

Contracts Finder Records

A Guide to the Dataset

August 2026

1. Overview

This dataset provides contract-level records of public contracts awarded through [Contracts Finder](https://www.contractsfinder.service.gov.uk/) (<https://www.contractsfinder.service.gov.uk/>), the UK government's portal for public sector procurement opportunities and awards, enriched with the project's classification of who won each contract. It was produced by the [UK Third and Civil Society Sector Database](https://uk-third-sector-database.github.io/) (<https://uk-third-sector-database.github.io/>) project, which collects, processes, and links public administrative data on civil society organisations throughout the UK.

Unlike the project's payments dataset, this file covers contracts awarded to *all* suppliers — private companies, public bodies and individuals as well as civil society organisations. Civil society suppliers (registered charities, community interest companies, co-operatives and mutuals, and other nonprofit companies) are identified by matching the supplier name to the TCSS Organisation Register; where a match is found the record carries the organisation's `uid`. Retaining the non-civil-society awards is what makes it possible to measure civil society's share of the public contract market and to test the reliability of the buyer's own voluntary-sector flags.

The dataset contains 406,113 awarded contract notices from 4,687 public buyers, naming 106,747 distinct suppliers. 19,591 awards (4.8%) went to 5,222 register-confirmed civil society organisations. Coverage is effectively continuous from 2015, when the current version of Contracts Finder was launched, to early 2026.

406,113

Awarded Contracts

106,747

Suppliers

4,687

Buyers

19,591

Awards to CSOs

Related dataset. This guide covers the **Contracts Finder** dataset (`contractsfinder-enriched-refined.csv`): one row per awarded contract notice, all suppliers. The project also publishes a **payments** dataset (`cso-procurement-dataset-refined.csv`): one row per payment (or Contracts Finder award) made to a register-confirmed civil society organisation, drawn from central government

transparency spending data, NHS payment records and Contracts Finder. Use this dataset to study the contract market — what was tendered, who won it, what it was for, and civil society’s share; use the payments dataset to measure money actually paid to civil society organisations over time. The payments dataset has its own guide: [Procurement Payments — A Guide to the Dataset](https://uk-third-sector-database.github.io/guidance/tcss-procurement-payments-guidance.html) (<https://uk-third-sector-database.github.io/guidance/tcss-procurement-payments-guidance.html>).

2. What is Contracts Finder?

Contracts Finder is the UK government’s free online service for publishing public sector contract opportunities and awards in England (and some UK-wide and devolved-administration notices). Since February 2015, under the Public Contracts Regulations 2015, public bodies have been required to publish on Contracts Finder any opportunity above a low-value threshold — currently £12,000 including VAT for central government and £30,000 for the wider public sector (£10,000 and £25,000 excluding VAT before 2022) — together with the subsequent award notice giving the winning supplier, the awarded value and the contract dates. A small number of older awards were back-published “for transparency purposes” when the current portal was launched and carry award dates before 2015.

How the data were collected

Awarded contract notices are retrieved from the Contracts Finder API by the project’s own [Contracts Finder collector](https://github.com/uk-third-sector-database/contracts_finder) (https://github.com/uk-third-sector-database/contracts_finder), which pages through the published notices and stores each award with its full metadata. Every record in this dataset has `noticeType = Contract` and `noticeStatus = Awarded`; opportunities that had not yet been awarded at collection time are not included. 20.6% of records (83,704) were added by a later “gap-fill” collection pass (`is_gap = TRUE`) that retrieved notices missed by the initial run.

How the data were enriched

Two kinds of enrichment are added to the raw notices (see [Appendix A1](#)). First, the buyer is classified into a funder type (UK Government, Local Government, NHS, and so on) using the same classification cascade as the payments dataset. Second, the supplier name is normalised and matched to the TCSS Organisation Register; where a civil society organisation is found the record receives its `uid` and a `supplier_type` (Charity, CIC, Co-operative/Mutual or Other CSO). Suppliers that do not match the register are labelled `Non-CSO`, and awards naming several suppliers at once, or carrying an unusable supplier string, are labelled `Excluded` and not matched.

3. Dataset Contents

The dataset is supplied as a single CSV file, `contractsfinder-enriched-refined.csv`, containing 406,113 records across 52 fields. Each row is a single awarded contract notice. The fields are organised into the groups below. Field names in lower camel case (e.g., `awardedValue`) come directly from the Contracts Finder API; fields in snake case (e.g., `supplier_type`) are added by the project.

