Charity field classification

Main project report

September 2021

Acknowledgements

This project was funded in part by a grant from the Esmée Fairbairn Foundation. We thank them for their support.

We would like also to acknowledge the following contributions:

Oliver Chan for his work as a research assistant on the project and especially his tireless work helping to code the manually classified sample dataset.

Caitlin Curran for some additional help coding the manual dataset as part of her placement at CRESR.

Lisa Hornung and her initial involvement in setting up the project at NCVO.

Lorna Dowrick for detailed comments on the UK-CAT.

We would also like to acknowledge the work of Lester Salamon, who passed away in August 2021. Lester was a professor at Johns Hopkins University who first designed the ICNPO classification along with Helmut Anheier.

Introduction

Non-profit organisations are united not only by their non-profit status, but also, depending on the definition used, their voluntarism and social purpose. At the same time, they vary greatly in their activities, size, legal form, targeted beneficiaries, structure, level of voluntarism, governance, human and financial resources, ownership arrangements, and undoubtedly much else (Kendal and Knapp, 1995). From large international aid charities running programmes across multiple continents, to small village halls, which provide spaces to community groups, the sector is nothing if not diverse. As such, many researchers are not just interested in the entire non-profit sector, but in either exploring the characteristics of one sub-sector in more depth or comparing several at once.

One of the most popular forms of classification is to divide charities on the basis of which cause, purpose or mission they seek to benefit with their charitable activities. A focus on activities may partly be due to the interest from many stakeholders in identifying the combined level of *economic activity* within the sector (Salamon and Anheier, 1992; Kendal and Knapp, 1995). Although this emphasis is sometimes criticised (Barman, 2013), the idea is presented that the size, scope, and importance of the non-profit sector is often overlooked and undervalued. For example, some have suggested that this may have contributed to the non-profit sector losing out in relative terms when the Government began distributing COVID-19 relief (Cooney, 2020). Researchers or infrastructure bodies within the sector itself, often, therefore, conduct mapping exercises of particular activity areas, in order to draw more attention to their scale and scope (Clifford et al. 2013; NCVO, 2012; Newbigging et al. 2017). Doing so tells us not only about their economic contribution, but also at least something about the scale of their contribution to social welfare and the collective good.

There are also conceptual reasons for wanting to 'home in' on sub-sectors of activity, where the whole sector may be too heterogeneous to provide an effective case study. As Kendal (2003) has pointed out, the regulatory and policy environment that non-profit organisations experience is often defined more by their activity area than their non-profit status, shaping their opportunities and limitations. An organisation's activities may also determine which other organisations it interacts with most often, helping to create fields of activity in which norms and institutions are able to spread more easily (Leiter, 2013; DiMaggio, 1983). Researchers may, therefore, wish to sample case studies from within a particular activity area, include 'activity-area' as a covariate within their modelling, compare one sub-sector with another, or compare for- and non-profit providers within a single activity area. Looking within an activity area does not eliminate the issue of internal diversity, of course, so much as shift it down a level. The most appropriate approach will depend on the research aim at hand.

In several countries, however, it has been noted that the most comprehensive data available on non-profits do not provide sufficient information on activity areas. The problems fall into one of two camps: either issues with the classification scheme used, or with its application and availability within the relevant datasets. In the US, a small but growing literature addresses concerns with the use of the National Taxonomy of Exempt Entities (NTEE) and its use with Form 990 data (Ma, 2020; Fyall et a., 2020; Lampkin et al. 2001). In European countries such as Austria, attempts have focussed on applying the International

Classification of Non-profit and Third Sector Organisations based on organisational names (Litofcenko et al, 2020). We are also aware of attempts to apply bespoke classification systems on Australian non-profits (Our Community, 2020).

In the UK, the largest and most used sources of data on non-profits are the registers of charities in England and Wales, Scotland, and Northern Ireland, maintained by the respective national regulators. As with other countries, however, there are several issues with both the classification schemes and their availability. Neither the classifications included within the various registers, nor alternatives such as the ICNPTSO, map well onto some of the areas of most interest to researchers in a UK context. ICNPTSO categories are not collected directly from charities at all, while the self-selection of the regulatory categories by charities brings its own challenges.

This report outlines our efforts to solve these problems by assigning each charity in the UK a category from the ICNPTSO, as well as creating a new, bespoke classification system against which charities can be allocated multiple relevant categories. As well as creating a training dataset using human coding, we have developed automated approaches to apply both classification systems based on the text within their regulatory records. The results are available for free on a shared commons licence, and should allow researchers, funders, member networks and others to create their own groups of charities to fit their needs.

Literature review

Controversy over classification schemes

Classifications are the conceptual building blocks of the social world, used to divide phenomena into workable, comprehensible segments (Bowker and Star, 1999; Barman, 2013). When conducting classification work, we metaphorically place phenomena either ‘inside’ or ‘outside’ conceptual boxes, and in doing so, we suggest how we might view or act towards them. To classify is, therefore, an everyday activity undertaken to make sense of the world. Nevertheless, it can be a controversial process. Classifications are socially constructed and negotiated, and will, therefore, inevitably embody a particular set of ethical and political values (Bowker and Star, 1999). Individual categories may become the centre of contention, especially when linked to emotionally resonant issues such as sexual identity or race (Bowker and Star, 1999). The classificatory system also invariably helps to highlight some entities, or their characteristics, at the expense of leaving others less visible. Some examples may be forced into ill-suited descriptors or combined together into ‘other categories’.

Barman (2013) argues that classification systems are always the product of struggles between groups for various forms of capital and as such are bound to inequality and domination. By this view, a classification system is necessarily hierarchical and tied to an unequal distribution of resources. Barman cites Bourdieu, who suggests that classifications possess “the power to make people see and believe, to get them to know and recognize, to impose the legitimate definition of the divisions of the social world and, thereby, to make and unmake groups” (Bourdieu 1991: 221 in Barman, 2013). For example, Barman (2013) suggests that the National Taxonomy of Exempt Entities (NTEE), used extensively in the US, was formed as part of a larger symbolic struggle by powerful elites and scholars to protect the tax-exempt status of foundations.

This line of argument is similar to those who critique the ‘mapping’ of non-profits, especially as a prelude to state intervention. Nickel and Eikenberry (2015) claim that mapping primarily serves the interests of the powerful, at the expense of the democratic needs of VSOs' participants and stakeholders. They argue that increased visibility makes organisations vulnerable to greater external discipline, even if ostensibly conducted for critical or emancipatory purposes. Carmel and Harlock (2010) similarly suggest that mapping is one means of making previously private spaces 'knowable' and therefore 'governable'. It means they are visible to state actors and can be more easily regulated, managed, and coordinated.

Perhaps unsurprisingly, we would not necessarily agree with these arguments in full, nor their underlying epistemology, which risks framing almost all activity and knowledge claims as acts of domination and power (Hay, 2002). This narrative also arguably underplays the extent to which mapping and categorisation exercises are driven from within the non-profit sector itself, rather than the state. Representative bodies, funded by their members, often seek to draw attention to the impact and scale of their particular sub-sector. Though even this view risks painting a somewhat reductionist picture, in which researchers, infrastructure bodies, the state and non-profits are seen as acting only in their self-interest. Many in both the state and the non-profit sectors are, of course, motivated to maximise the welfare of vulnerable individuals or other socially beneficial causes.

On the other hand, we would be wary of relying too heavily on metaphors about mapping the ‘invisible subcontinent’ of the non-profit sector (Salamon, 2010), which may imply neatly delineated units and categories waiting to be discovered by researchers without any need for social construction. We would agree that categorisation will inevitably privilege one conceptual framework at the expense of alternatives and that this will involve political and ethical choices as a result (Bowker and Star, 1999). Furthermore, some will have more power than others over which framework is privileged (Alcock, 2010), including ourselves as researchers. Appe (2012) is correct, therefore, to point out that who measures, and why they measure, matters.

For our part, we hope that the results of this project will be of most use to those conducting research on the sector from within the sector itself or academia. We are more wary of any suggestion that the classifications discussed should be used as a basis for targeting resources, though acknowledge that they might be used to match donors with charities. As researchers who subscribe to at least a minimal sense of independent reality, however, we would also argue that despite their socially constructed nature, some classifications are ‘better’ than others, or at least more useful to a much wider range of stakeholders. At the very least, therefore, we hope that the schemes and data provided offer improvements and alternatives to the existing options.

Bowker and Star (1999) offer helpful advice for the would-be classifier. First, it is important to deploy a healthy degree of self-reflexivity throughout the process, constantly asking which groups are being made more visible, and which are being made less so. Second, as far as possible, it is better to record the methods and decision-making process transparently and openly, so that systems can be critiqued and modified in the future. “The only good classification is a living classification” (Bowker and Star, 1999, p.326). As such, we seek to present an extensive account of our methods here and elsewhere, as well as making all the code and data from this project open source[1]. It should be possible for any external observer to be able to identify exactly how a given non-profit has been allocated to a particular category and to make suggested changes for future editions.

Charity classification in the UK

Authors such as Kendal and Knapp (1995) identify an extremely diverse range of concepts and categorisation systems with which non-profits can be classified. This paper focuses specifically on the charitable causes, purposes, or missions of non-profits. These are usually best summarised as a particular activity, beneficiary group or targeted problem, which define why the non-profit operates for the public benefit. The same general cause can, therefore, usually be presented in several ways, such as ‘addiction and dependency’, ‘people with addictions’, ‘addiction rehabilitation’, or ‘addiction clinic’. Understanding which non-profits are conducting activities for different purposes does not just help to map the economic activity of the sector, or even its quantified social impact. It also helps us to understand just what the

[1] https://github.com/drkane/ukcat/

non-profit sector is, its reach, boundaries, purpose, and the extent to which it can meet the many demands placed upon it.

Several prominent classification systems already exist to classify either UK charities or non-profits internationally according to their social mission. As part of their registration process, charities in England and Wales are asked to select from several different drop-down lists. The first asks '*what does your charity do?*' and includes options such as 'animals', or 'amateur sport'. It can be roughly characterised as measuring the charity's purpose. The second, '*who does your charity help?*' identifies beneficiary groups such as 'Children / young people' and 'Elderly / old people'. The final question, '*how does your charity operate?*', identifies whether a charity 'makes grants to organisations or 'provides services', amongst other activities.

The three groups can become quite conflated in practice and there is some overlap between the questions. For example, the question on purposes contains an option for 'disability', while the question on beneficiary groups includes 'people with disabilities'. Some categories are also very broad, including a 'general charitable purposes' category. More specific categories, which could be considered important in the UK context are not included, such as food banks, homelessness services or medical research.[2]

The Scottish Register of Charities contains a very similar set of three category lists, with similar issues. The main difference is that the categories are applied post-hoc by officials at the Office of the Scottish Charity Regulator, based on several longer, textual questions charities fill out as part of their application process. These include: '*What are the activities or projects the organisation intends to run?'*, '*How will these activities help achieve the organisation's charitable purposes?', 'Who will benefit from these activities?',* as well as the charities' formal purposes in full.

