RELX

RELX response to UK Intellectual Property Office Call for Views on AI and IP

RELX is a UK-based FTSE 20 provider of information and data analytics for professional and business customers across a range of sectors, including financial services, science, technology, medical, healthcare and energy.

We employ over 5,000 people in the UK and support customers in 180 countries. We utilise technology and data to help our customers improve their decision making across the sectors we serve. We help scientists make new discoveries, doctors and nurses improve the lives of patients, lawyers win cases, prevent online fraud and money laundering and insurance companies evaluate and predict risk.

RELX welcomes the opportunity to respond to the IPO's call for views on AI and IP. Artificial Intelligence (AI) and Intellectual Property (IP) are both key areas of interest for RELX's businesses. RELX has a unique perspective on the debate around the development of Artificial Intelligence (AI) and copyright. With an annual investment of more than £1 billion into technology and with 8,000 of our 33,000 employees being technologists RELX is both a significant consumer and developer of the latest technologies, including techniques which fall within the broad definition of Artificial Intelligence. At the same time, RELX is the rightsholder of large numbers of copyright works (most notably through our science, technology and medical journals publications within Elsevier). Hence, we are both a user and provider of data for the development of AI.

Do you agree with the above description of how AI may use copyright works and databases, when infringement takes place and which exceptions apply? Are there other technical and legal aspects that need to be considered?

The Intellectual Property Office's (IPO) description of how AI may use copyright works and databases provides an accurate account of the current legal framework surrounding the use of copyrighted works by AI systems, as we would expect given the IPO's expertise in this area.

We would add that, while they are often positioned in opposition to one another, copyright and AI actually complement each other. For AI systems to operate effectively and safely they have to be supplied with high-quality information. The best way to guarantee the quality of information used in an AI context is through copyright, which provides the underpinning mechanism that incentivises investment in its production and dissemination. This is particularly true when discussing AI systems used in the scientific and medical fields. For obvious reasons, scientists and clinicians need to have a high level of confidence that the information being fed into the algorithms that power the digital tools they are using is valid. To use an example from Elsevier, part of RELX: the Entellect platform allows search and analysis, driven by machine learning, across a wide range of databases including copyright content to assist with drug development and clinical trials. The value of such a product for end users derives from the combination of curated, valid content with the machine learning tool and analytics. Weakening copyright would increase the risk that the quality of underlying content used in an AI context deteriorates and in turn increases the risk of AI-based decisions in the drug discovery process being made on the basis of incorrect science or even fake science.

This call for views should be used as an opportunity to reset the debate between copyright and AI and to recognise how the two can work together to produce highly trusted and high quality AI systems which drive real change across society and the economy and yet incentivise the creation and protection of IP.

Within the IPO's description is the reference to an 'AI brain' and the suggestion that this is comparable to human brains in the context of temporary copying of copyrighted material. We would contend that

such an analogy is not appropriate. It suggests current AI systems are significantly more advanced than they are and that they are sentient beings, which is widely accepted not to be the case. Human brains making copies of works is a fundamental human reaction and could be described as involuntary to an extent. AI systems have an architecture which is deliberately and specifically designed to act in a certain way which means any copying of works would be premeditated. We therefore suggest the comparisons between human and AI brains are unsuitable.

Furthermore, we would add that the ability to read does not entail the ability to mine. It is argued that once a text is legally acquired and read then it should also be legally possible to apply text-mining techniques to it (such as those required in developing AI tools). One rational for this position is that this is what applies in the analogue world: that a book, once bought, can be read by as many times, or by as many people as possible. But this rationale fails because the machine-reading of the digital world is patently qualitatively and quantifiably different from human-reading in the analogue world. Subsequent readings can be made of a physical book without the need to incur the exclusive right to reproduction; there is little potential harm to rightsholders if a book, once purchased, is read by a number of people (the physical condition of the book quickly degrades and by around the sixth reader, they are better off buying a new copy). However, digital originals and their copies by their nature remain pristine and so there is no natural limit to their utility. Hence there is clear harm to rightsholders when multiple copies are made since there is never any requirement for a re-purchasing.

Is there a need for greater clarity about who is liable when an AI infringes copyright?

We do not believe any greater clarity is required for determining who is liable when an AI infringes copyright. As the IPO's document sets out clearly, when an AI infringes copyright, the person with control of the AI is liable.

Is there a need to clarify existing exceptions, to create new ones, or to promote licensing, in order to support the use of copyright works by AI systems? Please provide any evidence to justify this.

