

Making Law Open-Access to Improve Legal Competitiveness

* britishprogress.org/uk-day-one/making-law-open-access-to-improve-legal-competitiv

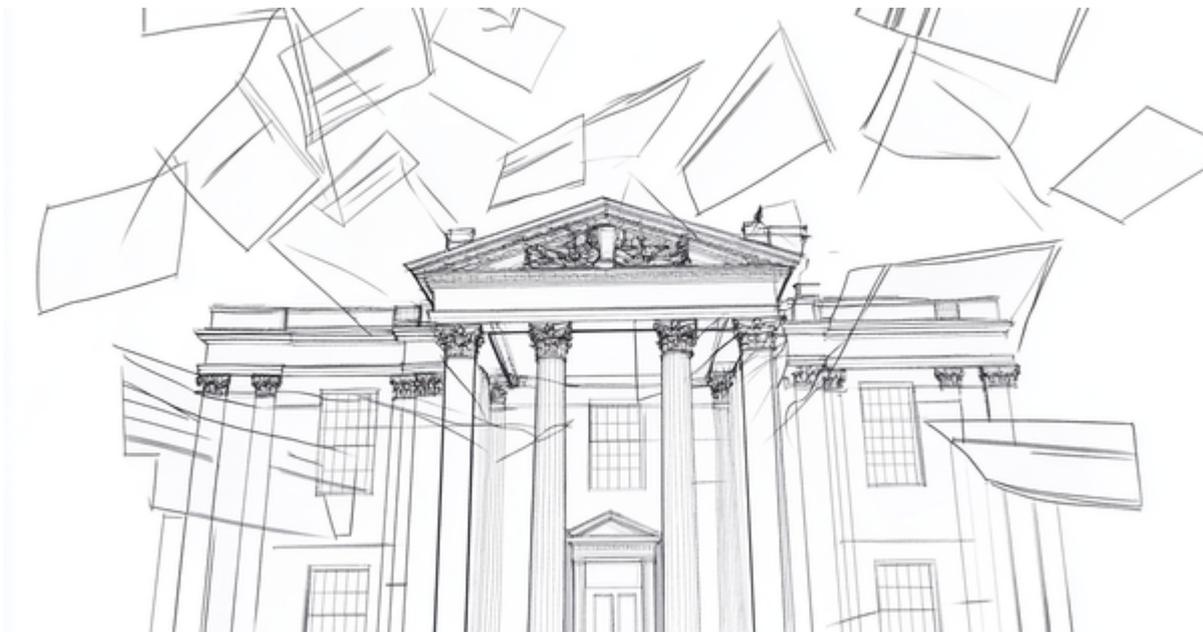


Table of Contents

- 1. Summary
- 2. Challenge and Opportunity
- 3. Plan of Action
- 4. FAQs
- 5. Authors

Summary

- The majority of legal cases in England and Wales are not freely available to the public. Instead, they are provided by two private companies—Incorporated Council for Law Reporting and Lexis Nexis Butterworths—which charge high fees for access and restrict their use.
- Given the precedent-based system of common law, this is a long-recognised problem in terms of access to justice; increasingly, it will also act as a drag-factor on technological innovation.
- Many new technological innovations require access to data. Concentrating an essential part of English legal text in the hands of a few private companies unnecessarily handicaps the development of legal technology.
- The English legal system competes on a world stage and is a major UK export. However, this dominance is not assured, and will depend on the ability of British law firms to adapt to new innovations.

- English and Welsh legal cases should be made publicly available through cheap and straightforward reforms to the National Archive's 'Find Case Law Service' (TNA). TNA currently publishes many court reports, but has significant omissions due to existing regulations.
- Expanding TNA's remit, when combined with the existing [legislation.gov.uk](https://www.legislation.gov.uk) project to make statutory legislation available, would make the vast majority of legal sources open source. This would constitute a valuable resource for creating new tools which could radically reduce the cost of legal services and make English law both more accessible and competitive.