The Northern Irish process is more similar to the version in England and Wales. Charities select pre-coded options to describe the 'charitable purpose', which are similar to those listed in the other national registers. They also select one or more 'main focus' for the organisation, which in practice is simply a longer list of purposes, and finally from a list of beneficiary groups. Again, there is considerable overlap between the three sections, with some such as 'animal welfare' appearing twice, nearly verbatim, and 'disability' appearing three times within different categories.

The National Taxonomy of Exempt Entities (NTEE) is also worth noting. As the most widely used classification system of tax-exempt non-profits in the US, it has a relatively high degree of prominence in the academic literature (Fyall, et al. 2018; Lampkin et al., 2001; Ma, 2020). It divides non-profits into categories based on their organisational purpose, using 10 top level categories, 26 second level categories and ultimately around 450 third level categories (Lampkin et al. 2001). The major groups include broad areas such as health, education, and arts and culture (Fyall et al. 2018). In 1995, the Internal Revenue Service, the US's tax collection agency, took responsibility for assigning the classifications to non-profits when they apply for tax-exempt status, similar to the approach employed in Scotland (Ma, 2020).

---

[2] During the course of this project, we have been involved in a consultation process run by the Charity Commission for England and Wales, to update and expand the list of categories that they use. In some cases, the new categories may help to address some of these gaps.

Finally, in contrast to the nationally tailored schemes found in the UK registers of charities and the NTEE, the International Classification of Non-profit and Third Sector Organisations (ICNPTSO) is designed to be applied internationally, to enable cross country comparisons and assist with the preparation of national accounts (Salamon and Anheier, 1992; Kendal and Knap, 1995). In the US or UK, and to our knowledge internationally, ICNPTSO categories are not recorded as part of the formal registration process for non-profits, which means they are usually allocated retrospectively. To work across many national contexts, the categories are quite broad, and the system is sometimes a poor fit for common areas of charitable activity in the UK. For example, there are no specific categories for food banks, drug addiction services, or domestic violence refuges.

One of the authors of this report, over a decade ago, previously attempted to apply the first version of these categories to the register of charities in England and Wales, using a mixture of manual classification, automated keyword searches, and the charity commission classifications.These classifications have been used in a number of academic research outputs, but their accuracy was limited by the methods available at the time and they are now substantially out of date.

All of these systems have their advantages and disadvantages and cover certain groups of purposes (or beneficiary groups) in more detail than others. When designing the ICNPTSO categories (then known as the International Classification of Non-profit Organisations, or ICNPO), Salamon and Anheier(1992, p1) drew on Deutsch (1963) to identify several evaluative criteria, including *economy* (the number of categories and parsimony achieved), *significance* (the importance of the distinctions drawn), *rigour* (the repeatability and reliability of the measurements), *combinational richness* (the number of analytically useful comparisons it allows), and organizing power (its generalisability to other contexts). Salamon and Anheier (1992) identify that no single classification system can successfully serve all purposes.

We argue that there is a clear gap within the classificatory infrastructure for a scheme that prioritises significance and combinational richness in the UK social policy context, even if this comes at a modest cost in terms of more categories, less generalisability and relying on post-hoc classifications. A balance will clearly still need to be struck between enough categories to capture the most important nuances, while not making a system that is unworkable in practice, though this scheme would likely be more extensive than the current options available.

A second important distinction between the different classification systems is whether categories are mutually exclusive. Both the NTEE and the ICNPTSO systems have traditionally applied a single classification per organisation, which avoids double counting any economic activity. Ensuring mutually exclusive classifications also makes it more straightforward to incorporate the classifications into modelling work. On the other hand, a single category can fail to capture the multiple or combined purposes of many charities and risks excluding relevant organisations from sampling efforts (Fyall et al. 2018; Lampkin et al. 2001; Ma, 2020). Fyall, et al. (2018) discovered that many homeless housing providers in Washington State were not picked up by the relevant NTEE category due to their multipurpose nature.

Allowing multiple selections, however, can also bring challenges. When UK charities self-select their purpose, activities, and beneficiaries, they sometimes select a high number, despite marginal relevance, perhaps to show that they are being particularly prolific or inclusive (DEI data group). When organisations have many classifications and no ranking system is incorporated, it can be difficult to interpret which are most meaningful.

Whether multiple or single classifications are best, therefore, depends on the research task at hand. A key point to note, however, is that whilst the NTEE and INCPTSO have traditionally applied a single category, and the register of charity classifications have allowed multiple selections, this is a practical rather than due to a conceptual barrier. A flipped coin is either heads or tails, never both. But an non-profit organisation can easily undertake social advocacy (ICNPTSO category G11) as well as services to the elderly (category D14).

Data sources

The second major limiting factor when applying a classification scheme is the data available. At one extreme, we might only have the organisations' name from which to make a classification decision. Litofcenko et al., (2020), in their attempts to apply ICNPO classifications to non-profits in Germany and Austria, were faced with this limitation. The situation is improved somewhat for the Austrian non-profits, as the law stipulates that their names must be related to their organisational purpose, though this is not the case in Germany. In general, this data brings the risk of misleading or uninformative names, acronyms, or unconventional language such as wordplay, neologisms, regional dialects, or foreign languages (Litofcenko et al., 2020).

At the other end of the spectrum, when an employee at the IRS applies an NTEE category to an applicant for non-profit status, with the help of an algorithm since 2007 (The Non-Profit centre, 2008), they have access to all the application information provided. Similarly, when charities apply classifications themselves, such as those included in the various UK charity registers, they have access to all the information they could possibly need.

For those wishing to apply classifications to UK charities post-hoc, the situation lies somewhere in between. We generally have access to organisations' charitable objects, a legally required paragraph within the charity's governing documentation, which sets out their purpose and objectives. In the case of England and Wales, and Northern Ireland, for currently registered charities, we also have access to a written description of their charities' activities. Compared to the objects, these may be written in more modern language, be more up to date and contain less legalese (Leung, 2020). Both these text fields are somewhat similar to the 'mission statement' that non-profits in the US include as part of their Form 990 when applying for tax-exempt status (Fyall et al. 2018).

Whilst very useful for the purposes of classification, these data can introduce several challenges (Fyall, et al. 2018; Lampkin et al. 2001; Leung, 2020; Ma, 2020). First, charities may omit certain important aspects of their work in their descriptions, or even fail to complete the activities section at all. Second, the quality is uneven. Charities will sometimes write extremely general, uninformative clauses, such as 'general charitable purposes', 'at the complete discretion of the trustees', or even just take the opportunity to try and dissuade any speculative grant applications. Third, these records are rarely updated as charities change their activities, and in the case of UK charities, explicit permission is required from the national regulator to change their formal objects. Fourth, in both the UK and US it should be noted that not all non-profits fall under the scope of the same regulator, and several groups are exempt from having to submit full documentation, including smaller organisations.

Classification methods

Regardless of the data available, there are several options for how to apply a set of activity classifications (Ma, 2020). Non-profits, or external individuals, can apply classifications manually. Alternatively, either human coders or a supervised machine learning process can create an algorithm, to automatically apply a predefined set of categories. Or finally, an unsupervised machine learning model can derive its own categories, based on recurring patterns in the data.

The first, and most self-explanatory method, is self-section by the non-profits themselves. As described above, charities in the UK select classifications from several lists when applying to their relevant regulator. The person completing the relevant form is likely to know the charity well, the amount of work per charity is relatively small (though not negligible, especially if the number of potential categories is large), and no external entity is imposing a potentially unwelcome choice on the charity. On the other hand, charities may select an unnecessarily high number of categories when permitted. They may also lack expertise on the classification system itself, and their reasoning for choosing a particular category may not, therefore, always be immediately clear.

A similar, but distinct option, is manual classification by someone external to the non-profit itself. Generally, this is likely to be either a researcher or an employee of a regulator such as the IRS or OSCR (Ma, 2020). As described above, this individual may still have access to a reasonable amount of data on the charity, including application data, annual reports, websites, and other online information. They will also presumably be very familiar with the classification system. The main limitation is likely to be the time they have available to engage with all the data available for each charity. As the number of classifications needed increases in size, the less feasible human classification becomes (Fyall et al. 2018; Leung, 2020; Litofcenko et al., 2020). Reclassification over time is also less likely as a result (Ma, 2020; Fyall et al. 2018).

Litofcenko, Karner and Maier (2020) created a manually classified dataset of 5,000 Austrian non-profits (as well as a separate sample of 1,000 German non-profits), to help assess the results of their automated methods. The ‘correct’ classification was allocated by consensus, with three different coders. Individual coders achieved between 79 and 87 per cent agreement with this final allocation. The authors stress the relatively high level of expertise needed for the task, the extensive amount of time it took to complete the coding, and the relatively low transparency of the process due to the many subjective judgements needed.

In contrast, there are a range of automated classification options available. Ma (2020) distinguishes between ‘dictionary methods’, and both supervised and unsupervised machine learning. All automated methods can make use of however much textual data is available for each charity (though including absolutely everything may not always be beneficial). The methods can also be applied indefinitely and remain consistent, whereas human coders might tire.

Dictionary methods search for keywords throughout the text that either increase or decrease the probability of an entity belonging to a particular category. For example, Fyall et al. (2018) searched for a list of keywords to determine the probability of a non-profit providing homelessness accommodation. Litofcenko, Karner and Maier (2020) also used keywords as part of multiple if-then statements, to allocate each non-profit in their sample to an ICNPO category. To achieve a single result per non-profit, these matching rules were ordered into different hierarchical tiers. In both cases, the keyword search rules were created by the researchers, who possessed a high degree of background knowledge.

The degree to which dictionary methods are successful depends partly on the quality of the underlying data, partly on how distinct and clearly defined the categories are, and partly on the effectiveness of the keywords chosen. Inevitably, there will be errors of both over-identification and under-identification (false positives and false negatives) (Fyall et al. 2018). Litofcenko, Karner and Maier (2020), using just information from non-profit names, suggest that their algorithm was 'correct' 85 per cent of the time, similar to the results achieved for an individual human coder. Fyall et al. (2018) did not have a manually classified dataset to compare against, but did achieve significantly more matches amongst their sample of Washington State non-profits than relying on NTEE categories or regulatory lists alone.

In an unsupervised machine learning model, the algorithm uncovers linguistic patterns in the texts without using a prior set of classifications (Ma, 2020). Leung (2020) used an unsupervised method of natural language processing and clustering to identify recurring terms within the activities and objects of organisations who had self-selected the 'arts, culture, heritage or science' category on the register of charities in England and Wales. The result was an automatically generated, hierarchical taxonomy of keywords on activities and beneficiaries for these charities. Although Leung's categories are generally meaningful and detailed, the 2,747 'end clusters' and 92 second tier categories may be difficult to work with for some purposes. There is also a risk with unsupervised models generally that not the categories may not be as theoretically meaningful as those developed a priori (Ma, 2020).

