Mapping Career Causeways: Crowdsourcing feasibility ratings of career transitions

George Richardson[1], Karlis Kanders[1]

Technical appendix
January 2021

This technical appendix accompanies the crowdsourced dataset of career transition feasibility ratings, obtained by Nesta as part of the Mapping Career Causeways project. We describe the methodology and results of the crowdsourcing study, and develop a machine learning model to predict feasibility ratings for career transitions between any two occupations from the European occupational framework ESCO.

Table of Contents

---

[1] Nesta, 58 Victoria Embankment, London, EC4Y 0DS; open.jobs@nesta.org.uk.
We especially thank Genna Barnett and Cath Sleeman for managing the project and providing valuable insights.

Introduction

As part of the Mapping Career Causeways project, funded by J.P. Morgan, Nesta has used data science and machine learning methods to develop an algorithm for recommending career transitions. This algorithm identifies viable and desirable career transitions for any occupation in the European multilingual framework of Skills, Competences, Qualifications and Occupations (ESCO). Viable transitions are to occupations that are similar to the worker's existing role, in terms of the skills that are required, as well as other occupational factors such as the typical work activities; interpersonal, physical and structural work context aspects; and the expected education and experience levels. Desirable transitions are the subset of viable options which offer comparable or higher levels of pay.

Additionally, the algorithm can adjust its outputs to take into account occupations' risk of automation, and suggest career moves only to occupations that reduce workers' exposure to automation risks - we call these safe and desirable transitions.

The algorithm is described in complete detail in the report Mapping Career Causeways: Supporting workers at risk, and the corresponding code for generating transition recommendations has been released on the project's github repository.

Here, we describe a crowdsourcing study that Nesta has subsequently carried out to determine whether the transitions recommended by the algorithm represent reasonable suggestions. That is, whether an imagined worker in the origin occupation could reasonably make the transition to the suggested destination, based on a "common sense" judgement by members from the public. There is of course a large degree of subjectivity and personal context that would affect what is considered reasonable and we attempt to consider how this affects both the experiment design and the interpretation of any results.

This technical paper is organised as follows. In the Methodology section, we describe the design of the crowdsourcing study, state our initial research questions, and describe a machine learning model that we developed to predict a feasibility rating for any transition between the ESCO occupations. In the Results section, we first explore the crowdsourced dataset and establish its suitability for further analyses. We then characterise the correspondence between estimates of occupation similarity derived by our algorithm with the crowdsourced feasibility ratings, and build upon the feasibility ratings to arrive at a refined model for recommending career transitions. Finally, we close with a Discussion section where we consider the interpretation and communication of the crowdsourcing results.

Methodology

We used a crowdsourcing platform to obtain feasibility judgements on occupation transitions from the UK population. The total number of all possible transitions between the 1,627 top level ESCO occupations is over 2.5 million, and the volume of the subset of all safe and desirable transitions identified by the algorithm is over 65,000. For the purposes of this study, we created a sample of 9,113 transitions, where we included all safe and desirable transitions from at-risk occupations[2] (4,813) and a further sample of transitions (4,300) from across the distribution of occupation similarity. The latter also included a portion of transitions that were not identified as viable by the algorithm.

To obtain crowd judgements, we used a crowdsourcing platform provided by 1715 Labs. Each contributor was asked a set of one-off pre-qualifying questions about their age, gender, occupational experience, qualifications and geographic location. Subsequently, they would make judgements on individual occupation transitions, being asked for their familiarity with each occupation, a feasibility judgement for the transition, and the reasons for their decision. To aid the contributors in making their judgments, for each transition they were presented with the titles and descriptions of both destination and origin occupations, and a comparison of their skills sets. The full survey design can be found in the Survey design section of the Appendix.

Each contributor could assess as many transitions as they desired up until the job was complete (when all transitions had been evaluated). Each transition was assessed by at least 5 contributors, although those with a higher level of variation in the feasibility rating were shown to more contributors to increase confidence in the results. Contributors were paid the local living wage for their time spent on this task.

To ensure the validity of the crowdsourced data, 1715 Labs use several methods for removing so-called 'bad actors' - those who use bots or supply answers simply to progress through the survey. These include contributors who:

- Are very fast and have a low variance in the time taken to answer;
- Provide inconsistent answers to particular pre-qualifying questions (e.g. age);
- Whose answers to questions appear to be random.

All responses provided by the bad actors have been removed from the published and analysed dataset.

