential for all Government and Government-funded bodies' work. Rights holders could have the option of being able to indicate their consent for use by companies who have achieved the Kitemark.

Case study

ORA

EKILA [Balan et al, 2023] is one of the earliest technical frameworks proposed for recognizing and rewarding data contributors to Generative AI ('Generative AI) model training. Given a synthetic image created from a Generative AI text-to-image model (Stable Diffusion 1.4), EKILA can identify the subset of training images most responsible for that generated image and issue micropayment royalties to the owners of those images. This provides a new way to generate value in the creator economy which may also incentivize data contribution, to meet the inexorable demand of AI for high quality data (Figure 1).

To achieve this, EKILA built upon the C2PA open standard for content provenance – combining it with AI data attribution and Distributed Ledger Technologies (DLT) to create a technical framework known as ORA (Ownership-Rights-Attribution). C2PA is a widely adopted standard used primarily to establish the authenticity of assets. By design, communicates neither the ownership of an asset, nor (beyond the ability to express opt-in/out permissions for AI) usage rights or licensing. ORA proposed a way to combine C2PA with emerging technologies for describing ownership and rights, leveraging provenance for applications beyond authenticity to novel models of value creation for asset re-use [Collomosse and Parsons, 2024].

At the time (circa 2022) there was considerable interest in the use of Non-Fungible Tokens (NFTs) for asserting and transferring ownership of digital assets. NFTs are digital tokens that represent a digital or physical asset registered on a Distributed Ledger Technology (DLT) or colloquially a 'blockchain'. ORA proposed to combine creation provenance (via C2PA) with ownership provenance (via NFT), as well as a way to register and issue licenses to those assets, also via NFTs. The ORA framework therefore provided not only a decentralised registry of ownership and rights, but also a way to create 'smart contracts' for the automatic compensation of content owners, via micropayments, for the re-use of their assets (in particular, the use of those assets in generative AI). Follow-up work has explored the monetization framework proposed by ORA through in-situ studies [Liddell et al, 2024] and creator workshops [Digital Catapult, 2024].

Section 5:
Limitations and Future Iterations of the Framework

If the technology families within the ACCCT framework are the jigsaw pieces to a puzzle we must collectively solve to make the UK a world leader in the ethical development of generative AI and copyright, then it is also not a definitive solution. With our workshop participants we examined how the jigsaw might evolve over time, identifying the limits of the framework that would need to be addressed in a future roadmap for the sector. Here we identify those limits and opportunities to ensure ACCCT can provide an ongoing framework of access, control, consent, compensation and transparency. An important point about the limitations we acknowledge here is that by starting to act, new behaviours and business marketplace norms would be incentivised, making the challenges we list below more easily surmountable than they may currently appear. For example, in a well-functioning marketplace for transparent access to licensed copyright material, the Charter and/or Kitemark, amongst other factors, would provide a strong incentive for AI developers to retrain models in line with the business and regulatory consent norms of the market. Through such norms, retrospective rights may be compensated in future iterations of an AI developer's model.

COMPOSITE, CASCADING AND DERIVATIVE RIGHTS

The framework assumes that there will need to be a series of mutual obligations on all involved in the value chain to guarantee provenance and permissions. This is a particularly complex area with many content types made up of a patchwork of interlocking IP rights (e.g. music, film and TV). This limited research exercise does not aim to present a complete solution in this area; further work is needed to understand what solutions are feasible to facilitate the licensing of this type of content. Some of these solutions may be technical, while others may be within the nature of the licences themselves – which is outside the scope of the framework suggested. Co-designing solutions with representatives of different industry verticals is crucial for further progress on this matter.

UNLEARNING

There is currently no reliable process for verifiably removing the influence of distinct assets on a trained generative AI foundation model [Du et al, 2024; Hu et al, 2025] – whether for infringement or changed licence terms. Instead, a model requires being retrained from a clean slate ensuring the distinct assets in question are not included in the training data. The 'carrot' of ACCCT is that AI developers retrain models of their own volitions and in cycles to engage in ethical practices and comply with the Charter and gain an industry 'Kitemark' for being a 'good actor'. The 'stick' would require companies found to infringe or fail to comply with the framework to retrain their models, as well as remaining open to legal action. With further work on unlearning it would be possible to create a more dynamic system that enables, for example, time-limited testing of content for generative AI training or licensing for partial data sets.