Finally, a supervised machine learning method is similar to the dictionary method, in that the list of categories is derived in advance. The rules and keywords used to apply NPOs into different categories, however, are derived inductively by a machine learning algorithm, attempting to apply the best possible predictive model to a learning dataset containing 'correct' classifications. This is then tested before being used on cases where the classification is unknown beforehand. Classification based on text is a relatively common machine learning task (Ma, 2020; Lantz, 2015).

In addition to their keyword-based rules, Litofcenko et al., (2020) experimented with a decision tree algorithm, generated based on statistical properties. They found the results unsatisfactory and suggested this may be because they lacked any long, high-quality texts such as mission statements, and were forced to rely on names and web scraped data from websites. Using a training sample of 1,068 cases to develop the algorithm, and a test sample of 750 to assess it, none of their models classified more than 50% of the test sample correctly. This increased somewhat by preselecting the most relevant words, based on their dictionary method, to 77 per cent, though this was still worse than the rules-based approach on its own.

Ma (2020), using relatively advanced machine learning methods, experimented with several different models and parameters to apply NTEE categories using Form 990 textual data. The most successful is called a BERT classifier (Bidirectional Encoder Representations from Transformers) and achieved 90 per cent overall accuracy using the nine broad NTEE categories and 88% for the 25 major groups. This research suggests that a very high level of success may be possible, given enough cases, high-quality textual data, and machine learning expertise.

Methodology

In response to the issues regarding both the existing classification schemes and their availability within key UK charity datasets, this project's aims were:

- To develop a new classification system for UK charities - the **UK Charity Activity Tags (UK-CAT)** - able to better capture the full range of charitable purposes in the UK context.
- To automatically apply the UK-CAT to all the charities registered in the UK.

The first step was to create a population level dataset of all UK charities with suitable data, using the various national registers of charities, and to then **create a sample of charities** from within that population. We then **designed the UK-CAT** classification scheme, in an iterative way, using the sampled charities. The researchers created new tags as they went through each charity in turn, with frequent iterations to combine and refine the classifications as a team. The sampled organisations could then be **manually classified** using the UK-CAT, as well as the ICNPTSO, to create a baseline dataset of organisations to train and test different methods for automatic classification. A **rules-based classification** was developed using regular expressions that could apply the UK-CAT tags to every charity in the UK based on their names and activities. The manually classified set of charities could then be used to test the accuracy of these rules.

1. Sampling strategy

The population used for sampling included all charities on the main registers for the three UK charity regulators (Charity Commission for England and Wales, OSCR - Office for the Scottish Charity Regulator, and the Charity Commission for Northern Ireland). After removing any inactive charities from the files, this left a population of around 200,000.

Keyword searches were performed for a small number of groups of very easily identified, homogenous charities, based on their name only. These groups could then be excluded in bulk straight away, meaning that the remaining charities would have more variety. The groups that were excluded were:

  - Parent Teacher Association (7,430 charities)
  - Village Hall (6,865)
  - Scout groups (4,266)
  - Girlguiding groups (2,805)
  - Parochial Church Council (2,746)
  - Playgroups (1,740)

The remaining 174,692 charities were then split into two income bands, with a threshold of £100,000, based on the latest available income of each charity. 39,826 (23%) of the remaining charities were above this threshold and 134,866 (77%) below it.

Between the different members of the research team 4,203 charities were ultimately classified using the UK-CAT. This included 1,328 charities with an income of £100,000 or over and 2,875 charities with a lower income.

Larger charities were, therefore, proportionately over-represented in the manual classification. This is partly due to their financial weight, to try and improve eventual estimations of financial totals for each service area. Secondly, some types of charities are disproportionately found amongst larger charities (for example medical research charities). In contrast, many smaller charities fall into a relatively small number of categories, such as places of worship, small grant makers and community associations. Without oversampling larger charities, therefore, some classification types would be likely to be underrepresented or missed out entirely from the manually classified sample.

2a. Developing the UK Charity Activity Tags

The UK Charity Activity Tags (UK-CAT) taxonomy system was developed iteratively whilst manually classifying the sample of charities. The list of categories is considerably longer than the ICNPTSO or the existing charity register classifications, at over 250 tags, in order to help capture more of the variation seen within UK charities. The tags sit within a hierarchy of 24 top-level categories, such as health, education or social welfare. In some cases, there are also mid-level subcategories to help structure the system further. A full list of the UK-CAT can be found in Appendix A.

To begin with, we created these tags from scratch each time they found a common charitable purpose that was not already covered. At regular intervals, the researchers met to consolidate, clarify and coordinate the full list. The aim was to help remove duplicates, fill in gaps, remove tags that were judged too ‘niche’, and make sure that the different coders were interpreting the tags in a reasonably coherent fashion. This iterative process of refinement continued throughout the project, until Version 1.0 was finally declared in order to write up the results.

Sometimes informal keyword searches were conducted on the population as a whole to get a sense of whether a tag was likely to be sufficiently common to include. Generally speaking, it was expected each tag would be applied to well over a hundred charities in the population. This was not an absolute rule, however, as a few categories with fewer matches were included due to their conceptual importance or for the sake of consistency with other tags.

In addition to developing categories inductively, the research team also drew on the existing classification frameworks in the various registers of charities, to make sure that there were no obvious gaps. The aim was generally to be more comprehensive than existing schemes. A lookup table of the different classification schemes against the UK-CAT is available in Appendix A.

Identifying charitable purposes

As discussed in the literature review, there is often some conceptual blurring which occurs around the idea of charitable purposes. Most commonly, charitable activities are geared towards either solving a particular problem, such as ‘poverty’, or to promote something seen as positive or valuable, such as a sport, a cultural activity or the environment. In many cases these are two sides of the same coin, such as ‘loneliness’ or ‘social activities.

Some groups of tags, however, do not fall under the overall heading of ‘charitable purpose’ as neatly. Beneficiary groups are sometimes identified separately within classification systems, but in reality many charities' purposes are to help a particular group, such as people with disabilities, in a whole host of ways. A number of these categories have been included in the UK-CAT under the general heading ‘beneficiary group’, though some others such as ‘veterans’ are found under different top-level categories.

Again reflecting some of the distinctions found in the national registers of charities, some categories account for a particular type of activity, such as giving grants, running charity shops or campaigning. Arguably, these stretch the boundary of the classification scheme, given that these activities are not charitable purposes in and of themselves. But when combined with other tags, they can help to significantly clarify what a charity does and why. In some cases, such as small grant makers, this type of classification may be the only one possible, given the organisation’s other purposes are fairly indeterminate or generally phrased.

Similarly, some charities frame their purpose almost entirely around providing a particular facility, such as a village hall, playing fields or a community centre. Again, providing such a venue is not inherently charitable (professional sports pitches hardly qualify, for example), but it can usually be inferred that these facilities are being offered to groups, or for activities, which justify a charitable status. A certain degree of inference is also needed for most types of ‘associations’, including scouting groups, women’s institutes and service clubs. These organisations generally provide a wide range of different charitable activities for their members and users.

We have not been too strict, therefore, with what constitutes a charitable purpose when forming the UK-CAT, to allow for the fact that different charities describe their objectives in different ways. By including all of these groups into a single scheme, however, we have tried to eliminate some of the duplication and ambiguity found in other classification systems. In all cases, we have attempted to focus on what it is about an organisation that actually makes it charitable, even if this requires a combination of tags to fully represent.

2b. Manual classification of a sample of charities

One of the main tasks the research team conducted was to manually classify the sample of charities using the UK-CAT, assigning the organisations as many ‘tags’ as were applicable. The classification exercise took place over several weeks, with each researcher allocated a random batch of organisations to work through. Frequent meetings between researchers took place in order to compare notes, discuss any difficult or ambiguous cases, and agree on shared rules for dealing with them.

As described above, this process was somewhat complicated by the fact that the UK-CAT was developed in parallel to the manual coding taking place, meaning that backwards revision was sometimes needed when making changes. Fortunately, in practice, this did not prove too difficult. First, most of the tags and the broad framework of the scheme was developed early on in the process. Second, changes to tag names or the hierarchy, and merging or removing tags, could be achieved relatively straightforwardly. Fourth, adding new tags generally only occurred the first time a new type of charity emerged, meaning that previous examples were not likely to have been missed.

Nevertheless, some retrospective checking of the finished dataset was required to try and refine the results as far as possible. This process, as well as the fact that multiple tags could be applied, make it difficult to apply a stringent test of inter-coder reliability using the UK-CAT coding. At the same time as applying the tags, however, we also applied a single category for each charity from the ICNPTSO, to ensure that these categories could also be automated and to allow further experimentation with machine learning techniques. Because we began with the ICNPTSO coding scheme ready formed, and applied only one category per charity, it was easier to test the inter-coder reliability using this data.

For this purpose, half way through the creation of the manually classified dataset, the three coders each coded the same subsample of 50 charities using the ICNPTSO categories, based on their name, activities and objects. All three coders achieved the same result for 60 per cent (30) of the cases. Two out of three matched for 28 per cent (14) of cases and all three used different categories for 12 per cent (six) of the charities. In a few cases, the disagreements were simple mistakes on behalf of one or more of the coders. More commonly though, a charity could be reasonably applied to more than one ICNPTSO category, especially where no obviously applicable category exists in the scheme. Partly, this reflects some of the concerns with the INCPTSO raised in the literature review and which motivated the development of the UK-CAT. On the other hand, the process is also inherently somewhat subjective. Nevertheless, having retrospectively agreed one category to act as the ‘correct’ classification for each of the 50 charities, the individual coder agreement rates were 86, 82, and 78 per cent, similar to those achieved by Litofcenko et al., 2020 (though using more textual data than they had available).

The exercise does introduce a reasonable level of caution concerning the manually coded dataset, and whether it can be said to constitute the ‘correct’ set of codings. Furthermore, unlike Litofcenko et al., 2020, this project did not have sufficient resources for all of the coders to code all of the charities, or to reconcile all of the differences which occurred as a result. The potential issues are likely to be even more true for the UK-CAT classifications, given the added parameter of how many tags to use, the higher number of opportunities (classifications) for human coders to potentially disagree on, and the higher number of categories to choose from.

Arguably, the process highlights some of the disadvantages of manual coding; namely human error, subjectivity and a lack of transparency. As a research team we were drawn, therefore, to the advantages of a dictionary method of coding. The manually coded dataset had not only allowed us to develop and refine the UK-CAT, but also provided a foundation from which to develop automated key-word search rules, and test their effectiveness. One of the main advantages to using the keyword searches was transparency, given some of the ethical risks involved in setting up classification schemes in the literature review. The keyword search process is relatively simple and easy to understand. Anyone is able to see which terms we used, and therefore exactly why a particular charity was matched. These rules can then be subject to feedback and improvement over time.