Notice & Contract Details

Field	Description	Type	Coverage
<code>id</code>	Contracts Finder internal notice identifier (unique in this dataset)	Text	100%
<code>noticeIdentifier</code>	Contracts Finder notice reference number, as shown on the portal. A small number of notices share a reference (2,063 rows); use <code>id</code> as the unique key	Text	>99.9%

Field	Description	Type	Coverage
parentId	Identifier of the parent notice (for call-offs under a framework or related notices)	Text	10.9%
isSubNotice	Whether the record is a sub-notice of a parent notice (True/False)	Boolean	100%
noticeType	Type of notice — always <i>Contract</i> in this dataset	Text	100%
noticeStatus	Status of the notice — always <i>Awarded</i> in this dataset	Text	100%
dept	Publishing department/buyer label from the source portal	Text	100%
organisationName	Buyer organisation name as published on the notice (not standardised; see <code>funder_name</code>)	Text	100%
title	Contract title	Text	>99.9%
description	Contract description	Text	>99.9%
sector	Sector tag, where provided by the buyer	Text	8.2%
cpvDescription	Primary Common Procurement Vocabulary (CPV) category description	Text	100%
cpvDescriptionExpanded	Expanded CPV category description(s)	Text	100%
cpvCodes	CPV code(s) for the contract	Text	100%
cpvCodesExtended	Extended CPV code list	Text	100%

Buyer (Funder)

Field	Description	Type	Coverage
funder_name	Canonical (standardised, uppcase) buyer name assigned by the project; consistent with the payments dataset	Text	99.9%
funder_name_alt	Alternative buyer name, where applicable — typically the parent central government department of an agency (e.g., DVSA mapped to DFTRANSPORT)	Text	9.7%
funder_id	Unique buyer identifier (e.g., F01726); consistent with the payments dataset	Text	99.9%
funder_type	Buyer classification: UK Government, Local Government, NHS, Education Institution, CSO, Police, Fire and Rescue, Private Sector, Other, or Junk/Invalid (see Appendix A2)	Text	99.9%
funder_type_alt	Classification of the alternative buyer name	Text	15.2%
funder_type_detail	A more granular buyer classification, refining <code>funder_type</code>	Text	99.9%

Supplier & Civil-Society Classification

Field	Description	Type	Coverage
awardedSupplier	Supplier name(s) as recorded on the award notice; several names separated by commas where a notice records more than one winner	Text	99.8%
normalized_supplier	Normalised supplier name used for register matching (uppercased, punctuation removed, whitespace collapsed)	Text	99.8%
supplier_id	Internal supplier identifier (e.g., S017036); consistent with the payments dataset	Text	99.9%
supplier_type	Supplier classification: <i>Charity, CIC, Co-operative/Mutual, Other CSO, Non-CSO</i> (did not match the register), or <i>Excluded</i> (multi-supplier or unusable supplier string, not matched)	Text	99.9%
supplier_type_detail	A more granular supplier classification, refining <code>supplier_type</code> (e.g., <i>Private Sector, Education, NHS, multi_supplier</i>)	Text	99.8%
uid	TCSS Organisation Register identifier; present only where the supplier matched a register entry, i.e. is a confirmed civil society organisation. This is the recommended indicator of a civil society award	Text	4.8%

Field	Description	Type	Coverage
organisation_name	Matched organisation name from the register	Text	4.1%
is_multi_supplier	Whether the award named more than one supplier (TRUE/FALSE)	Boolean	100%

Values & Dates

Field	Description	Type	Coverage
awardedValue	Awarded contract value (£) as entered by the buyer. Zero for 16,193 records (4.0%); see Section 6 for outliers	Numeric	100%
valueLow	Lower bound of the estimated contract value (£)	Numeric	100%
valueHigh	Upper bound of the estimated contract value (£)	Numeric	100%
publishedDate	Date and time the notice was published (ISO 8601, e.g., 2019-03-01T00:00:00Z)	Date-time	100%
approachMarketDate	Date the buyer approached the market, where recorded	Date-time	0.6%
deadlineDate	Submission deadline date	Date-time	>99.9%
awardedDate	Date the contract was awarded — the recommended date for time-series analysis	Date-time	>99.9%
start	Contract start date	Date-time	>99.9%
end	Contract end date	Date-time	>99.9%
lastNotifiableUpdate	Date of the last notifiable update to the notice	Date-time	100%

Voluntary-Sector & SME Flags

Field	Description	Type	Coverage
isSuitableForVco	Buyer flag — the contract was advertised as suitable for voluntary and community organisations (True/False)	Boolean	100%
awardedToVcse	Buyer flag — the contract was recorded as awarded to a voluntary, community or social enterprise (True/False)	Boolean	>99.9%
isSuitableForSme	Buyer flag — the contract was advertised as suitable for small and medium-sized enterprises (True/False)	Boolean	100%
awardedToSme	Buyer flag — the contract was recorded as awarded to an SME (True/False)	Boolean	>99.9%

Note: The four flags above are self-reported by the buyer and are not validated against any register. They are an unreliable guide to which contracts actually went to civil society organisations: across the whole dataset only 11% of contracts flagged “suitable for” the voluntary sector, and 34% of those flagged “awarded to” it, went to a register-confirmed civil society organisation — and the “awarded to” flag captures only 16% of the awards that did go to one. The register-matched `uid` and `supplier_type` fields provide the validated classification; see [Section 4](#) for the full comparison.