[2] 'At-risk' occupations were defined in the 'Mapping Career Causeways: Supporting workers at risk' report as having both a high level of overall automation risk and low prevalence of so-called bottleneck tasks (see the report for further details).

Research questions

We have restricted the scope of this technical appendix to describing the crowdsourced dataset, and answering the following initial research questions:

- What is the correspondence between the crowdsourced feasibility ratings and the occupation similarity measures derived by our career transition algorithm?
- How can the career transitions algorithm be refined to remove recommendations that the public may consider highly unfeasible.

Modelling crowd feasibility ratings

To extend the feasibility ratings to any transitions between two occupations, we trained a regression model using the ratings provided for the 9,113 transitions in the crowdsourced dataset. In this way, we are able to provide both actual feasibility ratings for a sample of the transitions and predicted ratings for those transitions, as well as all others recommended by the algorithm. The model was trained to predict the mean crowd feasibility rating using features derived from similarities between the skills, work activities and descriptions of the two occupations in each transition.

Procedure for rejecting transitions

Given the large number of transitions recommended by the algorithm, it was not possible to vet all of them. However, from inspecting the feasibility ratings of transitions assessed by the crowd it was clear that certain transitions did not meet the standard of reasonable judgement. We had anticipated that this might be the case, and undertook a process to automatically remove such transitions from the recommendations. We employed the predictive model described above to create estimated feasibility ratings for all 65,000 safe and desirable transitions[3], and established a minimum threshold for which transitions would be included. Those which were predicted a feasibility rating below this threshold were flagged and removed from the recommendations.

[3] More generally, this approach can be applied to any transition between the 2,942 different, broader and narrower ESCO occupations.

Results

Crowdsourced feasibility ratings

We obtained over 70,000 ratings for 9,113 transitions, between nearly 1,600 occupations. The ratings were generated by 387 unique contributors. In the following sections, we explore the crowdsourced dataset and establish its suitability for further analyses.

Contributors

Most of the study contributors were in their 20s or 30s (Figure 1) and located in England (Figure 2). 60% and 39% of the contributors indicated their gender as male or female, respectively.

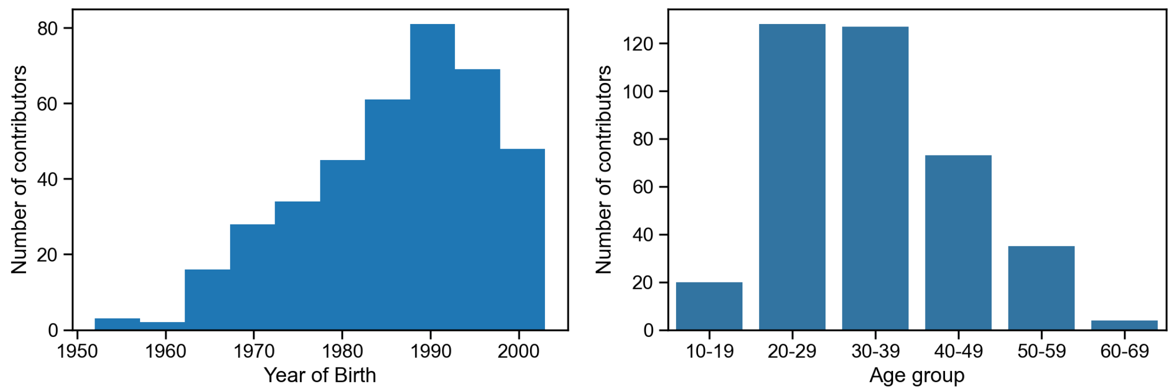

Figure 1. Year of Birth (left) and the corresponding age groups (right) of the contributors. The youngest contributor was born in 2013 and the oldest contributor in 1952.

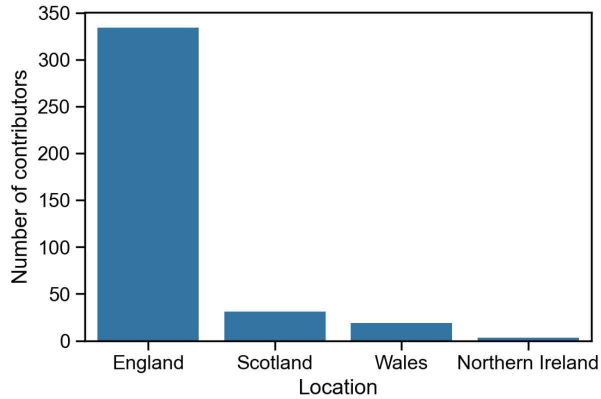

Figure 2. Locations of the contributors.