RIGHTS TRANSIENCE

It is currently standard practice for licenses to be given for set time frames or in perpetuity. As generative AI foundation models do not currently have a mechanism to forget, rights holders are effectively being asked to provide content licenses in perpetuity or at least until a model is re-trained. There is a need to provide greater transparency on when models will be retrained to enable fairer negotiations for all parties, recognising that smaller AI developers will not retrain models as frequently as large players.

RECOMMENDATION:

To provide support to urgent and underlying socio-technological challenges, enabling emerging technologies to scale rapidly. In particular, support should be given to address the linked problems of generative AI models unlearning where copyright permissions have not been given and/or withdrawn, and attribution algorithms which enable rights holders and AI developers to transparently link outputs from generative AI inputs. Whilst prototype technologies in these spaces exist, they currently don't scale with accuracy and raise important associated legal and social questions that need tackling hand-in-hand.

RETROSPECTIVE RIGHTS

Many creators of content are interested in understanding if their work has already been used in model training without their knowledge or consent. If so, they would like to find ways to protect their work and/or to receive value from its use. The ACCCT framework is limited in this regard as it focuses on resolving transparency issues going forward. This is an area where more work is needed but is not necessarily an impediment to moving forward with the ACCCT framework proposal, particularly because rights holders would still be left with recourse to the courts under current legal protections. The ACCCT framework's proposal to include a Charter would require companies to acknowledge previous unregulated practice and seek to make good on copyright infringement where it exists on the terms of licences established through ACCCT: incentivising a solution rather than punishing past behaviour.

NOTIFICATION

ACCCT requires an agreed mechanism for providing informed consent to enhance transparency where the asset owner is informed about intended and actual use of the content. The unit-based granularity of consent supported by ACCCT provides control over use in a range of circumstances. However, an agreed method for notifying rights holders about the use of their data is required. This may be direct communication with the rights holder or indirect (e.g. publishing which data sets have been used, providing an audit option). The range of scenarios contemplated in which notification of intended or actual use would be required suggested further consultation on how to develop an informed consent and notification mechanism that is scalable and use specific is important to rights holders.

THIS IS NOT A DRM SOLUTION

It is important to distinguish that at the core of our technological solutions are standards that support a form of *open rights expression* rather than a mechanism for technical access control. As noted previously, technical access controls—commonly known as Digital Rights Management (DRM)—have historically failed to prevent unauthorised content reuse. The danger of our current moment is that, without government intervention, large corporations will enact proprietary DRM solutions that stifle innovation and new players, fragment the marketplace and do little for end-user/consumer rights. Legal recognition for open standards for granular control and consent would significantly strengthen their enforceability, enabling creators to seek legal remedies if their rights and licences are violated.

DOWNSTREAM COPIES

The existing ACCCT roadmap works when good actors engage with ethical practice. Existing provenance technologies cannot protect assets from having provenance data stripped purposely or accidently through screenshots, cropping or processing by social media websites (for example). Whilst ACCCT takes steps towards embedding and enhancing the opportunity for good actors to access and respect content creation and ownership provenance, globally accepted standards are required to ensure good actors retain this information in the assets during processing. It is also understood that bad actors remain in the eco-system and that unlawful use of assets may occur. The legal processes embedded in ACCCT aim to support asset owners in proving and asserting their rights where such circumstances occur.

INVISIBLE CONTENT

Video Games are an anomaly that legal development around ACCCT could benefit from considering further. Video games contain visible content (e.g. the art, sound, UI/UX design etc.) and invisible content (e.g. technical frameworks and code that make the game playable, provide visual polish and engage the player). The visible content, at its core, is similar to other content forms such as film or music and thus have existing frameworks for copyright protection. The underlying invisible technical content, however, is a unique feature of software. End User Agreements attempt to minimise the unlawful use of content but there is a need to define the processes for protecting and licensing this unique and important “invisible” content within the generative AI setting. The need to factor this invisible content aspect into the prototype and demonstrator stages of the road map is a crucial component of moving ACCCT forward.

Section 6:
Next Steps

Arguably, the potential value of ACCCT is immense. Materialising the framework in a complete sense, however, is a long-term project. Though some of the technologies and standards required to make the ACCCT framework and ecosystem possible are proven to a degree, others are emergent and require further work before an end-to-end version of the framework can be realised at scale. In addition, much of the legal and business frameworks for working with these tools is nascent.

Nonetheless, there is a need to demonstrate the potential value of the ACCCT framework and resulting ecosystem in the context of what is nascent – ideally providing companies with a safe space to experiment and to establish trust in a new framework. We therefore propose a two-stage process of materialising the framework moving from a simulated ecosystem to a fully developed demonstrator (Figure 6).