Geography

Field	Description	Type	Coverage
postcode	Postcode entered on the notice (usually the contract delivery or buyer location, not the supplier’s registered address)	Text	32.4%
coordinates	Latitude, longitude pair associated with the notice; 0,0 (no location) for 69.3% of records	Text	100%
region	Region code(s) for the contract; may list several regions, or <i>Any region</i>	Text	87.5%
regionText	Region name(s) corresponding to <code>region</code>	Text	87.5%

Linkage Fields

Field	Description	Type	Coverage
is_gap	Indicates the record was retrieved by the gap-fill collection pass (TRUE/FALSE)	Boolean	100%
payment_year	Year field carried over for cross-dataset linkage (see note)	Numeric	>99.9%
payment_date	Date field carried over for cross-dataset linkage (see note)	Text	>99.9%
total_value_payments_to_org	Total value of payments to the matched supplier across the payments dataset (£); populated for all rows but meaningful only for register-matched suppliers	Numeric	100%
total_number_payments_to_org	Total number of payments to the matched supplier across the payments dataset	Numeric	100%

Note: For contract-level analysis, use the contract dates — `awardedDate` , `publishedDate` , `start` and `end` .The `payment_year` and `payment_date` fields are derived from `awardedDate` during the linkage process to align with the payments dataset and should not be treated as additional information; a handful of records carry implausible values (e.g., 24, 1898, 2108) inherited from the source award date. Records can be linked to the payments dataset and the TCSS Organisation Register using `supplier_id` and (for matched civil society organisations) `uid` .

4. Coverage & Completeness

Core Field Coverage

The contract fields from the portal — identifiers, title, description, CPV codes, values, dates and flags — are present in every record. The only exceptions are 74 records with no `awardedDate` and a few with no supplier name. The project's buyer classification is missing for a small number of records whose buyer name could not be resolved (`funder_id` for 583, 0.1%; `funder_type` for 390), and `supplier_type` for 464 records. The `uid` is present only for the 4.8% of awards won by a register-confirmed civil society organisation; geographic fields are the least complete.

id	awardedValue	funder_id
100.0%	100.0%	99.9%
supplier_type	regionText	postcode
99.9%	87.5%	32.4%
uid (CSO awards only)		
4.8%		

Who Buys: Buyer Type Distribution

Central government departments and agencies (37% of awards) and local authorities (35%) between them publish the large majority of notices. The picture for awards to civil society organisations is different: local government is the largest buyer from civil society (46% of CSO awards), followed by UK Government (32%) and the NHS (9%).

Buyer Type	All awards	Share	Awards to CSOs	Share of CSO awards
UK Government	148,977	36.7%	6,200	31.6%
Local Government	143,085	35.2%	9,109	46.5%
NHS	37,241	9.2%	1,696	8.7%
Education Institution	31,569	7.8%	922	4.7%
CSO	16,638	4.1%	513	2.6%
Police, Fire and Rescue	10,337	2.5%	376	1.9%
Private Sector	9,026	2.2%	353	1.8%
Other	6,827	1.7%	241	1.2%
Junk/Invalid	2,023	0.5%	168	0.9%
(not classified)	390	0.1%	13	0.1%

Who Wins: Supplier Type Distribution

Most awards go to suppliers that do not match the TCSS Organisation Register (`Non-CSO`) — overwhelmingly private companies, with some public bodies and universities. 10,085 awards (2.5%) are `Excluded` from matching because the notice names several suppliers at once (9,708) or carries an unusable supplier string (386). Among civil society winners, charities dominate.

Supplier Type	Awards	Share	Distinct suppliers
Non-CSO	375,973	92.6%	91,927
Charity	13,328	3.3%	3,454
Excluded	10,085	2.5%	9,713
Other CSO	3,592	0.9%	914
CIC	1,493	0.4%	612
Co-operative/Mutual	1,178	0.3%	253
(not classified)	464	0.1%	—

Note: From the August 2026 release the `supplier_type` label and the `uid` match are consistent by construction (Spine-only rule): every row typed Charity, CIC, Co-operative/Mutual or Other CSO carries a `uid`, and every row with a `uid` carries one of those types. Counts in this guide use the `uid`. Earlier releases had 1,546 CSO-typed rows without a `uid` and 166 the reverse.