We do not believe existing exceptions need clarifying or that new ones need to be created in order to support the development of AI. As the IPO has indicated, 'It is not clear that the possibility of copyright infringement by AI systems is a major impediment to their use'. Prior to suggesting any changes to the copyright framework, which would be significant and could have unintended consequences, there would need to be a clear demonstration of a problem, which we do not believe exists. The reputation of the UK as one of the most advanced countries for AI development does not suggest there are any fundamental problems with the underlying IP regime.

While it may be true that AI systems require access to large quantities of data, this fact should not call into question the ability of rightsholders to protect their copyright if they choose to do so via technical or legal means. We would contend that organisations, especially businesses, that are claiming that copyright needs to change to promote AI make this assertion in an effort to avoid the expense and inconvenience of seeking licences and are engaged in an attempt to appropriate the value in content created by other businesses. Furthermore, we would note that the vast majority of data which will be accessed by AI systems will not be covered by copyright: copyrighted content probably represents a tiny subset of global 'data'. Even if it could be successfully argued that it is overly cumbersome to agree licences to copyright-protected content for use in AI applications, and we have not seen evidence to confirm this, removing those 'barriers' is unlikely to advance AI development meaningfully.

In cases where AI systems do require, or would like, access to copyright content, there are sufficient mechanisms to facilitate that access. For example, as the IPO outlines, it is possible to enter licencing

arrangements should the rightsholder wish to provide access to their high-quality data or content to develop AI programmes or outputs. These systems, which are constantly evolving, are already capable of dealing with the development of AI. For example, Elsevier, a RELX business, has well established mechanisms and processes to enter into discussions with third parties who wish to access Elsevier's copyrighted material and different licencing arrangements are available depending on the type of activity the third party wishes to conduct. Equally, we have found as company that it is possible to secure through the negotiation of licences third party copyright material to be used within our own analytical products. For example, Elsevier's Scopus is the largest abstract and citation database of peer-reviewed literature: scientific journals, books and conference proceedings. Scopus has developed this database by integrating content licensed from about 12,000 publishers. The completion of this huge task demonstrates that the current system of licencing already provides a framework which facilitates access to copyright content.

Furthermore, we would contend that the current licencing framework provides welcome flexibility within the system to allow negotiation between different parties, reflecting the reality of commercial relationships. Having a market for data is an important factor driving the Digital Economy. With the development of AI attracting increasing levels of investment around the world it is one of the growth areas within the digital economy. Some make the argument that because AI will deliver so many positive benefits to society and the economy there should be minimal restraints on its development, and in particular there should be free access to data. But policy makers should not seek to favour AI developers at the expense of rightsholders in this fast-moving, nascent and unpredictable market. The digital economy should be exactly that, a market-driven process, not a system of state-sanctioned handouts. The development of data, be that in the form of text, images, music curated databases, machine-data or any other form, requires investment on the part of other parties. To suggest that this effort and endeavour should go in some way unremunerated in order that AI development can continue, is to create an unfair – and potentially highly damaging – bias in the economy. As a matter of principle in market economies, production inputs should be the subject of fair exchange.

Of course, we recognise that a full commercial relationship is not always appropriate, particularly when you are not dealing with commercial entities. We therefore support the UK's existing exception for non-commercial research, and fully facilitate this as part of our licencing framework. For example, RELX's Elsevier business engages in a number of research collaborations in the UK as part of this work, which allows researchers to fully text and data mine Elsevier's content for the purposes of their research.

However, extending exceptions beyond, for example, the existing UK TDM exception would call into question the UK's commitment to copyright and the ability of rightsholders to protect the value of their content and their investments in creating it.

Is there a need to provide additional protection for copyright or database owners whose works are used by AI systems? Please provide any evidence to justify this.

We are not convinced of the need for an additional protection for copyright owners whose works are used by an AI system, as long as the existing rules are enforced effectively. The Berne Convention provides international consensus that rightsholders are able to protect their copyright works. A recognition of this right of ownership by those who seek to undermine copyright would be more beneficial than the creation of any new rights. Stronger enforcement of the existing regime would also be more beneficial than any new rights.

Should content generated by artificial intelligence be eligible for protection by copyright or related rights? If so, what form should this protection take, who should benefit from it, and how long

should it last? Do other issues need to be considered in relation to content produced by AI systems?

The UK has one of the most progressive and advanced IP frameworks for dealing with AI. Since the Copyright, Designs and Patents Act 1988 the UK has provided protection for computer-generated works with no human author. Given these provisions already exist in UK law, we see little reason to change a system which already addresses these questions.

Does copyright provide adequate protection for software which implements artificial intelligence? Does copyright or copyright licensing create any unreasonable obstacles to the use of AI software?

We do not see any problems within the UK with regard to copyright protections for software or obstacles to licensing AI software.

For further information on this submission, or for further discussions please contact:

[redacted]