Contributors had a broad span of qualifications and employment backgrounds. Their highest qualification ranged between GCSE and Doctorate, with most of the participants having completed a Bachelor's or Master's degree (Figure 3). Among the most popular areas of contributors' work experience were sales and services; business and administration; education; food, drink and tobacco; and ICT (Figure 4).

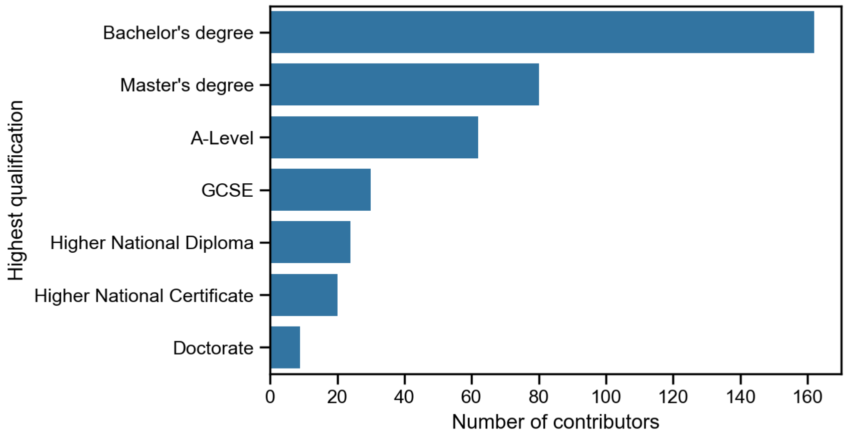

Figure 3. Highest level of contributors' qualifications.

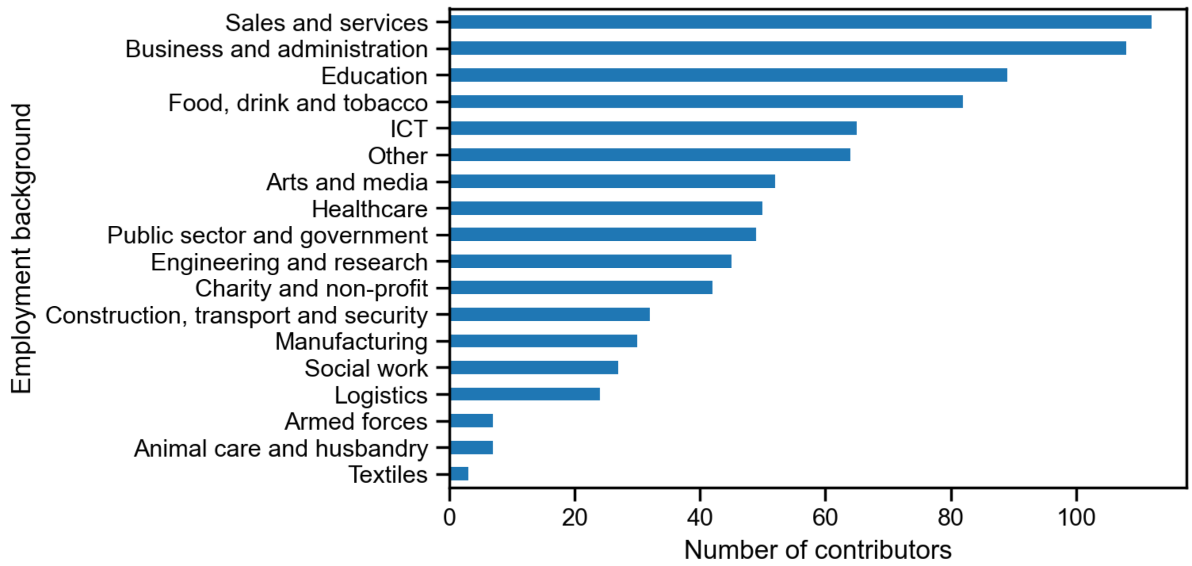

Figure 4. Sectors in which the contributors have worked in (note that contributors could choose more than one sector).

Individual contributor activity

Individual contributors provided widely varying levels of contributions, with the median number of ratings per contributor being 3, and about 30% of the contributors providing only one rating (Figure 5). 10% contributors (90th percentile) provided 216 ratings or more, with one respondent making 6786 contributions. This indicates that although each transition was evaluated by on average eight different contributors (Figure 6), a relatively small percentage of 'core contributors' generated the majority of results. Therefore, the aggregated results for any transition will be necessarily skewed by the opinions of this group. Nonetheless, the study still samples a broad range of opinions, with 57 people making at least 100 contributions each.