We are also interested in whether machine learning is able to outperform our own efforts at creating these automated classification rules. Given the availability of textual data from the activities and objects fields in the registers of charities, there is the potential to achieve a higher success rate than Litofcenko, Karner and Maier (2020) were able to achieve with just name data and web scraping websites. Indeed, Ma (2020) appears to have achieved a higher success rate in the US context using more data.

Not all machine learning algorithms are particularly transparent, however, or even comprehensible to humans. Nor can new categories easily be added without creating new training data. For these reasons, and due to time constraints, we decided to focus resources on the human derived keyword searches as part of a 'dictionary' method, at least in the first instance. We hope that providing a robust and simple to understand implementation, based on keyword searching, will help to establish the UK-CAT system and lay a strong foundation for further experimentation with more advanced machine learning techniques.

4. Rules based classification

The manually classified entries provided a pool of baseline data from which to start developing the automated keyword searching, using regular expressions[3]. As a first step, we took all of the charities from the manual sample linked to each tag and ran frequencies on the most common words and pairs of words (bigrams). This provided an initial indication of the most commonly associated keywords, though in practice the research team used their own subjective judgement and knowledge to come up with many of the terms used.

Secondly, using an online tool created specifically for this purpose, we worked through each tag examining the 'false negatives'. These were charities which we had linked manually to a tag, but which were not yet being caught by our search terms. Examining the activities or objects of these charities often revealed necessary modifications or additions to the search terms. In this way, we were able to iteratively improve the keyword terms.

At the same time, we kept a close eye on those charities that were being included by the search terms, particularly those charities which the human coders had not matched to the tag. In many cases, these were entirely reasonable and had either been missed during the manual classification, left out due to other tags being prioritised, or were just slightly less central to the charity's purposes than the standard the manual coders would have used.

Unsurprisingly, the eventual search terms included for each tag was a balance between whether to prioritise avoiding false positives or false negatives. This had to be struck fairly intuitively by the research team, though in general there was a somewhat stronger emphasis on avoiding false positives and an acceptance that keywords would not be able to identify every single relevant charity.

---

[3] A regular expression, sometimes known as a regex, is a set of characters describing a search pattern that can be matched against text. Regular expressions use a syntax that offers several advantages over keywords by allowing complex expressions, such as matching any one of a list of keywords. https://en.wikipedia.org/wiki/Regular_expression

Table 1 shows some examples of the regular expressions used alongside their tags. Note that the ‘|’ symbol represents ‘or’, ‘\b’ represents the start or end of a word, and the ‘?’ means that the preceding element is optional.

Table 1: Example regex expressions / keyword search rules

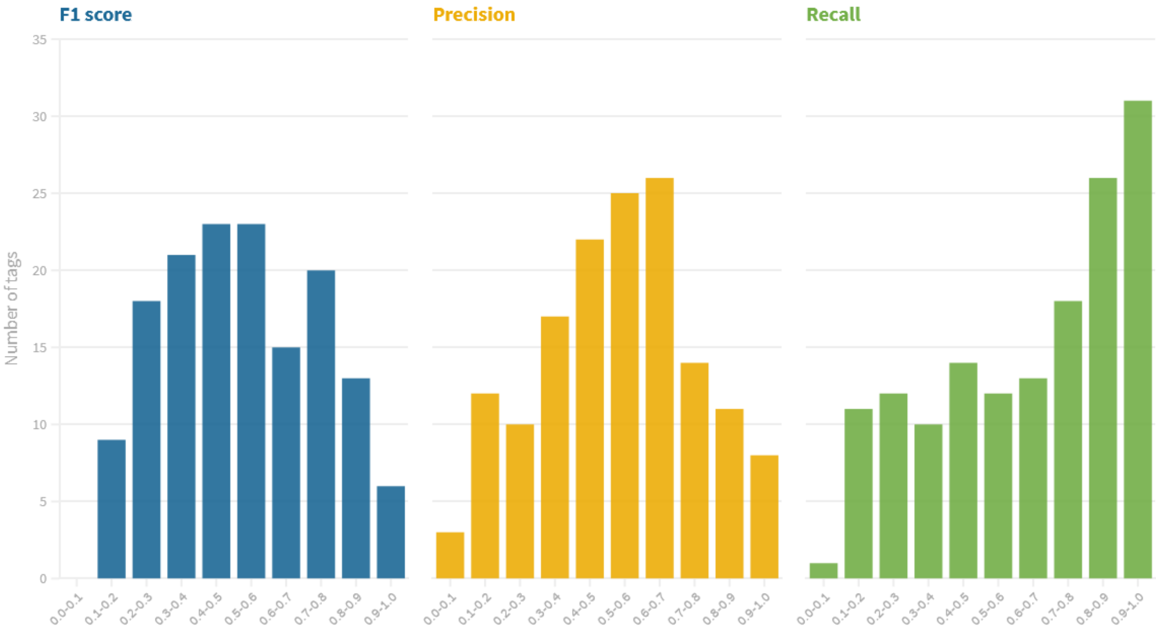
Arts - Performing art - Choirs	\b(choirs?|choral|chorus|choristers|singing|singers?)\b
Crime and Justice - Prevention and safety	\b(crime (prevention|reduction)|public safety|(prevention|reduction) of crime)\b
Education - scholarships	\b(scholarships?|bursar(y|ies)|grants for (student|pupil)s?)\b

The method for refining the tags allowed us to produce two measures of success: precision and recall, as well as the f1 score, which combines the two. Each measure was scored between 0 and 1. In our case, these measures are defined as follows:

- **Precision** shows what *proportion of the charities selected by the tag keyword* were also selected by the human coders. A low precision score would suggest that the keywords had selected lots of charities not matched by the manual ociding (and so potentially false positives).
- **Recall** shows the *proportion of the charities matched by the human coders* were also matched by the keyword search. A high recall score means that the keywords did well at finding a large proportion of the relevant charities. A low recall score means that the keyword found a small proportion of the relevant charities - lots of false negatives.
- The **F1 Score** combines these, using the harmonic mean. It generally reflects the lower of the two figures.

Figure 1 shows the distribution of tag results in bands using these three measures. It demonstrates that the current tags are better at maximising recall, and less so with precision. This means we would expect to see more false positives in the result, but fewer false negatives, which matches the general strategy adopted by the research team when deriving the keyword search terms.

Figure 1: Distribution of tags across different measures of quality of the results



https://public.flourish.studio/visualisation/6706687/

Findings

Overall results

Applying the key-word search rules to the dataset of 200,000 active charities[4] resulted in 800,000 matches across all 254 UK-CAT tags.

An initial ‘eyeball’ examination suggests that the results provide a reasonable summary of each charity’s activities. Table 2 shows a randomly selected group of five charities, along with their activities and their matched tags.

Table 2: 5 example charities and associated UK-CAT classifications

Charity name	**Activities**	**UK-CAT tags**
Corporation of The High School of Dundee	“The advancement of education.”	Education, Schools, Secondary education
Northampton Scottish Association Fund	“Providing charitable donation to local charities on an annual basis”	Associations, Charity and VCS support
Craven police charity fund	“Supporting local causes in raising money”	Fundraising, Emergency services
The windfall centre limited	“The Windfall Centre is a not for profit organisation of professsionals [sic] with expertise in the field of children's and young people's health, welfare and development. We provide therapeutic support to children and youung [sic] people through the medium of play and creative activities.”	Children, Young people, Health
1st Culter Rainbow Unit	Promoting the instruction of girls of all classes in the principles of discipline, loyalty and good citizenship.	Girls, Women, Citizenship

[4] May 2021 download

To conduct a more quantifiable test, we created an additional sample of 50 randomly selected charities[5], coded manually by the research team using the finalised UK-CATs. The research team allocated 63 tags in total across all 50 of the charities, applying no more than two tags per charity (though no upper limit was imposed). The keyword searching matched 162 tags directly when run against the same 50 charities, about 2.5 times more than the research team. Perhaps unsurprisingly given its automated nature, therefore, the keyword searching process is less parsimonious than the use of human coders.

As described further below, additional group and sub-group tags are sometimes applied automatically at the end of the automation process, based on *indirect* matches against lower level tags. So if 'Museum' is matched, its group, the group tag 'Heritage', is matched automatically. The following figures, however, only consider the tags *directly* applied by the research team or the keyword searching to avoid inflating agreement levels. Note, however, that this makes the comparison quite strict. For example, no credit will be given if the keyword search matches 'grant making' and the research team matches 'grants to individuals', for example.

Of the 63 tags applied directly or indirectly by the research team, 48 were also matched using the keyword searching (76 per cent). Of the 15 which were not matched by the keyword searching, reasons included:

- The activities field was very vague, in which case the human coders normally infer that it is a small grant making trust (1 case)
- The keyword search did not pick up on a key phrase, but could potentially do so with further modification and refinement to the regular expression (6 cases)
- The keyword search did not pick up a key phrase and is unlikely to do so, even with modification (sometimes due to key words having multiple meanings). Or, the human coders are able to infer a meaning not directly stated in the text (4 cases).
- The activities text lists a high number of activities, but the human coders summarised the charity as a 'community association' (2 cases).
- A spelling or formatting error in the activities field prevented the keyword match (2 cases)

These results are quite positive, even if they suggest that the keyword search terms have room for further improvement and refinement over time. Typing mistakes in the underlying data are harder to fix, though the idea of using automatic spelling correction may be possible for future iterations. It must be acknowledged, however, that there is always likely to be an upper limit to performance, either because the English language is too ambiguous or because the underlying activities description is too vague.

[5] Note that these are not the same 50 charities used for an inter-coder reliability test half way through coding, as described in the methodology.

Of the 113 tags matched by the keyword searching, but not by the research team, it would be misleading to assume that these were necessarily ‘incorrect’ (though they would reduce the ‘precision’ rate described above). Some of these are likely to be near misses, hitting the group or subgroup, but not the exact tag. Some lower order tags are also relatively closely related, such as community associations and community development, or community centres and village halls. Whilst we have tried to limit such ambiguities, some overlap is inevitable. Finally, some tags may be superfluous, but not necessarily erroneous. For example we tagged one charity using ‘playground’, but in contrast to the keyword search, did not also feel the need to include the additional tag for ‘young children’.

