Donoghue v Stevenson Global Citation Dataset Introduction
Legal Citation Lab

TABLE OF CONTENTS

README – Donoghue v Stevenson Dataset ***1***

Introduction ***1***

Data Collection ***3***

Introduction to the Case Law Citation Lab - Donoghue v Stevenson 🍺 ***4***

README – DONOGHUE V STEVENSON DATASET

Created by the Legal Citation Lab

Date: 2026/10/10

Version: Release 2.1

Purpose
This dataset compiles judicial decisions from around the world that cite *Donoghue v Stevenson*, [1932] UKHL 100. It is intended to support research in case law citation networks, comparative tort law, and legal precedent analysis.

Data Format Notes:

- Dates are in ISO 8601 format: YYYY-MM-DD.
- Some cases include full links to CanLII, AustLII, or other databases.
- 'Citations' are aggregated strings and may include multiple reporters.

Change Log: Initial release (2025/07/08) with 5,290 cases citing *Donoghue v Stevenson*. Second release (2026/08/13) with 5,525 cases citing *Donoghue v Stevenson*. Updated second release (2026/10/10) with 5,532 cases citing *Donoghue v Stevenson*.

Release Notes: In Release 2.1, errors in case names and dates were corrected, duplicate entries were removed, hyperlinks were added or updated, and newly identified citing cases were included.

Licence: This dataset is for research and educational use. Attribution required when publishing analyses.

INTRODUCTION

For the initial project of the Legal Citation Lab, we picked *Donoghue v Stevenson*, the 1932 landmark decision of the House of Lords and likely the best-known case in the common law, at least in countries belonging to the Commonwealth of Nations. The 3:2 decision in that case heralded the modern law of negligence, particularly that of consumer products liability, and continues to be cited nearly a century later.

We used Excel to populate the dataset, starting with CSV (comma-separated value) files downloaded from Westlaw in 2021 and comprising about 2200 cases citing *Donoghue*, chiefly from Canada, the UK, Australia and New Zealand. We supplemented this with electronic searches from 17 other proprietary databases, 33 Legal Information Institute collections; 130 court websites or other public-domain judgment portals; 90 print report series (not all of which were complete) and associated indexes or digests; and several secondary

sources such as books and articles. As of mid-2026, the dataset had grown to over 5500 cases, with more being added monthly, albeit at a slower rate.

In the resulting spreadsheet, the citing cases are the unit of analysis; in other words, each case occupies one row of the spreadsheet, and the primary entry in a row is the case name. This includes individual instances of a case; if it proceeds through appeal, each instance is treated as a separate case for the purpose of the dataset.

However, there are many pieces of information about a case beyond its name, such as the date, jurisdiction, level of court, and citations to law reports in which the citing case appears. In a study that analyzed over 16,000 Israeli Supreme Court cases of all types over a specific period, the dataset had 61 variables, which gives a sense of how much information could be collected if one were inclined to do so (Weinshall & Epstein, 2020). The question becomes what additional data is manageable and useful for conducting citation analysis. For the initial dataset, we created fields for 16 data items which we have narrowed down to 8 fields for the purpose of this release. These fields are explained below:

Case name: This is the style of cause of the citing case, rendered, where possible, in a recognized format for case names (e.g., McGill Guide, Bluebook or OSCOLA) and using the *sub nom* technique to identify potentially competing case names. An example is an Ontario Court of Appeal decision in which one defendant applied to add another in a claim relating to the failure of a car's brake system. The name for the underlying case is *Stewart v Domingos*, and this is how both CanLII and LexisNexis cite the appeal decision that allowed the applicant to add the proposed defendant, one Heppel. Westlaw, on the other hand, cites the case as *Heppel v Stewart*, and this is also how the case was named on further appeal, where the Supreme Court of Canada reversed the decision. The case is therefore entered as *Heppel v Stewart*, sub nom *Stewart v Domingos*. In both Westlaw and CanLII, the same case turned up more than once owing to different case names, and the above method was used to eliminate the duplication.