Challenge and Opportunity

The majority of case law in England and Wales is not freely available to the public. In a common law, precedent-based system where cases create binding law, this inaccessibility hinders (1) access to justice and (2) technological innovation in law (LawTech).

Legal systems are an obvious use-case for recent developments in technology like large language models (LLMs). The English legal system has several advantages that make it uniquely well placed to benefit from recent technological developments. The UK is unique in its combination of an extensive and unitary legal corpus, globally competitive legal system, willingness to allow data-mining on legal material, and advanced university sector. Despite this advantage, a crucial part of the data needed to train new AI models is locked away by private bodies. This is a wasted opportunity, and could mean that the UK falls behind international competitors in developing legal technology.

LawTech has the potential to make legal services more accessible and would further strengthen the global position of the English legal system. The UK could unlock this potential through a series of straightforward reforms to existing institutions, at minimal cost.

The immediate effect of these reforms would be to render the entire legal corpus publicly available. In the mid-term this corpus will stimulate academic research into new legal technology currently held back by the lack of usable datasets. Such research can then be developed and made commercially available by law firms, improving the accessibility, cost, speed, and global presence of English law. In the long-term, these tools will strengthen the competitiveness of English law firms, drastically lower the transaction costs imposed by the legal system across the economy, and radically improve legal access for ordinary citizens.

The UK's legal system is a significant asset to the economy, generating a turnover of [£40 billion](#), and employing [more than 333,000 individuals](#) in 2023. The UK's legal system produces some of the world's largest and most prestigious law firms, contributing to the UK's economic strength and international reputation.

However, without reform, the English legal system risks falling behind other systems that make their legal sources easily and freely accessible. The US serves as a prime example, with several states offering state-funded, public-access legal sources. If the UK fails to capitalise on

its natural advantages in LawTech, it may be outpaced by the US or other common law systems, and the entire English legal profession could lose its competitive edge in the global market.

Plan of Action

Unlocking relevant English legal datasets requires three straightforward reforms to the National Archive Find Case Law Service, a service offered by the National Archives in collaboration with His Majesty's Courts and Tribunals Service (HMCTS), an agency of the Ministry of Justice. Currently, it provides a significant, but nonetheless limited, portion of English case law online for free.

The purpose of these reforms would be to remove these limitations and to improve TNA's database for the purpose of training new legal technology. The Lord Chancellor and Secretary of State for Justice should:

1. **Change the right to republish.** This would require HMCTS to alter its contracts with the court shorthand writers responsible for producing court reports. The contracts should be changed to retain the right of HMCTS to republish judgement transcripts and make them publicly available. These reports are used by TNA in its database but the corpus is incomplete – especially in the case of the high court – because of uncertainty over the rights of use of the transcripts. Changing these contracts would not require primary legislation, but is within the purview of HMCTS.
2. **Clarify copyrights of existing judgements** The UK Government should pass a clarifying statute declaring the copyrights of judgement text to reside with their judicial authors, thereby freeing up TNA to duplicate and upload the pre-1996 case law which is currently absent from its database. Legally, the copyright status of these judgments is not formally settled, though the opinion of the legal profession is strongly in favour of the judicial-authorship interpretation.
3. **Expand remit of TNA** The MoJ should increase the remit of TNA to publish all its reports and make them, with its API, available to third parties for a variety of research uses. Currently, TNA restricts the use of its dataset to those holding special licences. The process of applying for such licences should be made as smooth and straightforward as possible, on the principle that small hurdles to access [radically reduce](#) participation rates.

Budget

The main costs of this proposal lie in Parts (1) and (2). They are unlikely to be large, but they cannot be assessed externally from the Ministry of Justice. These are:

1. The cost of renegotiating contracts with court shorthand writers. Currently, part of the cost of employing the shorthand writers is reduced by giving them rights of reuse over their transcripts. The extent to which altering their contracts would require greater remuneration lies entirely in their respective bargaining position with HMCTS.