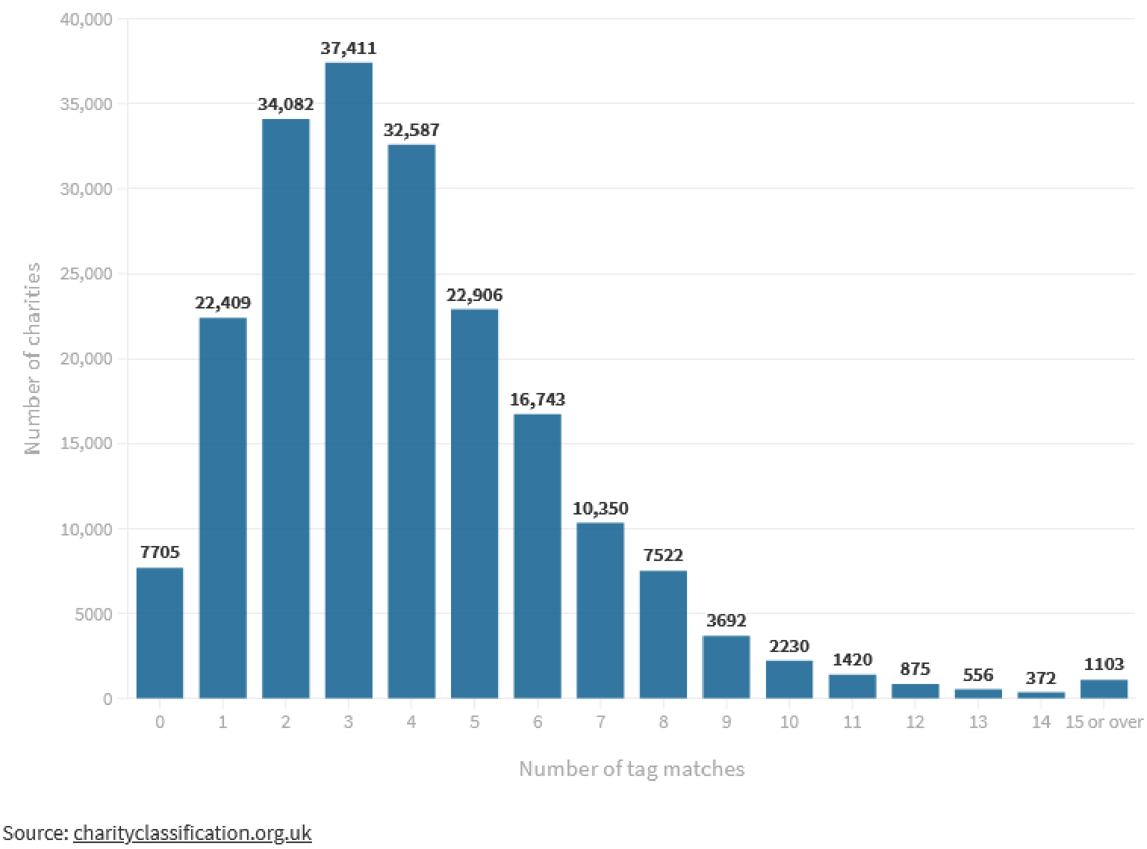
We manually reviewed all the 162 tags applied by the keyword searches to the 50 sampled charities. Although some tags were more important to describing the charities’ activities than others, we found only seven (4.3 per cent) which we felt could be classified with confidence as ‘false positives’ (all of which can be fixed in the first official update of the UK-CAT keyword terms).

Overall, we consider the estimated false negative rate (24 per cent) to be manageable and the false positive rate (4.3 per cent) to be very small. The key difference with coding by keyword matching is that it is about 2.5 times less parsimonious than the coding conducted ‘manually’ by the research team, which should be borne in mind when using the results.

Number of tags per charity

Figure 2 shows the number of direct UK-CAT matches charities have in the final keyword matched dataset. The mean and median number of tag matches per charity is four. 73.5 per cent of charities have between one and five tag matches, a further 20.5 per cent have between 6 and 10, and just 2.3 per cent have more than 10 matches. Overall, therefore, there is a modest amount of skew caused by charities with many matches, but most charities have a relatively small number.

Figure 2: Number of charities with different numbers of tags



https://public.flourish.studio/visualisation/7288933/

Charities with no matches

Four per cent of charities (7,705) have no matches at all. Table 3 shows a randomly selected ten charities from this group along with their activities field, which reveals a number of recurring themes.

Table 3: A selection of charities with no tag matches

Charity name	Activities
COMMUNITY CONCERN ARK	Providing care within the community
Netherdale Trust	for such charitable or philanthropic purposes as the trustees shall decide.
THE RAINE FAMILY CHARITABLE TRUST	GENERAL CHARITABLE PURPOSES
THE LORD CARADON LECTURES TRUST	To extend understanding of international affairs and tolerance of ideologies, races and religions, primarily through the series of annual lectures

CYLCH MEITHRIN LLANBERIS A NANT PERIS	Addysg i blant dwy a hanner i bedair mlwydd oedd
SREE AYYAPPA SEVA SANGAM (LONDON)	ADVANCE AYYAPPAN FAITH TO UK COMMUNITY.HELP THE NEEDEDPROVIDE SERVICE TO THE LOCAL COMMUNITYYOUTH WORK
LUKE SENIOR HALL	Provides a range of activities for the benefit of the whole of the community. Provides a venue for other organisations and individuals to hire for social and educational purposes.
THE NDL FOUNDATION	TO FURTHER SUCH OBJECTS OR PURPOSES WHICH ARE EXCLUSIVELY CHARITABLE ACCORDING TO THE LAW OF ENGLAND AND WALES IN ANY PART OF THE WORLD AND IN SUCH MANNER AS THE TRUSTEES MAY IN THEIR ABSOLUTE DISCRETION THINK FIT.
LEONARD JEROME CHARITABLE TRUST	At the Trustees' discretion
THE TOWN LANDS	Townlands Trust is a charity with general charitable purposes for residents/organisations within the villages of Halvergate and Tunstall.

Cylch Meithrin Llanberis A Nant Peris (a children’s nursery) has clearly written it’s activities in Welsh. In future, we hope to explore the option of using an automatic translation service (such as google translate) to allow us to match against these welsh language fields, or to use welsh language expertise to translate particular keywords for searching, but for the time being they are likely to have zero matches.

Sree Ayyappa Seva Sangam is a temple which has not been picked up by the ‘place of worship’ tag. It also mentions ‘youth work’, but a space is missing, which means this has not been picked up either. Again, it is possible that more advanced text and natural language processing techniques may be able to solve this difficulty in future iterations.

A number of foundations and trusts also have no matches. Their activities are often written in very vague or general terms, whilst using the key words ‘trust’ and ‘foundation’ in the search would include too many charities incorrectly. An alternative option, that can be applied post-hoc to the data download we provide, is to use a supplementary variable from the charity register. In their annual return, charities with an income over £10,000 are asked:

- *Was grant making the main way your charity carried out its purposes?*

We have included the data from charities’ answers to this question in the download, so that users can apply the grant making tag retrospectively if they wish to do so. Around 1,200 of the 7,700 charities with no tags have answered yes to this question.

Charities with many tags

A minority of charities matched a large number of tags. As an example, the charity with the most tags was a Northern Irish charity that matched 30 different tags. The following is an extract from their activities description:

“[The charity] raised funds for The Haiti Earthquake Appeal, for Breast Cancer groups and Ethnic minority families effected through bereavement . [The charity] has been delivering 10 weeks English Language classes to migrant worker, IT courses, Anti- Racism and Diversity training and Polish classes for children U12 from the Polish background and also children from the local community. [The charity’s] cultural events have become very popular in the area Multi -Cultural Food Art & Music Night, Polish Night, 3 days Diwali Festival Bollywood dancing, Music, Fashion show, Indian Food cooking Demo, Fire works Primary Schools Diwali story presentation and St Patrick’s Day parade...”

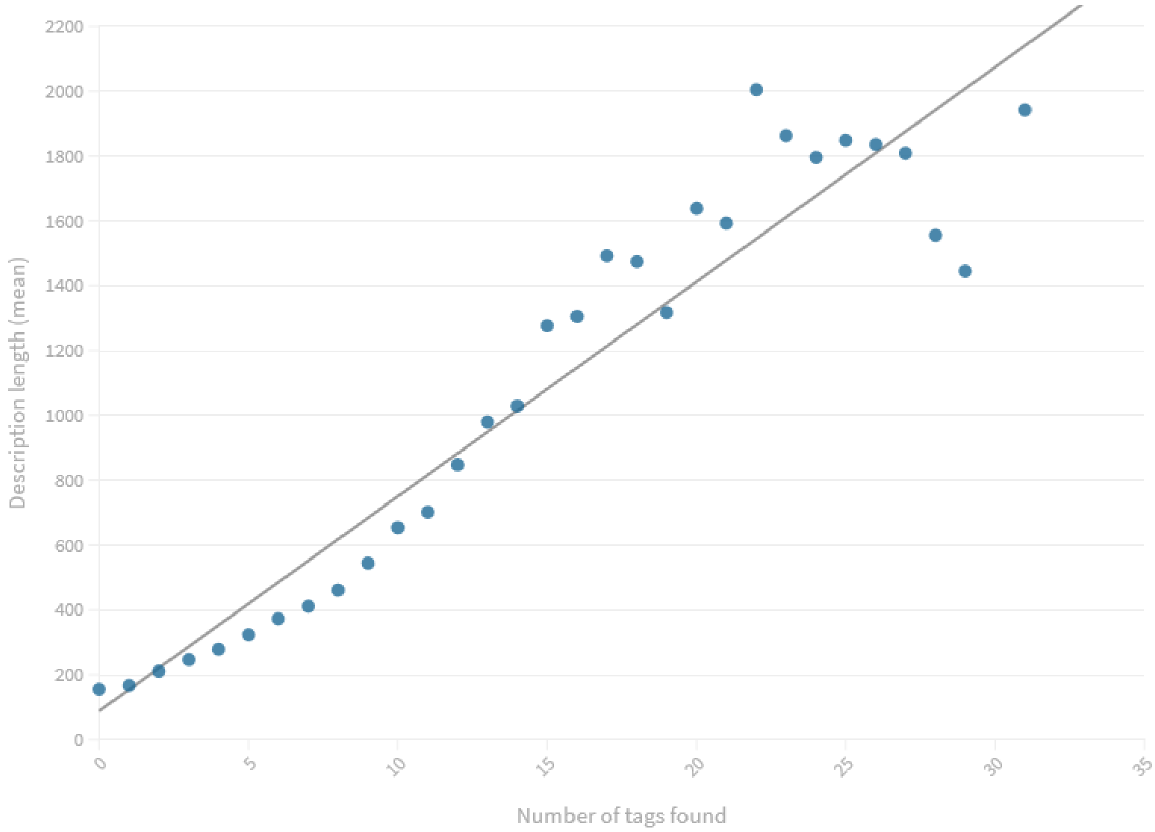
This charity is a multi-purpose community charity which outlines all of its various activities in some detail. As shown in Table 4, it is worth noting that most of the charities with the highest number of tags (over 15) are from Scotland and Northern Ireland, which may reflect the slight differences in how the underlying textual data is collected.