Decision Date: This is the date of the citing case converted to ISO 8601 format (e.g., 2015-07-23). This allows for the creation of time-series data and other date-related analysis. In some cases, the date differs from the original Westlaw dataset. For example, a court may have rendered a decision on a motion or a trial as well as a later decision on costs or a variation of the original order. Westlaw or other services may use the later of these dates as the reported date on the basis that this was the decision that brought the case to a close. But for determining whether *Donoghue* played a role in the decision, we have generally used the earlier date. An example is *Alberta Caterers Ltd v R Vollan (Alberta) Ltd.* Westlaw shows the date as December 8, 1977, but that is the date on which the costs decision was varied: (1977) 11 AR 181. In the dataset, we have changed the date to 28 October 1977 (1977-10-28), that being the date of the judgment: 10 AR 501, 1977 CanLII 1854 (AB QB). Where the year of a decision is available but not the precise date within the year, the dataset uses the year and July 1 as the median date within the year, until such time as the precise date can be ascertained. For the Israeli cases, we converted the date from the Hebrew calendar to the Gregorian calendar if that version of the date had not already been provided in the case.

Citations: This field is for parallel citations to the case, including a neutral citation or identifier from an LII or other non- proprietary database, or an entry in a recognized format for unreported cases, with each entry separated by a comma. For example, the three citations to *Heppel v Stewart* show as 62 DLR (2d) 282, [1967] 2 OR 37, 1967 CanLII 29 (ON CA).

Primary Jurisdiction: For federal countries, this field holds the country name—e.g., Canada, Australia, USA. It also holds the name for those countries or regions that, while not federations, are comprised of political units within the country or region— e.g., the UK or the Pacific Islands.

Secondary Jurisdiction: This field holds the sublevels of the entries under Jurisdiction (1). For example, for Australia, it comprises Australian Capital Territory, New South Wales, Northern Territory, Queensland, South Australia, Tasmania, Victoria and Western Australia, as well as "Commonwealth" (i.e., for cases heard in federal courts). The same principle applies to Canada and the USA, with the designation "Federal" indicating

the jurisdiction of the federal courts. For countries or regions that, while not federations, are comprised of individual political units, this field holds the names of those units. For example, for the UK, the dataset uses three sub-levels: (i) England and Wales (typically considered together for legal citation purposes), (ii) Scotland, and (iii) Northern Ireland.

Court: This field holds the abbreviated name of the court that decided the citing case; e.g. SCC (Supreme Court of Canada).

Court Full Name: This field holds the full name of the court that decided the citing case; e.g. Supreme Court of Canada.

Primary Link: This field is for a uniform resource locator ("URL" or "hyperlink") to the citing case, where a publicly accessible URL exists.

DATA COLLECTION

The first stage of the project is to collect all available cases in which *Donoghue* is cited, directly or indirectly, by a court or tribunal when rendering a decision. This excludes cases in which *Donoghue* is cited by counsel in argument (where the report of the decision lists such cases) but is not mentioned in the reasons for decision. Similarly, where a report's editors annotate the case and include a reference to *Donoghue*, that case does not form part of the dataset unless the court also references *Donoghue*. To that extent, the dataset may be seen as under-inclusive, since citation by counsel or by law report editors is still evidence of *Donoghue*'s reach, as is citation by scholars in law review articles and books. However, the project focus is on judicial citation analysis, a field in existence for at least 150 years (Ogden, 1993), and for which methodologies have been developed through various studies, especially in the US but also in Canada (Alschner et al, 2024; Fournier, 2021, 2023; Neale, 2013; Warchuk, 2024). For something to count as a judicial citation, then, the criterion adopted for this project is that there must be some evidence of a court having considered *Donoghue*, even if only in a fleeting or formulaic way, or even if only declaring that *Donoghue* did not apply to the facts at hand.

On the other hand, the dataset may be seen as over-inclusive because it includes cases in which *Donoghue* does not appear in a list of cited cases at the beginning or end of a judgment, or in a table of cases judicially considered in a print reporter or citator. Often, such lists or tables do not show *Donoghue* as one of the cases cited in a law report, despite it being directly cited in one or more judgments in the report. In other judgments, *Donoghue* appears in a citation to another cited case. An example is Lord Wilberforce's speech in *Anns v Merton London Borough Council*, [1977] UKHL 4, referencing the "trilogy" of House of Lords cases for the proposition that a prima facie tort law duty of care exists in certain situations. That passage has been widely cited and is often followed by a discussion of the principles of proximity and foreseeability as expressed in *Donoghue* and other cases. Including such sub-citations casts a broader net than methodologies that return results based only on whether a case lists *Donoghue* as having been applied, followed, considered, distinguished, doubted or overruled (the terminology varies according to the methodology). The idea here is that to exclude *Donoghue* because it is cited only indirectly or because it has been distinguished or held not to apply would risk missing judicial treatment of the principles promulgated in the case. See, e.g., Landes et al, 1998 at 273 (in measuring influence, deciding not to differentiate among favourable, critical or distinguishing citations). This is also the philosophy behind AustLII's LawCite project and is the approach used when noting up *Donoghue* in CanLII.