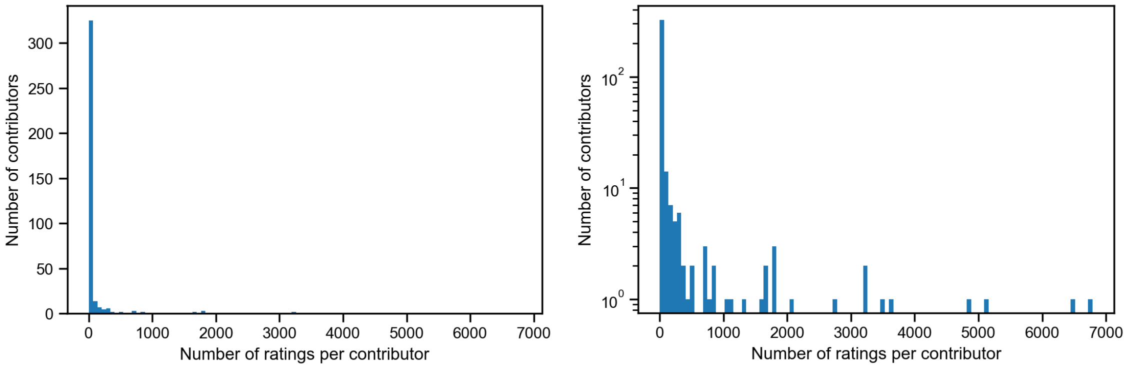

Figure 5. Histograms of the number of ratings per contributor, shown using linear (left) and logarithmic vertical axis (right). Note the "super users" at the tail of the distribution.

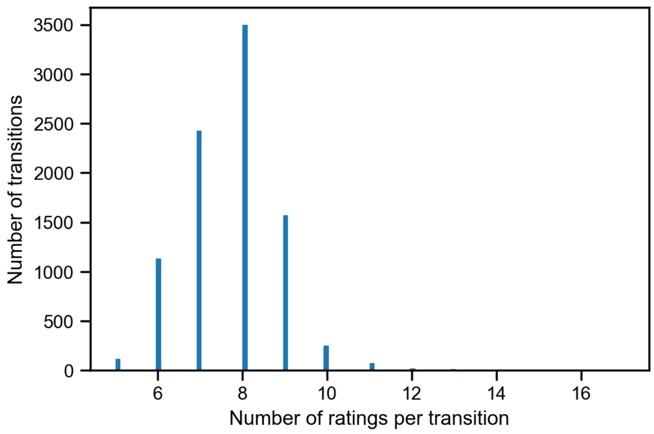

Figure 6. Number of ratings per transition (median number of ratings is 8).

Familiarity with the occupations

Contributors rated their familiarity with the origin and destination occupations on an integer scale between 1 (not at all familiar) and 5 (very familiar). For the majority of ratings, contributors had at least some familiarity with both transition occupations, and the distributions of familiarity with the origin and destination occupations appear to be roughly normal and equivalent (Figure 7). In about 40,000 cases, the contributors felt equally familiar with both the origin and destination occupation, and large differences in familiarity between the origin and destination occupations were generally rare (Figure 8).

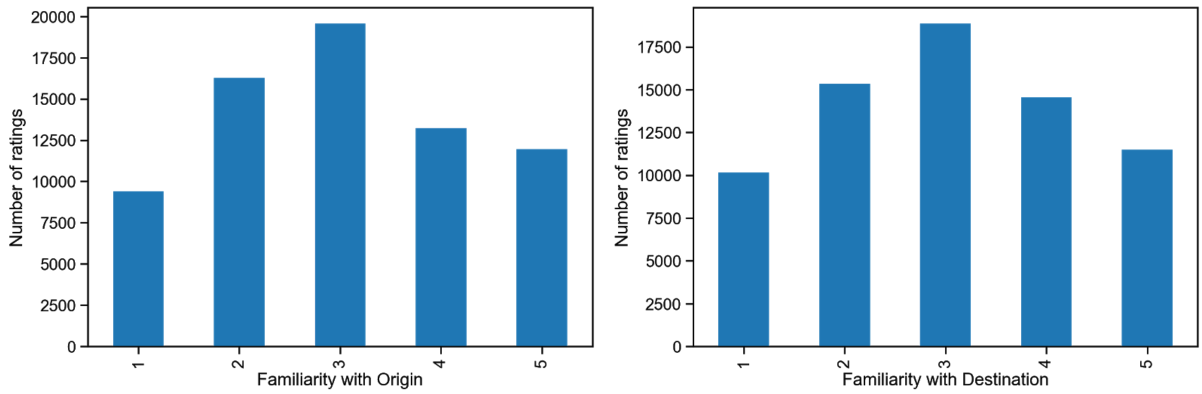

Figure 7. Contributor familiarity with origin (left) and destination (right) occupations across all ratings (mean≈3.03 for both distributions).