Which Registers the Civil Society Winners Come From

Among the 19,591 awards to register-confirmed civil society organisations, the `uid` prefix shows the source register. Registered charities in England and Wales account for 66% of CSO awards; nonprofit companies registered only at Companies House (community interest companies, companies limited by guarantee) for 26%.

Prefix	Source Register	Awards to CSOs	Distinct organisations	Share
GB-CHC	Charity Commission for England & Wales	12,957	3,347	66.1%
GB-COH	Companies House	5,073	1,518	25.9%
GB-MPR	Mutuals Public Register (FCA)	754	224	3.8%
GB-COOP	Co-operatives UK	424	29	2.2%
GB-SC	Office of the Scottish Charity Regulator	337	92	1.7%
GB-NIC	Charity Commission for Northern Ireland	46	12	0.2%

Year-by-Year Award Counts

Award counts by `awardedDate` year. Coverage is effectively continuous from 2015, when the current Contracts Finder service launched and publication became mandatory, to early 2026 (collection date). The 1,666 awards dated before 2015 pre-date the current service (they were migrated from its predecessor or back-published for transparency) and should not be read as full coverage of those years; 39 records carry award dates after 2026 (source data errors), 147 fall outside 2010–2026 altogether, and 74 have no award date.

Award year	All awards	Awards to CSOs	CSO share
2010	32	3	9.4%
2011	91	6	6.6%
2012	108	1	0.9%
2013	214	15	7.0%
2014	1,113	65	5.8%
2015	14,627	733	5.0%
2016	22,093	1,054	4.8%
2017	25,481	1,325	5.2%
2018	28,787	1,404	4.9%
2019	32,667	1,550	4.7%
2020	34,283	1,568	4.6%
2021	46,914	2,345	5.0%
2022	52,927	2,741	5.2%
2023	55,785	2,959	5.3%
2024	52,817	2,386	4.5%
2025	37,214	1,392	3.7%
2026	738	24	3.3%

Buyer Flags versus Register Match

Each notice carries four buyer-entered flags. The table compares them with the register match: how many flagged contracts actually went to a register-confirmed civil society organisation (precision), and what share of all CSO awards each flag captures (recall). The voluntary-sector flags are poor on both counts — the “suitable for VCO” flag is mostly attached to contracts won by other suppliers, and the “awarded to VCSE” flag is set on only 9,390 notices and misses most awards that did go to civil society.

Flag	Awards flagged	Share of all awards	...of which register-matched CSO	CSO share among flagged	Share of all CSO awards captured
isSuitableForVco = True	79,721	19.6%	9,127	11.4%	46.6%
awardedToVcse = True	9,390	2.3%	3,152	33.6%	16.1%
isSuitableForSme = True	272,403	67.1%	14,355	5.3%	73.3%
awardedToSme = True	138,006	34.0%	6,876	5.0%	35.1%

What Civil Society Organisations Are Contracted For

The primary CPV category gives a broad-brush view of the services procured. Health and social work and education and training services together account for the largest share of awards to civil society organisations.

CPV category (primary)	Awards to CSOs	Share of CSO awards
Health and social work services	2,886	14.7%
Education and training services	2,136	10.9%
Research and development services and related consultancy services	619	3.2%
Health services	518	2.6%
Other community, social and personal services	474	2.4%
Social work and related services	341	1.7%
IT services: consulting, software development, Internet and support	286	1.5%
Software package and information systems	273	1.4%
Business services: law, marketing, consulting, recruitment, printing and security	257	1.3%
Health and social work services and Other community, social and personal services	244	1.2%

Largest Buyers from Civil Society

The buyers awarding the most contracts to register-confirmed civil society organisations, using the standardised `funder_name`.

Buyer (<code>funder_name</code>)	Buyer	Awards to CSOs	Distinct CSOs
MOJUST	Ministry of Justice	919	213
DEFRA	Department for Environment, Food & Rural Affairs	715	197
NHS_ENGLAND	NHS England	689	394
LEEDS_CITY_COUNCIL	Leeds City Council	457	192
DFEDUCATION	Department for Education	406	177
BRISTOL_CITY_COUNCIL	Bristol City Council	364	149
UK_SHARED_BUSINESS_SERVICES_LIMITED	UK Shared Business Services	300	140
MODEF	Ministry of Defence	278	134
DARLINGTON_BOROUGH_COUNCIL	Darlington Borough Council	273	88
UNIVERSITY_OF_THE_WEST_OF_ENGLAND	University of the West of England	236	65

Buyer (funder_name)	Buyer	Awards to CSOs	Distinct CSOs
DFTRANSPORT	Department for Transport	216	63
DOHEALTH	Department of Health & Social Care	202	102
HAMPSHIRE COUNTY COUNCIL	Hampshire County Council	196	88
CITY OF LONDON COUNCIL	City of London Council	169	61
UKRI	UK Research & Innovation	161	58

Contract Values

Awarded values are entered by buyers and are highly skewed; a few framework agreements carry headline values in the tens of billions of pounds, and 4.0% of awards record a value of zero. Medians are therefore more informative than totals. The total awarded to civil society organisations is £33.9 billion, of which 3 awards valued at £1 billion or more contribute a large share; excluding those, the total is £25.7 billion.