Table 4: Percentage of charities with many tags from each regulator and percentage of all charities in the population

Regulator	Percentage of all charities	Percentage of charities with over 15 tags
England and Wales	84.1	23.1
Northern Ireland	3.4	31.3
Scotland	12.5	45.6

Again, as a feature of the underlying data, it is difficult to avoid a minority of cases being ‘over-tagged’. In particular, the number of tags found is correlated with the length of the combined name and activities description for a charity, so charities with a longer description are more likely to find more tags (see figure 3).

Figure 3: Mean length of name and activities description combined (in number of characters), compared to the number of UK-CAT codes found, per charity.



https://public.flourish.studio/visualisation/7289006/

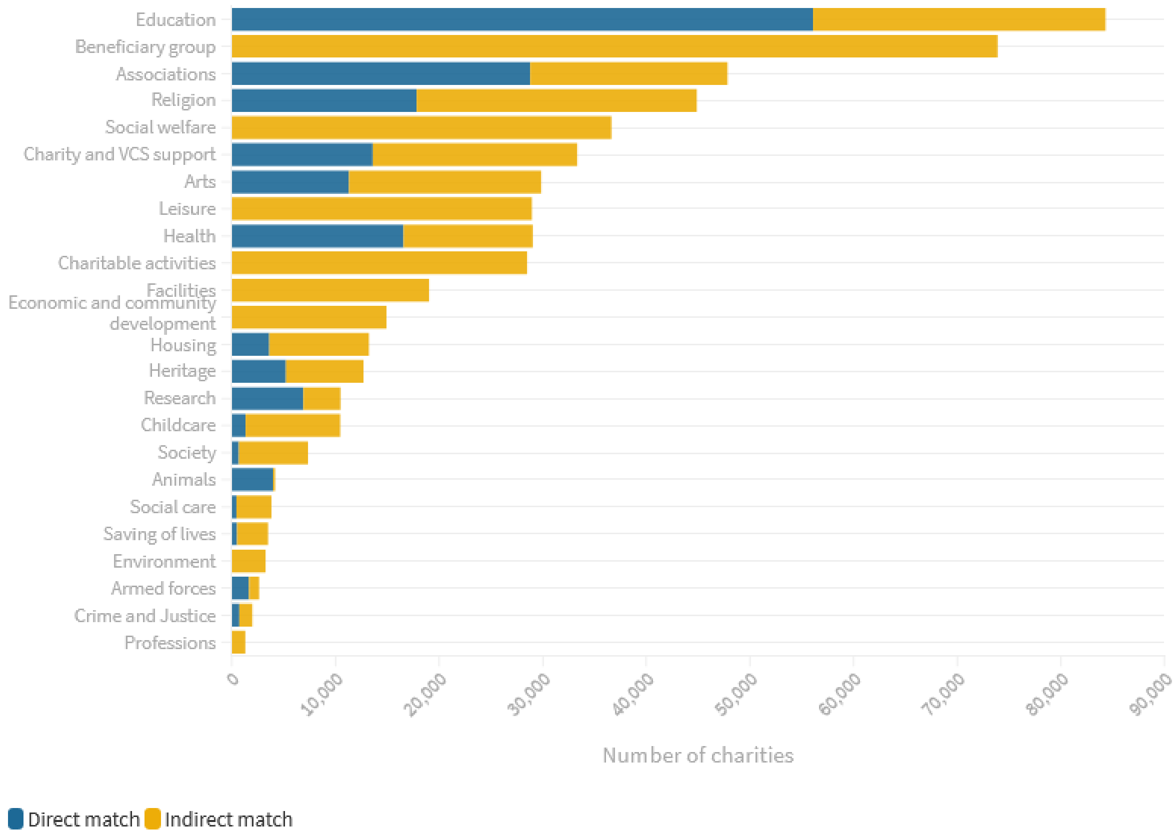
Charities per tag

It is difficult to neatly summarise the results for all of the 254 different tags in the UK-CAT. The full list of results, including the number of charities tagged directly or indirectly (due to lower level tags), is shown in Appendix A.

Comparing groups is more manageable, and the combined matches for all the tags within each group, including the group and sub-group tags, are shown in Figure 4. Some caution is needed, however. Some groups contain a wider range of tags than others, reflecting a wider conceptual scope. ‘Health’, for example, contains 33 tags (including itself and subcategories), whereas ‘Research’ and ‘Saving of lives’ both contain four. Somes tags are also more specific than others, and if a group contains a particularly broad tag, then it will be allocated a high number of ‘indirect’ matches.

Nevertheless, this section examines each of the groups, in descending order of the number of charities matched, whilst also drawing out some of the results for individual tags.

Figure 4: number of charities matched against each UK-CAT group (not mutually exclusive), directly and indirectly



https://public.flourish.studio/visualisation/7289092/

‘Education’ is the most common group, matched by ***42 per cent*** *of charities (28 per cent from direct tags)*. It should be noted, however, that the keyword ‘education’ itself is very broad. Whilst it does tend to indicate that charities are providing some form of education, this can range from formal education through to public awareness raising. The ‘Schools’ tag is also applied very frequently (16 per cent) as well as ‘training’ (10 per cent).

‘Beneficiary group’ is the second most common group, encompassing all those charities that listed any of the specifically beneficiary group tags (***37 per cent*** *of charities, no direct tags*). This is of interest, insofar that it shows that many charities do actually list a defined group, but it is clearly not quite as meaningful as some of the other group categories.

The *‘Associations’* group (***24 per cent*** *of charities, 14 per cent from direct tags)*. The keywords for the tag itself are again quite general, including ‘association’ and ‘club’. Even so, it is revealing that so many charities do use this type of language. Billis (2010) argues that the charity sector’s root identity is associational, capturing the idea of normal people voluntarily banding together to achieve a common task or goal. This also matches the fact that almost all charities are small and local in nature. Individual tags in this group include youth groups (five per cent), leagues of friends (four per cent), and scouting (two per cent).

As with education, charity has long been connected to religion in the UK, the fourth most commonly matched group. ***22 per cent*** of charities match at least one tag within the *‘Religion’* group (*nine per cent from direct matches*). 15 per cent of charities are tagged with ‘Religious activities’ and 11 per cent with ‘Church or place of worship’. ‘Christianity’ of various kinds is the most common religion tagged (16 per cent), with Islam second at 1.2 per cent, and Judaism third at 0.8 per cent. Many of these charities are again quite associational in nature, listing a range of social welfare activities as well as their religious orientation.

Tags from the ‘*Social welfare*’ group were matched by ***18 per cent*** *of charities*, which is in part due to 11 per cent matching the tag for ‘Individual poverty’, another cornerstone of charitable activity in the UK and perhaps the one most commonly associated with what it means to be ‘charitable’. This is a fairly broad group, however, which also includes ‘Social activities’ (three percent of charities), and ‘Food’ (also three percent).

Although it is more difficult to objectively identify areas where the number of tags is lower than expected, it is perhaps worth noting that within the ‘Social Welfare’ group, only one per cent of charities matched the ‘Abuse’ tag, including 0.15 per cent of charities who matched the ‘sexual abuse’ tag and a tiny 0.05 who matched the ‘Child abuse’ tag. This does not match public attention in recent years towards both historic child abuse cases and sexual harassment, which may reflect the age and historical nature of many charities.

A large number of charities (***16 per cent***) were also matched against ‘*Charity and VCS support*’ tags. Seven per cent of charities matched this group tag directly, generally because they mentioned supporting or donating to other charities in their activities. In most cases, these appear to be organisations and associations involved in fundraising and grant making on a relatively small scale. As such, the ‘Fundraising’ tag also matched eight per cent of all charities.

Tags in the arts groups are also matched relatively frequently (***15 per cent*** of charities, six per cent via direct tags), with tags in the 'Performing art' sub-group matching against seven per cent of charities, and the sub-group 'Media and Publishing' matching five per cent. The most common non-group tag is 'Literature' (four per cent of charities) though the keywords 'literature' and 'books' are quite general, matching against references to religious literature, as well as providing school books.

As mentioned above, the 'Health' group (***14 per cent,*** *8 per cent from direct tags*) contains many individual tags, including those in the 'Health condition' sub-group (five per cent of charities), 'Health services' (four per cent), 'Healthcare provider' (two per cent). Interestingly, the most common applied individual tags are 'Mental Health' (two per cent) and 'Counselling and therapy', suggesting that the emphasis on physical health might be slightly less pronounced in the charity sector than in public sector spending.

The 'Leisure' group of tags are matched by ***14 per cent*** of charities (*no indirect tags)*, with the individual tags 'Sports' (six per cent) and 'Recreation' (nine per cent) both common. The latter applies to many clubs and associations, highlighting the important role that charities play in developing social capital within many communities. Arguably, the connection between charity and leisure is sometimes somewhat neglected in public discourse, with a greater focus on charities involved in welfare services.

As discussed in the methodology, the next two groups of tags, 'Charitable activities' (***14 per cent*** of charities) and 'Facilities' (***nine per cent***), sit somewhat outside the usual focus on charitable purposes. Policy and campaigning keywords are matched by three per cent of UK charities, but by far the most prominent tag is 'Grant making' at 9 per cent. This is probably an underestimate, given that many of the organisations which match no tags are in reality likely to be very small grant makers. Only 0.01 per cent of charities mention 'Social investment' in their activities, which is again perhaps out of sync with the level of attention it receives within public debate on the sector.
In terms of facilities, 'Village Hall' is the most popular tag (four per cent of charities), but 'Green space' and 'Community Centre' are also relatively frequently matched (two percent each).

Beyond this point, the matches for the groups become less frequent. Again, it is worth reiterating that in some cases this may be due to the choice of keywords or how broadly the group has been drawn conceptually. Nevertheless, it is clear that the language involved in the following groups is less popular than that of education, association, religion, poverty, leisure, and the other groups addressed above.

Amongst the economic and community development tag group (***seven per cent*** *of charities*), the most popular tags are 'community development' (three per cent), which matches the frequency of community based organisations mentioned above. More obviously 'economic' tags are less common, though 'Unemployment' is matched against two per cent of charities, and 'Rural and farming areas' is also matched against two per cent. As with social investment, however, 'Social enterprise' is matched relatively rarely, at 0.2 per cent of charities.

The ‘Housing’ group of tags is matched by by only ***seven per cent*** of charities, though it should be noted that many medium and large housing associations will not be registered charities, as they are generally regulated by the national regulators of social housing. Only 1.1 per cent of charities match keywords associated with ‘Homelessness’.

The ‘Heritage’ group of tags is somewhat narrowly drawn and is often combined in other schemes with arts and culture. Nevertheless it is still matched by ***six per cent*** of UK charities, most often via the ‘history’ tag (two per cent) and the ‘Historical conservation and restoration’ tag (also two per cent). Childcare is another relatively small and homogenous group of tags, matched by ***five per cent*** of charities, but it does contain the relatively common individual tag for ‘Nursery’ (four per cent). And similarly, the small ‘Research’ group matches five per cent of charities. Only 0.5 per cent, however, are tagged with ‘medical research’ which emphasises the extent to which this activity tends to be consolidated into a few major charities such as Cancer Research.