2. If the MoJ is unwilling to take a firm stance on the copyright status of court reports, it will be necessary to buy the missing reports from private reporters to complete TNA's coverage. The cost of this will depend on the value of this relatively small slice of reports to the private reporters. The trustees of the British and Irish Legal Information Institute (BAILII), a separate non-profit which currently licences judgments from TNA and publishes them online, suggest these reports are no longer central to the business model of the private reporters.

The overall cost of these reforms would not be significant, and could be less than £5 million.

FAQs

Will this data actually be useful?

Yes. The main, and only remaining, obstacle to machine-learning research in English law is the availability of datasets. Whilst making the full corpus online is not a sufficient step for the creation of new LawTech, it is a necessary one.

Why has this not been done already?

The short answer is that there has been insufficient political will to complete TNA, and that the traditional (19th century) private-funding model of law reporting is apt in an era of extensive budget cuts to the legal system. TNA itself has been unable to complete its coverage due to a lack of funding and limits thrown up by uncertainties in the policy of HMCTS.

The traditional arguments for open-accessing the law have been to improve access to justice. Given the highly technical nature of the legal system, the main users of TNA are legal academics and blog writers. Accordingly, there has been relatively little public pressure for open-access, even if there is general public support for making the law available (and, indeed, indignation when learning of the current position). It is a central and novel part of this proposal, therefore, that in addition to these existing concerns, the novel technological and economic dimensions are emphasised.

Who are the main stakeholders?

Three key stakeholders in this process are:

1. **The National Archives**, which works in collaboration with The Ministry of Justice and the Judicial Executive Board (JEB). The National Archives, following its Open Justice Principle, will support the expansion of its holdings. The MoJ and JEB will insist on a licensing scheme for case research, but will be open to loosening it.

2. **The BAILII trustees and advisory council.** Prior to 2022, the MoJ licensed the judgments produced by the judiciary (represented collectively by the Judicial Executive Board) to the British and Irish Legal Information Institute (BAILII). In order to save money, and move control of the reports to the government, the MoJ changed this arrangement in April 2022. Following this move, BAILII provided many of its reports to TNA, and now the former licences new reports from the latter. BAILII is a non-profit public charity run by trustees. Given the stated aims of BAILII, it is extremely likely they will support an extension of TNA's database. The Chairman of BAILII, Guy Berringer CBE, has written on the importance of supporting English law's competitiveness globally.
3. **Research councils.** It is possible that the completion of TNA's database will be assisted by charitable funding, such as from the AHRC or Joint Information Systems Committee. They have funded such extensions in the past, even when the prospect for research output was less promising.
4. **The Institute of Advanced Legal Studies**, where BAILII is based, has the core mission of expanding "access to justice by publishing searchable databases of legal materials from the United Kingdom." The IALS is another potential source of funding to complete TNA's coverage.

Which groups might oppose these reforms?

Although private reporting was initially justified in terms of preventing governmental influence over the legal profession, it is unlikely the current profession will oppose extending TNA's coverage provided private reports continue to exist as an alternative. There are no significant principled arguments against making law more publicly available in general.

The main opposition will come from Butterworths and the ICLR. Their main objection will be to the second part of the proposal: the clarifying statute declaring the copyright of judgments to reside with the judge. This would allow TNA to copy the content of the ICLR and Butterworth reports freely.

Their objections will be two-fold: first, that the legal position is that they own the copyright of the judgments, making the 'clarifying statute' a piece of reform; and, second, such reform is undesirable because it will nationalise their (intellectual) property. Even if one were to concede their first point, however, the only part of the private corpus missing from TNA and still within copyright (due to the 70-year limit) is from 1953-1992. If the government was unwilling to progress with the statute, this 38-year corpus could be purchased from the ICLR or Butterworths, overcoming this obstacle albeit raising the cost of the project.

Is the government better placed to make use of the law reports than private reporters?