	25th percentile	Median	75th percentile	90th percentile	Total (£bn)
All awards	£31,199	£80,000	£296,895	£1,541,570	2,856
Awards to CSOs	£35,771	£90,000	£344,646	£1,497,751	33.9

Geography

Geographic fields describe the contract (usually where it is delivered or the buyer's location), not the supplier. 12.5% of records have no region, 17.8% are tagged *Any region* and 20.5% list more than one region; a postcode is present for 32.4%. For supplier location, link to the TCSS Organisation Register on `uid` (see [Appendix A4](#)).

regionText	Awards	Share
(more than one region listed)	83,056	20.5%
Any region	72,249	17.8%
(missing)	50,631	12.5%
London	29,499	7.3%
Yorkshire and The Humber	28,479	7.0%
South East	26,042	6.4%
South West	25,306	6.2%
North East	17,959	4.4%
North West	17,905	4.4%
West Midlands	17,362	4.3%
East of England	16,359	4.0%
East Midlands	16,162	4.0%
Wales	2,453	0.6%
Scotland	2,229	0.5%
Northern Ireland	196	<0.1%

5. What Can You Learn?

Because the dataset retains every awarded notice, not only those won by civil society, it supports questions about civil society's place in the public contract market as a whole.

Research Questions

- **Market share** — What share of public contracts, by number and value, is awarded to civil society organisations, and how does this vary by buyer type, service category (CPV) and over time?
- **Who commissions civil society** — Which departments, councils and NHS bodies award the most contracts to charities, CICs and co-operatives? Are there buyers that never do?
- **What is commissioned** — Using the CPV codes, titles and descriptions, what kinds of services are civil society organisations contracted to deliver, and how does this differ from the services won by private companies?
- **Tendering practice** — Do contracts flagged as suitable for the voluntary sector or for SMEs actually go to those suppliers? How long do buyers allow between publication and deadline, and does this differ for contracts civil society wins?
- **Text analysis** — The title and description fields support classification of contracts by policy area, beneficiary group or delivery model using text-mining or large-language-model methods.
- **Cross-dataset linkage** — By linking on `uid` to the TCSS Organisation Register and financial datasets, researchers can compare contract winners with the wider civil society population in size, age, location and activity, and relate contract income to organisations' accounts.

Example: Testing the Buyer Flags

A simple first analysis is to ask whether the `awardedToVcse` flag can be trusted. Of the 9,390 awards flagged as won by a voluntary, community or social enterprise, 34% match a register-confirmed civil society organisation; conversely, of the 19,591 awards that did go to a register-confirmed organisation, only 16% carry the flag. Researchers relying on the portal's own flags would therefore both over-count awards to some suppliers and miss the great majority of genuine civil society awards. Section 4 gives the full comparison for all four flags.

Tip: This dataset can be linked to the [TCSS Organisation Register \(https://uk-third-sector-database.github.io/guidance/tcss-organisation-register-guidance.html\)](https://uk-third-sector-database.github.io/guidance/tcss-organisation-register-guidance.html) using the `uid` field, enabling enrichment with organisation characteristics such as location, registration dates, and industrial classification codes. See [Appendix A4](#) for worked examples.

6. Limitations & Caveats

Publication Thresholds and Scope

Contracts Finder captures contracts above the publication thresholds (currently £12,000 for central government and £30,000 for the wider public sector, including VAT) awarded by English and some UK-wide buyers. Smaller contracts, grants, and most procurement by the Scottish, Welsh and Northern Irish public sectors (which use their own portals) are not covered. Compliance with the publication duty is incomplete, so the dataset understates the true number of public contracts.

Awarded Values

The `awardedValue` is entered by the buyer and is not audited. Values are sometimes the total ceiling of a multi-supplier framework rather than the amount any one supplier will receive (the largest recorded value is £150 billion), sometimes zero or a placeholder, and occasionally in the wrong units. Sum values with care; use medians or trimmed totals, and treat the 16,193 zero-value awards as missing.