The ‘Society’ group, which is matched by just ***four per cent*** of UK charities, includes ‘Citizenship’ as its most common tag (two per cent of charities), whereas ‘Conflict resolution’, ‘Equality and diversity’ and ‘Human rights’, ‘Democracy’ and ‘Religious; racial or cross-border harmony’ all match against fewer than one per cent each. Arguably, given the scope and importance of these issues, this is an area where the charity sector is less prevalent than one might expect. Perhaps most notably, only 0.5 per cent matched keywords relating to ‘Racial Justice’, which surely has relevance for ongoing public discussions such as ‘#charitysowhite’.

Some of the remaining groups are quite small or narrow by definition. ‘Armed forces’ is one of these, arguably a sub-category of ‘beneficiary group’. Its keywords were matched by ***one per cent*** of UK charities. Charities which support particular ‘Professions’ are quite difficult to match using keywords (***0.7 per cent***), but even the more specific individual tags for the ‘Clergy’ and ‘Emergency service workers’ do not match high numbers of charities.

Some of the other less commonly matched groups, however, were arguably rarer than we had expected given the debates and discussion regarding the charity sector which occur in a UK context. Notwithstanding the possibility of false negative results, social care, matching just ***two per cent*** of charities, was much less frequent than we had anticipated given the attention paid to this activity area in policy debates.

Defying stereotypes of donkey sanctuaries and cat shelters, only ***two per cent*** of charities matched our list of tags under the ‘animal’ group. Crime and justice related tags are also relatively uncommon, matching just ***one per cent*** of charities, perhaps reflecting the challenges of raising donations for work in this area. Arguably, this low frequency further challenges the stereotype that all charities necessarily work with the most disadvantaged individuals in a society.

The ‘Saving of lives’ group (***two per cent*** of charities) contains tags for ‘Humanitarian relief’ (one per cent) and ‘Search and rescue’ (0.2) per cent, which might have been expected to be higher given the role of many charities in international aid. This is another area where additional variables in the registers of charities may be able to supplement the tags, as international activities are relatively difficult to capture using keywords.

Finally, we were particularly struck by the ‘Environment’ tag group matching against only two per cent of charities’ activities and objects. Only **0.1 per cent** of charities appear to mention terms related to the climate emergency or climate change. Again, this is perhaps out of proportion to the scale of the threat posed to society by the climate emergency.

Discussion

This project was established to try and provide new, alternative ways to classify the activities of charities in the UK, according to their charitable purpose. To do so, we have introduced a new, UK focussed classification scheme, as well as providing a keyword matching method to apply these new categories automatically. Our hope is that by doing so, we have provided a means to view the charity sector in a new light, helping researchers and their audiences to better understand what thousands of charities do day-in, day-out across the UK.

Mapping metaphors are ubiquitous in classification work and research and capture our interest in the ‘shape’ of the charity sector in the UK well. On the other hand, they can obscure the fact that there are many different ways in which the same data could be segmented. Throughout the process of determining the classifications themselves and the keywords used to apply them, we have had to make subjective judgements, which risk privileging our own perspectives and biases on the sector. Our aim has been to include those categories which will be useful and relevant as possible for a wide range of researchers working on the UK charity sector, but only time and feedback will reveal how successful we have been. Certainly, the results have confirmed the great variety, occasionally even quirkiness, of charity purposes to be found in the UK.

Nevertheless, our own expectations about the frequency with which different tags have been matched have been challenged on occasion. Partly, this may be because charities infrequently update their activities, which means that the textual data available is somewhat historic. Some of the most common tags relate to purposes and activities that would not have looked out of place in Victorian times. Partly that is because they are somewhat timeless, such as individual poverty or community-based social activities, partly it is because these causes have had the most time to attract new charities compared to newer causes. Today’s prominent causes may lead to tomorrow's charities. On the other hand, the results are an important reminder that while income and media attention may flow to controversial areas of work, such as social care services, these areas account for a tiny proportion of charities in terms of numbers.

Regardless, it is hard to avoid the impression that whilst the charity sector covers a range of important causes, its organisational resources are not necessarily targeted according to any overwhelming logic of need or importance. Climate change poses an existential threat to modern societies with profound consequences for vulnerable groups, and yet there are relatively few references to it in the activities of UK charities. This fits with theoretical arguments that the voluntary sector suffers from philanthropic insufficiency (Salamon, 1987), with resources sometimes being lowest when and where they are needed most. Charities associated with the most ‘sympathetic’ causes, such as children and young people, find it easier to attract resources than more marginalised groupings or the ‘undeserving poor’. Rather than the putative safety blanket of the welfare state, charity provides a welcome but uneven patchwork.

Developing over 250 tags and using them to manually tag over 4,000 charities, as well as creating keyword search tools for each tag, has been no small task. Whilst we hope the endeavour has been successful overall, the project inevitably has its limitations.

The fact that we arrived at so many tags, despites efforts to rationalise and avoid those with very few matches, is testament to the number of different causes covered by the UK charity sector. Hopefully, the increased number of tags, and the fact they have been developed inductively from UK charity data, will help to capture the sector's activities in a way that is useful to a wide range of users. On the other hand, as a result of its length, the scheme may comparatively lack what Salmon and Anheier refer to as economy (parsimony) or organising power (generalisability).

On the other hand, no classification scheme is ever quite long enough, and the UK-CAT is likely to have missed categories that some may find important. Once you start listing examples of individual religions, sports, animals, you inevitably have to draw a cut-off line which some may find frustrating. Whilst we would be keen to hear of any significant gaps in the UK-CAT, we would also encourage anyone in this position to try and develop their own keyword search rule for that activity area and share the results.

The development of keyword matching rules brings a different set of challenges. Firstly, we have a renewed appreciation for the level of ambiguity and subjectivity involved in developing and applying a classification system. The process is as much art as it is science, particularly when there is no limit on how many tags to apply to each charity. Even when discussed between several individuals, all with knowledge of the sector, it can occasionally be hard to find consensus when conducting the manual matching. And in practice, we generally had to rely on a single individual to make the call. Clearly, therefore, it is hard to ever determine a ‘correct’ classification. One of the key advantages of keyword searching, therefore, is its transparency. Anyone is able to view the keywords we have chosen, identify any mistakes, and suggest improvements. We could welcome this feedback on either the UK-CAT or the keyword search rules[6].

We made the decision to apply as many tags as felt necessary to capture the charitable purpose of each charity. As discussed in the methodology, this has advantages and disadvantages. Whilst limiting its usefulness for some types of research, multiple tags help to avoid the challenges otherwise posed by multi-purpose charities. The analysis provided in this paper helps to show that the keyword matching is working relatively well, with few obvious false positives and a respectable rate of false negatives (if we assume that the manually coded classifications are ‘correct’). The vast majority of charities have at least one tag, and the numbers with many tags are manageable. What is clear, however, is that the keyword searching process applies many more tags than a human coder. The relatively simple matching process has no means of determining how important any individual tag is to the overall meaning of what is being described.

[6] https://charityclassification.org.uk/contact/

We hope that in future, it may be possible to apply a more sophisticated method which provides a ‘relevance’ score for each tag, for every charity. By setting various thresholds, this could provide the best of all worlds, allowing us to apply multiple tags, whilst ranking their relevance and picking a single tag, ‘most relevant tag’, where needed.

Natural Language Processing (NLP) techniques, potentially combined with machine learning, may in time be able to outperform our own efforts at classification rules and solve some of the issues to do with context, spelling and language which we experienced. Results from the US suggest there is great promise from this approach (Ma, 2020), though it should be borne in mind that much less training data is available in the UK (automatically applying the charity register classifications would be the most analogous comparison). It may be that a hybrid approach is possible, with human coders checking and correcting the results of machine learning in order to create an iteratively better training dataset over time.

Applying NLP provides a logical next step to help address some of the issues with language, spelling and words with multiple meanings we have encountered. Indeed, we have already begun experimenting with some of these methods. Machine learning and NLP may also come at the cost of transparency and human oversight, and can only ever be as good as the underlying training dataset. It is also difficult to add new categories (as there is no training data), a concern given the UK-CAT is still in its infancy.

Finally, it should be acknowledged that charities do not make up the entirety of the UK non-profit sector, which also includes CICs and other forms of social enterprise, as well as informal and unregistered organisations.

References

Salamon and Anheier, 1992

Kendal and Knapp, 1995

Barman, 2013

Cooney, 2020

Clifford et al. 2013

NCVO, 2012

Newbigging et al. 2017

Hay, 2002

Bowker and Star, 1999

Carmel and Harlock, 2010

Salamon, 2010

Appe (2012)

DEI data group

Fyall et al. 2018

Lampkin et al. 2001

Ma, 2020

Litofcenko et al., (2020)

The Non-Profit centre, 2008

Appendix A: UK Charity Activity Tags (UK-CAT) classification system

The system is organised into 24 categories, with 17 subcategories and 230 tags in total. It is designed to accommodate charities having more than tag applied.

The most up-to-date version of this classification system is available at https://charityclassification.org.uk/data/tag_list/.

Details of ICNP/TSO and other classification schemes are available at https://charityclassification.org.uk/other-work/#other-classification-schemes

Category	**Subcategory**	**Tag**	**Code**	**Matched charities**
Animals		Animals	AN	3,961
Animals		Cats	AN101	393
Animals		Dogs	AN102	914
Animals		Donkeys	AN103	59
Animals		Horses	AN104	619
Armed forces		Armed forces	AF	1,663
Armed forces		Army	AF101	657
Armed forces		Navy	AF102	543
Armed forces		RAF	AF103	410
Armed forces		Veterans	AF104	354
Arts		Arts	AR	11,310
Arts		Culture	AR101	
Arts		Festival	AR102	2,460
Arts		Languages	AR103	291
Arts	Media and publishing	Media and publishing	AR200	
Arts	Media and publishing	Film	AR201	1,141
Arts	Media and	Literature	AR202	7,600

	publishing			
Arts	Media and publishing	Media	AR203	393
Arts	Media and publishing	Print media	AR204	571
Arts	Media and publishing	Radio	AR205	454
Arts	Media and publishing	Television	AR206	370
Arts	Performing art	Performing art	AR300	887
Arts	Performing art	Choirs	AR301	2,601
Arts	Performing art	Dance	AR302	2,636
Arts	Performing art	Music	AR303	9,255
Arts	Performing art	Musical theatre	AR304	292
Arts	Performing art	Opera	AR305	696
Arts	Performing art	Orchestra	AR306	717
Arts	Performing art	Theatre	AR307	4,142
Arts		Visual arts	AR104	2,338
Associations		Associations	AS	28,809
Associations		Community association	AS101	3,505
Associations		Fraternal societies	AS102	681
Associations		Inner Wheel	AS103	293
Associations		League of Friends	AS104	8,350
Associations	Service clubs	Service clubs	AS200	329
Associations	Service clubs	Lions club	AS201	606
Associations	Service clubs	Rotary club	AS202	1,351
Associations		Social club	AS105	2,066
Associations		Townswomen's	AS106	220