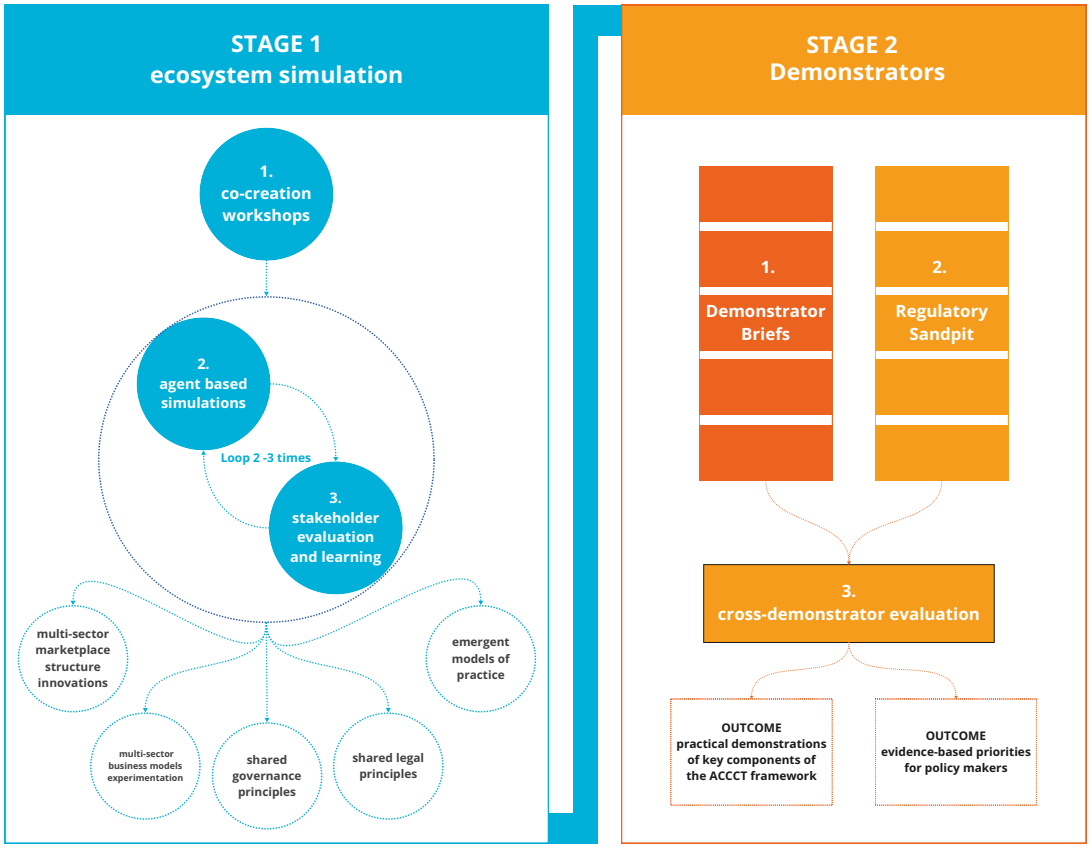


Figure 6: The two-stage process proposed towards enacting ACCCT: Ecosystem simulation with key stakeholders and demonstrator development of practical and policy outcomes.

STAGE 1: ECOSYSTEM SIMULATION

Initially, there is a need to simulate marketplaces to understand how value materialises for stakeholders within sub-sets of the creative industries, and across the ecosystem itself. Key activities include:

1. **Co-creation workshops and focus groups.** Undertaking a more detailed dive into the proposed framework with different users from different industries – to begin to map user journeys, pain points and practical considerations that need to be considered. Applying design thinking approaches with key stakeholders – content creators, rights holders, AI developers and policy makers – will help define the detail necessary to materialise the ACCCT framework in terms of governance, legal frameworks, terms of trade and challenges. Key provocations would include:

   a. What is the right level of granular **control** and **consent** required for provenance and ownership validation compared to the labour and value involved? [Digital Catapult, 2024]

   b. What business models and marketplaces provide appropriate forms of **compensation** for different sectors?

   c. What are the fair and equitable (smart) contract terms of trade for different sectors to allow **transparent** and efficient **access** to copyrighted content?

   d. What body or bodies need to be created/adapted for regulation and enforcement, royalty and revenue collection?