Supplier Matching

Supplier names were matched to the TCSS Organisation Register using exact and fuzzy name matching (Jaro-Winkler similarity, threshold ≥ 0.90) followed by alias resolution (see [Appendix A3](#)). Some false positives (incorrect matches) and false negatives (missed matches) are possible, particularly for organisations with common or ambiguous names, trading names that differ from the registered name, or awards to consortia. Awards naming several suppliers (24,196 records, `is_multi_supplier = TRUE`) are not matched, so a civil society organisation that won a place on a multi-supplier framework will not be identified.

Buyer Classification

Buyers are classified using a multi-layer cascade of metadata signals, external lookups, and keyword rules (see [Appendix A2](#)). The `organisationName` field is free text and the same buyer appears under many spellings (e.g., “Ministry of Justice” and “Ministry of Justice.”); use the standardised `funder_name` / `funder_id` for grouping. Edge cases may be misclassified, and 0.1% of records have no buyer classification at all.

Duplicate and Related Notices

10.9% of records are sub-notices of a parent notice (call-offs under a framework, lots, or re-published notices); a contract may therefore appear more than once. Use `parentId` and `isSubNotice` to identify these, and `id` as the unique key.

Dates Outside the Core Range

Award dates range from 1898 to 2108 because of legacy back-publication and source data errors. For time-series analysis, filter on `awardedDate` to 2015–2025 and note that 2026 is a partial year.

Geography

The postcode, coordinates and region fields describe the notice, not the supplier, and are incomplete (69.3% of coordinates are `0,0`). Do not use them to locate civil society organisations; link to the Organisation Register instead.

What’s NOT in the Data

The dataset does not include:

- Opportunities that had not been awarded at the time of collection, or notices of other types (pipeline, early engagement)
- Contracts below the publication thresholds, or procurement published only on Public Contracts Scotland, Sell2Wales or eTendersNI
- Actual payments made under each contract (see the [payments dataset](https://uk-third-sector-database.github.io/guidance/tcss-procurement-payments-guidance.html) (<https://uk-third-sector-database.github.io/guidance/tcss-procurement-payments-guidance.html>))
- Bidder lists, unsuccessful tenders or award criteria
- Supplier characteristics beyond the register match (link to the Organisation Register for these)

7. Citation & Licence

Licence: This dataset is licensed under the [Creative Commons Attribution 4.0 International License \(CC BY 4.0\)](https://creativecommons.org/licenses/by/4.0/) (<https://creativecommons.org/licenses/by/4.0/>). You are free to share, adapt, and build upon this data for any purpose, provided you give appropriate credit. The underlying notices are published by HM Government on Contracts Finder under the [Open Government Licence v3.0](https://www.nationalarchives.gov.uk/doc/open-government-licence/version/3/) (<https://www.nationalarchives.gov.uk/doc/open-government-licence/version/3/>).

Suggested Citation

McDonnell, D. et al. (2026). *TCSS Contracts Finder Records*. UK Third and Civil Society Sector Database. Available at: <https://uk-third-sector-database.github.io/data/> (<https://uk-third-sector-database.github.io/data/>). Licensed under CC BY 4.0.

If you would like to learn more about this dataset and how it can be applied to your project or research programme, please contact research@brawdata.com (<mailto:research@brawdata.com>).

8. Changelog

Version	Date	Changes
1.1	August 2026	August 2026 data release (406,113 notices, unchanged). Supplier classification now follows the Spine-only rule: <code>supplier_type</code> and <code>uid</code> are consistent on every row (1,546 CSO-typed rows without a <code>uid</code> demoted to Non-CSO; 13 charities recorded under compound register identifiers, e.g. British Red Cross, RNIB, Sense, NatCen, now classified Charity); 19,591 awards to 5,222 register-confirmed organisations. Supplier name matching also recovered notices whose supplier names contain apostrophes. Coverage tables re-derived. <code>supplier_id</code> values are re-minted each release; use <code>uid</code> to link across releases.
1.0	August 2026	First standalone guide for the Contracts Finder Records dataset (June 2026 data release, 406,113 records). The dataset was previously documented as Section 3B of the combined Procurement Records guide (v1.1, June 2026); this guide adds full coverage tables, the buyer-flag comparison, research uses, limitations and linkage examples.

A1. Pipeline Overview

The dataset is produced in three stages.

Stage	Component	Purpose
1	Contracts Finder collector (contracts_finder (https://github.com/uk-third-sector-database/contracts_finder))	Retrieves awarded contract notices from the Contracts Finder API, including an initial full collection and a later gap-fill pass (<code>is_gap</code>), and stores each notice with its full metadata.
2	Buyer classification (<code>02-build-funder-lookup.R</code>)	Standardises the buyer name to <code>funder_name</code> / <code>funder_id</code> and classifies it into a funder type using the multi-layer cascade described in Appendix A2 ; shared with the payments dataset.
3	Supplier matching and enrichment (<code>03-build-supplier-lookup.R</code> , <code>03a - 03c</code> alias scripts, <code>04-assemble-final-datasets.R</code>)	Normalises <code>awardedSupplier</code> , flags multi-supplier and unusable strings as <code>Excluded</code> , matches the remainder to the TCSS Organisation Register (exact then fuzzy), applies validated alias merges, assigns <code>supplier_id</code> , <code>supplier_type</code> and <code>uid</code> , and attaches the per-organisation payment totals from the payments dataset.