Yes. For two reasons:

1. First, the private reporters have little incentive to make this data available for technological innovation. They exist as a monopoly and derive most of their income from subscriptions purchased by legal professionals; consequently, they do not proactively make this data available to researchers or small start-ups. This is particularly evident in the ICLR's lacklustre attempts to produce legal AI.
2. Second, this data is incredibly valuable and will likely play a central role in the future of the legal profession. As legal systems change to accommodate these changes, the Government may want to control the use of this corpus. An exclusively private system of reporting strictly limits the possibility of such control.

What if this data is used by foreign actors?

The main concern here is that a foreign company will use the public database to train new legal tools which can be used to compete with English law firms. The most plausible candidate here is the US, a significant concern given (a) US law firms are already currently in a fierce competition with London-based firms; and (b) the American lead in machine-learning technology.

These concerns are not related per se to making the datasets public, but with the effects of subsequently created law-tech. Three basic points can be made here.

1. First, the current status quo strongly benefits well-capitalised US companies who can afford to pay high prices to the private reporters. The alternatives are not between making the reports available for producing law-tech or not, but between making them available freely, in a manner which can be controlled, or privately, in a manner which cannot.
2. Second, it is entirely possible for British start-ups to compete in producing LawTech. This could be achieved through the TNA licensing scheme – e.g., limiting the training rights of the TNA's case law database to UK-based companies. This may be unnecessary however: it is extremely unlikely legal sources alone are sufficient to create practical legal tools, and this gives incumbents with large quantities of other data – such as contract precedents and legal advice – a large advantage. Securing a British advantage here would require other policies to enable and encourage law firms to partner with British tech start-ups.
3. Third, even in the undesirable scenario that all major LawTech firms are American (if legal specialists in English law sold their data to US tech firms, for instance), this would still benefit the UK. Its legal system would become extremely efficient, attracting more business to UK law firms (even if those firms have replaced many domestic lawyers with software), drastically reducing the cost of legal services across the economy. In this third scenario, although the British economy would not directly benefit from new LawTech companies, it would reap large downstream benefits from reduced transaction costs.

What recent government initiatives and proposals have aimed to make English case law open-access?

The most recent proposals to make English cases public date to the early 2010s. BAILII was set up in 2000 following the 'Free Access to Law' movement's success in Australia. It has slowly grown since then (with a large expansion between 2005-7 following funding from the Joint Information Systems Committee and the Higher Education Funding Council).

As noted above, BAILII, and therefore TNA, is now complete for all the judgments post-1996, with the exception of a few high court cases. As such, pressure for reform has fallen off. The most recent change was in 2022 with the movement from BAILII to TNA, emphasising government control over the reports and making the reports more easily available for machine analysis. This move was also accompanied by a call for expanding TNA's coverage but, as yet, no expansion has occurred.

In the field of legislation, a more recent project '[Big Data for Law](#)' run by the National Archive was completed in 2015. This was funded by the Arts and Humanities Research Council (for around £550,000) to produce a complete database of legislation. It is worth noting that this predated the recent deep learning paradigm shift, and, in any event, only covers legislation.

Whilst there has been little recent activity in the government regarding open-sourcing the law, the 2010s reforms have left clear pathways for implementation. The MoJ is the obvious choice of department given its existing relationship with TNA, and in any event must necessarily be involved. The proposal is a natural development of the previous drive to widen access to the law; it also fits in with Government policies to increase the UK's competitive edge.

In terms of departmental capacity, MoJ will probably be unable to complete the proposals by itself: it will likely need either (a) extra funding, either from public or private sources; and/or (b) an act of legislation enabling TNA's expansion at a low-cost.

Authors

[Alec Thompson](#)

[Alec Thompson is a PhD student at the University of Cambridge specialising in legal evolution and legal complexity. He has published articles on the computability of law, legal history, and the ethics of legal natural language processing \(NLP\). He also teaches several undergraduate courses at the University of Cambridge and Oxford, such as a Roman Introduction to Private Law and Legal History.](#)