2. **Agent-based simulations.** To bypass the need to implement nascent technology stacks, rapid simulations will be used to model interactions between stakeholders, based on their current business models, practices etc. as guided by policies framed around one or more of the market-types outlined in the framework. The resulting simulations provide insight into the variety of potential scenarios and conditions that might or might not enable marketplaces and the wider ecosystem to work in practice (including an understanding of unintended consequences).

3. **Stakeholder evaluation and learning.** Working with stakeholders, the resulting scenarios can be used to (a) evaluate the degree to which the objective of co-creating and distributing value among all parties (or stakeholders) in a fair manner is achieved; (b) explore potential changes to current stakeholder business models; and (c) determine the practicalities of implementation – evaluating labour required and value generated through the different models.

Steps 2 and 3 are iterative, running through 2-3 cycles to enable different content types/rights holders and AI developers to experiment with marketplace structures and business models (e.g. Table 2) whilst developing detailed shared governance and legal principles.

Convening Body (established control)	**Content Form** (established variable)	**Standards & Technologies** (variable)	**Regulatory Sandpit** (co-designed control)	**Marketplace** (co-designed variable)
Innovate UK	Publishing and Journalism	TDM +XX	The regulatory environment would be the same in each experiment, with each prototype focused on developing recommendations for their specific content form	Decentralised licensing
CoSTAR	Audiovisual and Games	C2PA + Jpeg		Royalty and revenue
Creative Cluster	Music and text			Tokenised content and access
Digital Catapult	GLAM	ORA		Auction and Negotiation-Based Licensing

Table 2: Hypothetical example of simulation experiments

STAGE 2: DEMONSTRATORS

Next, there is a need to test outcomes from Stage 1 **at scale** and evaluate their impact on and existence within emergent marketplaces. Key activities include:

1. **Demonstrator briefs/specification.** Development of a series of demonstrator briefs that test Stage 1 outcomes and their *ability to support* scalable marketplaces or important aspects thereof (e.g., fair allocation algorithms). Incentives including pitch-based funding and tax breaks would be provided to content creators, rights holders and AI developers to promote participation (for example via the AI Opportunities Action Plan) to de-risk applied experimentation and to ensure **transparent** dissemination of results.

2. **Regulatory sandpits.** In parallel with the demonstrators, the emerging marketplace specifications can be evaluated under varying policy and regulatory conditions. The process can be disseminated through a convening body that took on the regulatory sandpit function (governed by contracts), including reporting on outcomes and making recommendations to ensure regulatory clarity going forward.

3. **Cross-demonstrator evaluation.** To ensure that the emerging ecosystem and regulatory priorities are fit for purpose, outcomes are evaluated by key stakeholders against their current policies and practices, business models, value propositions, and key challenges. This evaluation would enable the development of the Charter in the recommendation in *Section 4: Overarching Support of the ACCCT Framework*, so that future organisations complying with the ACCCT framework could receive a 'Kitemark' style certification.

Rights holders could choose to indicate that they are happy for their content to be used by companies who have committed to adhere to these standards.

The outcome of Stage 2 is thus a combination of evidence-based priorities for policy makers and practical demonstrations of key components of the ACCCT framework that, together, provide creative industries and AI developers with certainty and trust around the access, control, consent, compensation and transparency of creative works.

RECOMMENDATION:

Undertake a series of scaling prototype and demonstrator projects that enable the technical, legal and social solutions proposed in ACCCT to rapidly progress. The need is for safe and supported trials to test technologies and legal solutions that can then be implemented simultaneously to protect all parties and enable a vital, innovative creative ecosystem.

References

[Balan et al, 2023a] K. Balan, S. Agarwal, S. Jenni, A. Parsons, A. Gilbert and J. Collomosse. **EKILA: Synthetic Media Provenance and Attribution for Generative Art.** In *Proc. CVPR Workshop on Media Forensics (CVPRW).* 2023. Available at: https://arxiv.org/abs/2304.04639

[Balan et al, 2023b] K. Balan, A. Black, S. Jenni, A. Gilbert, A. Parsons and J. Collomosse. **DECORAIT - DECentralized Opt-in/out Registry for AI Training.** In *Proc. Conf. Visual Media Production (CVMP).* 2023. Available at: https://arxiv.org/abs/2309.14400

[Balan et al, 2025] K. Balan, A. Gilbert and J. Collomosse. **Content ARCs: Decentralized Content Rights in the Age of Generative AI.** In *Proc. Intl. Conf. on AI and the Digital Economy (CADE).* 2025. Available at: https://arxiv.org/pdf/2503.14519