Note: The civil-society awards in this file (those with a `uid`) are the same awards that appear in the payments dataset with `data_source` = *Contracts Finder* or *contractsfinder*, where the `awardedValue` is carried as the payment `amount` . Analyses that combine the two datasets should use one or the other for Contracts Finder awards, not both.

A2. Buyer Classification

Buyers are classified into nine types (eight substantive types plus a Junk/Invalid catch-all) using the same multi-layer cascade as the payments dataset. Classification is applied independently to `funder_name` (producing `funder_type`) and `funder_name_alt` (producing `funder_type_alt`). The cascade proceeds in order; the first match wins.

Classification Cascade

Pre-filter: Junk/Invalid Detection

Entries with purely numeric names, hash-like strings, or very short names (fewer than 3 meaningful characters) are flagged as `Junk/Invalid` and excluded from subsequent cascade layers. In this dataset 0.5% of awards have a Junk/Invalid buyer.

Layer 1: Metadata Signals

Information already present in the source data is used as the first classification signal:

- Funders from the `NHSSpend` data source (payments dataset) are classified as **NHS**
- Funders from the `Central Government` data source (payments dataset) are classified as **UK Government**; Contracts Finder buyers manually refined to a parent department carry that department in `funder_name_alt`
- The `note` field in the funder masterlist may contain signals such as "ministerial" or "non-ministerial", indicating **UK Government**

Layer 2: External Lookups

Unclassified funders are matched against two external reference databases:

- findthatcharity** (<https://findthatcharity.uk/>) — a comprehensive lookup of UK organisations. The `organisationType` field is mapped to funder types (e.g., `nhs-trust` → NHS, `local-authority` → Local Government). Both exact and fuzzy matching (Jaro-Winkler, threshold ≥ 0.90) are used.
- TCSS Organisation Register** — funders that match the Organisation Register are classified as **CSO** (a civil society organisation acting as a funder).

Layer 3: Keyword Rules

Remaining unclassified funders are matched using pattern-matching rules applied to the funder name. Rules are applied in priority order; the first match wins:

- NHS** — names containing "NHS" combined with "TRUST", "CCG", "ICB", etc.
- UK Government** — known department abbreviations (FCDO, DFID, etc.) and patterns like "BRITISH EMBASSY", "SCOTTISH GOVERNMENT"
- Local Government** — names containing "CITY COUNCIL", "COUNTY COUNCIL", "BOROUGH COUNCIL", etc.
- Police, Fire and Rescue** — names containing "CONSTABULARY", "POLICE", "FIRE" + "RESCUE"
- Education Institution** — names containing "UNIVERSITY", "COLLEGE", "ACADEMY TRUST", school patterns
- CSO** — names ending in "CIC" or containing "TRUST" or "CHARITY" (after NHS and Academy Trusts have been captured)
- Private Sector** — commercial companies identified by the external lookups (registered companies that are not civil society organisations)
- Other** — all remaining unclassified funders

Funder Type Taxonomy

Funder Type	Description	Examples
UK Government	Central government departments, agencies, and arm's-length bodies	DEPARTMENT FOR EDUCATION, MINISTRY OF DEFENCE, DVLA

Funder Type	Description	Examples
NHS	NHS trusts, clinical commissioning groups, integrated care boards	NHS ENGLAND, BARTS HEALTH NHS TRUST, NHS HRAW CCG
Local Government	County, district, borough, and unitary councils; combined authorities	MANCHESTER CITY COUNCIL, KENT COUNTY COUNCIL
Education Institution	Universities, colleges, academy trusts, schools	UNIVERSITY OF OXFORD, HARRIS FEDERATION
CSO	Civil society organisations acting as funders (grant-makers, intermediaries)	THE NATIONAL LOTTERY COMMUNITY FUND
Police, Fire and Rescue	Police forces, fire and rescue services	METROPOLITAN POLICE, LONDON FIRE BRIGADE
Private Sector	Commercial companies appearing as the buyer — typically procurement intermediaries or managed-service providers acting on behalf of public bodies (identified mainly through the external lookups)	ATAMIS LTD, EU SUPPLY LIMITED
Other	Funders not classifiable into the above categories	Various unclassified public bodies
Junk/Invalid	Clearly erroneous entries (numeric strings, hash tokens)	3149053, 37300

A3. Supplier Matching & Alias Resolution

The `awardedSupplier` field contains supplier names as entered by the buyer on the award notice — often inconsistent in spelling, abbreviation, and formatting. The pipeline matches these names to organisations in the TCSS Organisation Register to assign a standardised `uid` to each supplier.