		Guild		
Associations		Women's Institute	AS107	2,247
Associations		YWCA / YMCA	AS108	207
Associations	Youth Groups	Youth Groups	AS300	1,907
Associations	Youth Groups	Cadets	AS301	732
Associations	Youth Groups	Girlguiding	AS302	2,006
Associations	Youth Groups	Scouting	AS303	4,528
Beneficiary group		Beneficiary group	BE	
Beneficiary group		Asylum seekers and refugees	BE101	1,196
Beneficiary group		Children	BE102	26,728
Beneficiary group		Families	BE103	
Beneficiary group		Girls	BE104	4,596
Beneficiary group		LGBTQ+	BE105	279
Beneficiary group		Men	BE106	1,234
Beneficiary group		Migrants	BE107	532
Beneficiary group		Older people	BE108	8,235
Beneficiary group		Parents and guardians	BE109	10,305
Beneficiary group	People with disabilities	People with disabilities	BE200	13,967
Beneficiary group	People with disabilities	Riding for the disabled	BE201	405
Beneficiary group		People with learning disabilities	BE110	2,263
Beneficiary group		Racial; ethnic or national communities	BE111	595
Beneficiary group		Widows; widowers and orphans	BE112	2,169

Beneficiary group		Women	BE113	9,944
Beneficiary group		Young children	BE114	8,911
Beneficiary group		Young people	BE115	21,681
Charitable activities		Charitable activities	CA	
Charitable activities		Advice and individual advocacy	CA101	3,276
Charitable activities		Charity shops	CA102	889
Charitable activities	Grant making	Grant making	CA200	18,526
Charitable activities	Grant making	Grants to individuals	CA201	1,158
Charitable activities	Grant making	Grants to organisations	CA202	8,684
Charitable activities		Policy campaigning and advocacy	CA103	6,433
Charitable activities		Social Investment	CA104	23
Charity and VCS support		Charity and VCS support	CV	13,581
Charity and VCS support		Financial investment	CV101	54
Charity and VCS support		Fundraising	CV102	16,395
Charity and VCS support		Umbrella bodies	CV103	857
Charity and VCS support		Volunteering	CV104	6,257
Childcare		Childcare	CC	1,341
Childcare		Nursery	CC101	7,139
Childcare		Out of school club	CC102	1,547
Childcare		Playground	CC103	853
Childcare		Playgroup	CC104	2,393

Crime and Justice		Crime and Justice	CJ	733
Crime and Justice		Offender support and rehabilitation	CJ101	782
Crime and Justice		Prevention and safety	CJ102	186
Crime and Justice		Road safety	CJ103	274
Crime and Justice		Trafficking and modern slavery	CJ104	157
Crime and Justice		Victim support	CJ105	100
Economic and community development		Economic and community development	EC	
Economic and community development		Community development	EC101	5,459
Economic and community development		Economic development	EC102	1,402
Economic and community development	Infrastructure	Infrastructure	EC200	
Economic and community development	Infrastructure	Energy	EC201	171
Economic and community development	Infrastructure	Water	EC202	305
Economic and community development		International development	EC103	1,032
Economic and community development		Planning and architecture	EC104	186
Economic and community development		Rural and farming areas	EC105	3,961

Economic and community development		Social enterprise	EC106	436
Economic and community development		Unemployment	EC107	4,243
Economic and community development		Urban areas	EC108	835
Education		Education	ED	56,101
Education		Adult education	ED101	387
Education		Further education	ED102	2,875
Education		Higher education	ED103	2,300
Education		Primary education	ED104	5,414
Education	School support	School support	ED200	4,111
Education	School support	Parent teacher	ED201	5,355
Education	School support	School fundraising	ED202	5,032
Education		Schools	ED106	31,287
Education		Secondary education	ED107	1,544
Education		Student support	ED105	2,236
Education		Student union	ED108	132
Education	Training	Training	ED300	16,946
Education	Training	Basic skills	ED301	123
Education	Training	ESOL	ED303	256
Education	Training	Employability training	ED302	4,504
Education	Training	IT and digital	ED304	350
Education	Training	Mentoring	ED305	1,220
Education	Training	Vocational training	ED306	394

Education		University of the Third Age	ED109	832
Environment		Environment	EN	
Environment		Climate Emergency	EN101	263
Environment		Conservation and sustainability	EN102	1,843
Environment		Recycling	EN103	400
Environment		Wildlife	EN104	1,213
Facilities		Facilities	FA	
Facilities		Cemetery	FA101	589
Facilities		Community cafe	FA102	816
Facilities		Community centre	FA103	3,348
Facilities		Green space	FA104	3,486
Facilities		Open spaces	FA105	1,087
Facilities		Playing fields	FA106	2,598
Facilities		Village hall	FA107	8,861
Facilities		Youth centre	FA108	260
Health		Health	HE	16,573
Health	Health condition	Health condition	HE200	1,124
Health	Health condition	Addiction and dependency	HE201	1,149
Health	Health condition	Cancer	HE202	1,827
Health	Health condition	Cerebral palsy	HE203	87
Health	Health condition	Chronic Fatigue Syndrome	HE204	33
Health	Health condition	Dementia	HE205	553
Health	Health condition	Fibromyalgia	HE206	20
Health	Health condition	HIV / Aids	HE208	387

Health	Health condition	Hearing loss	HE207	628
Health	Health condition	Maternity	HE209	483
Health	Health condition	Mental health	HE210	3,401
Health	Health condition	Motor Neurone Disease	HE211	28
Health	Health condition	Multiple Sclerosis	HE212	146
Health	Health condition	Sickle Cell	HE213	46
Health	Health condition	Strokes	HE214	252
Health	Health condition	Visual impairment	HE215	1,187
Health	Health services	Health services	HE300	
Health	Health services	Alternative medicine	HE301	51
Health	Health services	Ambulance service	HE302	371
Health	Health services	Complementary therapies	HE303	134
Health	Health services	Counselling and therapy	HE304	3,075
Health	Health services	Health and wellbeing	HE305	1,866
Health	Health services	Nursing	HE306	864
Health	Health services	Palliative care	HE307	417
Health	Health services	Physiotherapy	HE308	127
Health	Health services	Surgery	HE309	668
Health	Healthcare provider	Healthcare provider	HE400	1,020
Health	Healthcare provider	Hospice	HE401	607
Health	Healthcare provider	Hospital	HE402	3,200
Health	Healthcare provider support	Healthcare provider support	HE500	362

Health	Healthcare provider support	Friends of healthcare provider	HE501	664
Health	Healthcare provider support	Healthcare provider fundraising	HE502	
Heritage		Heritage	HR	5,225
Heritage		Archaeology	HR101	493
Heritage		Historical conservation and restoration	HR102	3,419
Heritage		History	HR103	4,607
Heritage		Monuments; statues and memorials	HR104	601
Heritage		Museum	HR105	2,356
Heritage		Natural history	HR106	233
Housing		Housing	HO	3,604
Housing		Accommodation	HO101	8,353
Housing		Almshouse	HO102	1,296
Housing		Homelessness	HO103	2,251
Housing		Housing association	HO104	227
Housing		Temporary or emergency housing	HO105	545
Leisure		Leisure	LE	
Leisure		Exercise and fitness	LE101	2,549
Leisure		Gardening	LE102	1,175
Leisure		Hobbies	LE103	408
Leisure		Outdoor pursuits	LE104	1,027
Leisure		Recreation	LE105	18,022
Leisure		Sports	LE106	12,981
Professions		Professions	PR	

Professions		Clergy	PR101	543
Professions		Emergency service workers	PR102	62
Professions		Healthcare workers	PR103	376
Professions		Miners	PR104	342
Religion		Religion	RL	17,850
Religion		Baha'i	RL101	77
Religion		Buddhism	RL102	352
Religion	Christianity	Christianity	RL200	32,805
Religion	Christianity	Church of England	RL201	2,244
Religion	Christianity	Church of Ireland	RL202	393
Religion	Christianity	Church of Scotland	RL203	1,450
Religion	Christianity	Jehovah's Witnesses	RL204	1,359
Religion	Christianity	Roman Catholic	RL205	504
Religion	Christianity	Society of Friends (Quakers)	RL206	216
Religion		Hinduism	RL103	365
Religion		Islam	RL104	2,429
Religion		Jainism	RL105	35
Religion		Judaism	RL106	1,723
Religion	Religious activities	Religious activities	RL300	7,345
Religion	Religious activities	Chaplaincy	RL301	241
Religion	Religious activities	Church or place of worship	RL302	22,722
Religion	Religious activities	Parochial Church Council	RL303	2,937
Religion	Religious activities	Religious education	RL304	1,954
Religion	Religious activities	Religious ministry	RL305	7,225

Religion		Sikhism	RL107	280
Religion		Spiritualism	RL108	275
Research		Research	RS	6,915
Research		Medical research	RS101	1,065
Research		Philosophy	RS102	314
Research		Science	RS103	3,969
Saving of lives		Saving of lives	SL	511
Saving of lives		Emergency services	SL101	952
Saving of lives		Humanitarian relief	SL102	1,708
Saving of lives		Search and rescue	SL103	443
Social care		Social care	SC	453
Social care		Adult day care	SC101	777
Social care		Carer support	SC102	302
Social care		Children in care	SC103	145
Social care		Children's homes	SC104	642
Social care		Domiciliary care	SC105	215
Social care		Residential care	SC106	798
Social care		Residential care with nursing	SC107	57
Social care		Respite	SC108	730
Social welfare		Social welfare	SW	
Social welfare	Abuse	Abuse	SW200	1,218
Social welfare	Abuse	Child abuse	SW201	95
Social welfare	Abuse	Domestic abuse	SW202	638
Social welfare	Abuse	Refuge or shelter	SW203	293
Social welfare	Abuse	Sexual abuse	SW204	297
Social welfare		Benevolent Society	SW101	1,674

Social welfare		Bereavement	SW102	964
Social welfare		Clothes	SW103	1,339
Social welfare		Community transport	SW104	861
Social welfare	Food	Food	SW300	5,484
Social welfare	Food	Food banks	SW301	1,059
Social welfare		Individual poverty	SW105	22,607
Social welfare		Loneliness	SW106	2,698
Social welfare		Social activities	SW107	5,360
Society		Society	SO	689
Society		Citizenship	SO101	4,555
Society		Conflict resolution	SO102	318
Society		Democracy	SO103	61
Society		Equality and diversity	SO104	1,289
Society		Human rights	SO105	848
Society		Racial justice	SO106	109
Society		Religious; racial or cross-border harmony	SO107	970