[C2PA, 2024] **Coalition for Content Provenance and Authenticity.** "Technical Specification v2.1". 2024. Available at: https://c2pa.org

[CC, 2008] **Creative Commons. "CC REL: The Creative Commons Rights Expression Language."** 2008. Available at: https://creativecommons.org/ns

[Collomosse and Parsons, 2024] J. Collomosse and A. Parsons. **To Authenticity, and Beyond! Building Safe and Fair Generative AI upon the Three Pillars of Provenance.** *IEEE Computer Graphics and Applications (IEEE CG&A).* 2024. Available at: https://personalpages.surrey.ac.uk/j.collomosse/pubs/Collomosse-IEEECGA-2024.pdf

[DECADE, 2025] **The Future of AI and Copyright Law.** Available at: https://decade.ac.uk/2025/02/27/the-future-of-ai-and-copyright-law-decades-response-to-uk-government-consultation

[DeFilippi, 2016] De Filippi, P. and Hassan, S. **Blockchain Technology as a Regulatory Technology: From Code is Law to Law is Code.** First Monday, vol. 21, no. 12, 2016. Available at: https://arxiv.org/pdf/1801.02507

[Digital Catapult, 2024] **Emerging Futures for Media Tokenization and Digital Media Rights.** Technical Report, Digital Catapult. 2024. Available at: https://www.digicatapult.org.uk/wp-content/uploads/2024/12/Catapult_ORAgen_Emerging_Futures_Report.pdf

[Du et al, 2024] J. Du, Z. Wang, and K. Run. **Textual Unlearning Gives a False Sense of Unlearning.** Available at: https://arxiv.org/abs/2406.13348

[ERC 721, 2018] Entriken, W., Shirley, D., Evans, J., and Sachs, N. (2018). **ERC-721: Non-Fungible Token Standard. Ethereum Improvement Proposals.** Available at: https://eips.ethereum.org/EIPS/eip-721

References

[Hu et al, 2024] H. Hu, S. Wang, T. Dong, and M. Xue. **Learn What You Want to Unlearn: Unlearning Inversion Attacks against Machine Unlearning.** Proc. IEEE Symp. On Security and Privacy 2024. Available at: https://arxivorg/abs/2404.03233

[IPTC, 2023] **International Press Telecommunications Council (IPTC). "Photo Metadata Standard: IPTC Core and IPTC Photo Metadata."** 2023. Available at: https://iptc.org/

[Liddell et al, 2024] F. Liddell, E. Tallyn, E. Morgan, K. Balan, M. Disley, T. Koterwas, B. Dixon, C. Moruzzi, J. Collomosse and C. Elsden. **ORAgen: Exploring the Design of Attribution through Media Tokenisation.** In *Proc. ACM Designing Interactive Systems (DIS).* 2024.

[TDMRep, 2024] W3C Community Group. **TDM Reservation Protocol (TDMRep).** 2024. Available at: https://www.w3.org/community/reports/tdmrep/CG-FINAL-tdmrep-20240202/

[Ocean] **Meet Ocean: Tokenized AI and Data.** Available at: https://oceanprotocol.com/

[ODI, 2019] Open Data Institute (ODI). **Data Trusts: Lessons from Three Pilots.** ODI White Paper, 2019. Available at: https://theodi.org/insights/explainers/what-is-a-data-trust/

[ODRL, 2018] **Open Digital Rights Language**, W3C Technical Report. Available at: https://www.w3.org/TR/odrl-model/

[Story] **Introducing Story Protocol.** Available at: https://www.story.foundation/blog/vision

[JPEG Trust, 2025] Joint Photographic Experts Group (JPEG) Trust. **JPEG Media Tokenization.** Available at: https://jpeg.org/jpegtrust/

[NCSC, 2025] Five Eyes Intelligence Agencies CSI Report. **Joint guidance on content credentials and strengthening multimedia integrity in the generative artificial intelligence era.** January 2025. Available at: https://media.defense.gov/2025/Jan/29/2003634788/-1/-1/0/CSI-CONTENT-CREDENTIALS.PDF

[Wang, 2024]. J. Wang, Z. Deng, H. Chiba-Okabe, B. Barak, W. Su. **An Economic Solution to Copyright Challenges of Generative AI.** Available at: https://arxiv.org/abs/2404.13964

[VC, 2022] Verifiable Credentials Working Group, W3C. **"Verifiable Credentials Data Model 1.1."** W3C Recommendation, 2022. Available at: https://www.w3.org/2022/06/verifiable-credentials-wg-charter.html