Matching Process

Matching proceeds in two stages:

- Exact matching** — supplier names are normalised (uppercased, punctuation removed, whitespace collapsed) and matched exactly against the Organisation Register.
- Fuzzy matching** — unmatched suppliers are compared to the Register using Jaro-Winkler string similarity. Matches with a similarity score ≥ 0.90 are accepted. This captures variations in spelling, abbreviation (e.g., “LTD” vs “LIMITED”), and minor data entry errors.

Alias Resolution

After initial matching, the pipeline identifies potential aliases — cases where the same organisation appears under different supplier names. This is common when departments record the same supplier differently (e.g., “ST LUKE’S HOSPICE” vs “SAINT LUKES HOSPICE”).

Alias resolution proceeds in three steps:

- Candidate generation** (Script 03a) — deterministic rules identify supplier name pairs that may refer to the same organisation, based on shared UIDs, similar names, or overlapping funder relationships.
- Batch review** — candidate pairs are grouped into batches and reviewed using a combination of LLM-assisted classification and manual checks. Each pair is labelled as a confirmed alias or a false positive.
- Merge** (Scripts 03b–03c) — confirmed aliases are assembled into a validated alias file, and the supplier lookup is updated to consolidate duplicate entries under a single canonical record.

Note: The LLM review step occurs externally between Scripts 03a and 03b. In a fully automated run where reviewed batches already exist, all scripts execute in sequence without manual intervention.

A4. Linking to Other TCSS Datasets

The dataset can be linked to other datasets in the UK Third and Civil Society Sector Database using the `uid` field (register-matched civil society suppliers) and to the payments dataset using `supplier_id` or `uid`.

Available Linkages

Dataset	Join Key	What It Adds
Procurement Payments (https://uk-third-sector-database.github.io/guidance/tcss-procurement-payments-guidance.html)	<code>uid</code> or <code>supplier_id</code>	Payments made to the same civil society organisations by central government and NHS bodies, for comparing contract awards with money actually paid
Organisation Register (https://uk-third-sector-database.github.io/guidance/tcss-organisation-register-guidance.html)	<code>uid</code>	Organisation name, postcode, registration and removal dates, source registers, SIC codes
Charity Financial Records (https://uk-third-sector-database.github.io/guidance/tcss-charity-financial-records-guidance.html)	<code>uid</code>	Annual income, expenditure, and detailed financial breakdowns for registered charities (GB-CHC, GB-SC, GB-NIC prefixed UIDs)
Nonprofit Financial Records (https://uk-third-sector-database.github.io/guidance/tcss-nonprofit-financial-records-guidance.html)	<code>uid</code>	Companies House accounts data — balance sheets, profit and loss, employee numbers — for nonprofit companies (GB-COH prefixed UIDs)
CIC Founding Purposes (CIC 36/37 forms) (https://uk-third-sector-database.github.io/guidance/tcss-cic-statements-guidance.html)	<code>uid</code>	Community interest statements, beneficiary descriptions, and activity summaries for Community Interest Companies

Example: Civil Society's Share of Awards by Buyer Type

```
import pandas as pd

cf = pd.read_csv("contractsfinder-enriched-refined.csv", low_memory=False)

# A register-confirmed civil society award is one with a uid
cf["is_cso"] = cf["uid"].notna()

# Use the award date, not payment_year, for time series
cf["award_year"] = pd.to_datetime(cf["awardedDate"], utc=True, errors="coerce").dt.year
core = cf[cf["award_year"].between(2015, 2025)]

# Share of awards (by number) won by civil society, per buyer type
share = core.groupby("funder_type")["is_cso"].mean().sort_values(ascending=False)
print((100 * share).round(1))
```

Example: Linking CSO Awards to the Organisation Register

```
import pandas as pd

cf = pd.read_csv("contractsfinder-enriched-refined.csv", low_memory=False)
spine = pd.read_csv("TSCS_spine.spine.csv")

cso_awards = cf[cf["uid"].notna()]
merged = cso_awards.merge(
    spine[["uid", "organisationname", "postcode", "dateregistered"]],
    on="uid",
    how="left"
)

# Example: which registered organisations win the most contracts?
print(merged.groupby(["uid", "organisationname"]).size().sort_values(ascending=False).head(10))
```

Tip: When linking datasets, use a left join from the Contracts Finder data to preserve all award records, even if some organisations are not found in the target dataset. Check for missing values after the join to assess linkage coverage.