Therefore, as well as using the noting-up features of various legal databases, we also performed Boolean or text searches (e.g., "Donoghue NEAR Stevenson" or "[1932] AC 562") and reviewed and reconciled the results with those obtained from citation searches using print resources or other digital resources that were available.

A research project of this scope—spanning the world and dating back to 1932—faces many challenges of access and coverage. Much of the world's caselaw is still behind a proprietary paywall, and where public

access exists through non-proprietary databases, coverage may be sparse. For instance, Singapore's courts have been citing *Donoghue* since at least 1973, but a search on the Legal Information Institute databases yields only a few decisions dating from 2000 or later, depending on the court level. Through the generosity of the Singapore Academy of Law, we were able to narrow the gap with short-term complimentary access to its proprietary database, allowing us to generate a set of nearly 60 decisions from all levels of the Singapore courts. We had similar help from database providers in India and Africa.

Print reports have also been valuable, particularly for cases predating the 1990s, and work is still underway to identify and access relevant reports. In some cases, access to this law is now being addressed through the scanning and uploading of early print reports. For example, the Asian Legal Information Institute has uploaded scanned versions of early law reports from India and Burma (as it was then called), leading to the discovery of citing cases unavailable through other sources. The uploading of early print reports, and in some cases, typewritten judgments from court files, is a development that, while proceeding slowly, should aid those doing research such as is entailed by this project.

The lesson for those conducting citation analysis is that one must be creative and combine multiple search methods over multiple sources to ensure robust results when collecting cases.

INTRODUCTION TO THE CASE LAW CITATION LAB - DONOGHUE V STEVENSON 🐌

Full Case Name	M'Alister (or Donoghue) (Pauper) v Stevenson
Date	1932-05-26
Citations	[1932] UKHL 100; [1932] All ER 1; [1932] SC (HL) 31; [1932] SLT 317; (1932) 37 Com Cas 350; (1932) 48 TLR 494; (1932) 147 LT 281; [1932] Sol Jo 396; (1932) 101 LJPC 119; [1932] SC 31; [1932] WN 139; [1932] AC 562
Court	House of Lords (UK)
Majority	Lord Atkin; Lord Thankerton; Lord Macmillan
Minority	Lord Buckmaster; Lord Tomlin
Primary Link	https://www.bailii.org/uk/cases/UKHL/1932/100.pdf
Secondary Link	https://canlii.ca/t/gghx0
Summary	The case of *Donoghue v Stevenson*, [1932] AC 562 (HL), is a foundational decision in tort law that established the modern doctrine of negligence. The case arose when May Donoghue fell ill after consuming ginger beer that allegedly contained a decomposed snail. Since a friend had purchased the drink, Donoghue had no contractual relationship with either the café owner or the manufacturer, a factual matrix that generally was thought to bar a legal claim at the time. Indeed, her claim was initially dismissed in the Scottish Court of Session, but she appealed to the House of Lords on a question of law, that being whether the manufacturer of the drink could owe a duty of care to the ultimate consumer despite the lack of a contractual relationship. In a 3–2 decision (Lords Atkin, Thankerton, and Macmillan concurring; Lords Buckmaster and Tomlin dissenting), the Court held that the manufacturer, David Stevenson, owed a duty of care to the consumer. This judgment articulated two key principles. The narrow principle established that a manufacturer owes a duty to ensure that products intended to reach consumers without possibility of intermediate inspection are free from harmful defects (ibid at 599–600, per Lord Atkin). More significantly, Lord Atkin expressed what came to be known as the "neighbour principle," stating: "You must

	take reasonable care to avoid acts or omissions which you can reasonably foresee would be likely to injure your neighbour," defining neighbours as those "closely and directly affected by my act" (ibid at 580). This principle laid the groundwork for the modern law of negligence and was later built upon in cases such as *Anns v Merton London Borough Council*, [1978] AC 728 (HL), *Caparo Industries plc v Dickman*, [1990] 2 AC 605 (HL), and *Cooper v Hobart*, 2001 SCC 79.
Additional Resources	https://www.scottishlawreports.org.uk/resources/donoghue-v-stevenson/