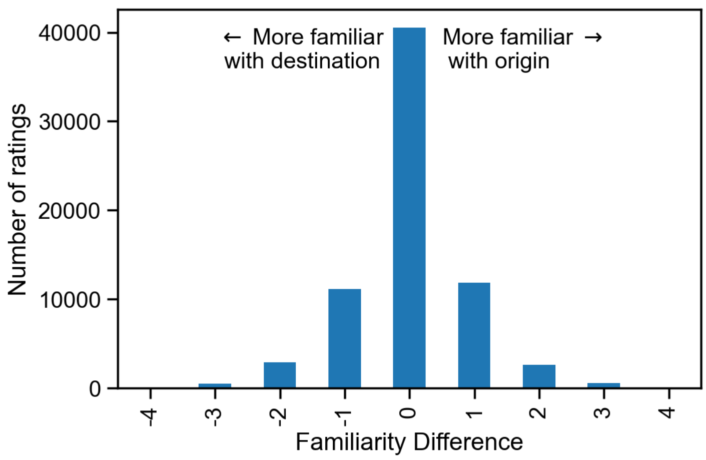

Figure 8. Difference between the familiarity with the origin and destination occupations.

Among the occupations that were on average the most familiar to contributors are taxi and delivery drivers, art models, actors and actresses, and astronauts, whereas the least familiar occupations are mine safety officers, electrolytic cell makers and 'on foot aquatic resources collectors'. Generally, it appears that occupations with technical, or long and very specific titles were less familiar to the contributors.

While contributors' familiarity with the occupations might influence their judgement on the feasibility of the transitions, we found only a weak - albeit statistically significant - preference for contributors to assign transitions with more familiar occupations a higher feasibility rating (Figure 9). Correlation coefficients between feasibility and familiarity with either origin or destination occupations were around 0.20, which can be interpreted as familiarity explaining about 4% of the variance in feasibility ratings.

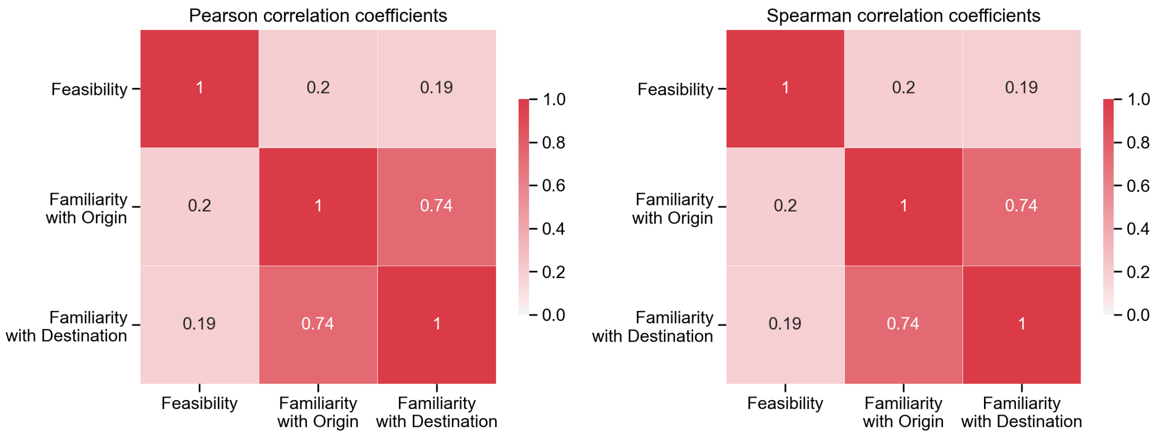

Figure 9. Pearson (left) and Spearman (right) correlation coefficients between feasibility and familiarity variables across all ratings. All correlations are statistically significant (p-value=0).

When the individual ratings are aggregated and the mean familiarity and feasibility for each transition is assessed, we find similarly weak correlations (Figure 10). The standard deviation of the transition feasibility ratings was also only weakly correlated with the familiarity ratings.

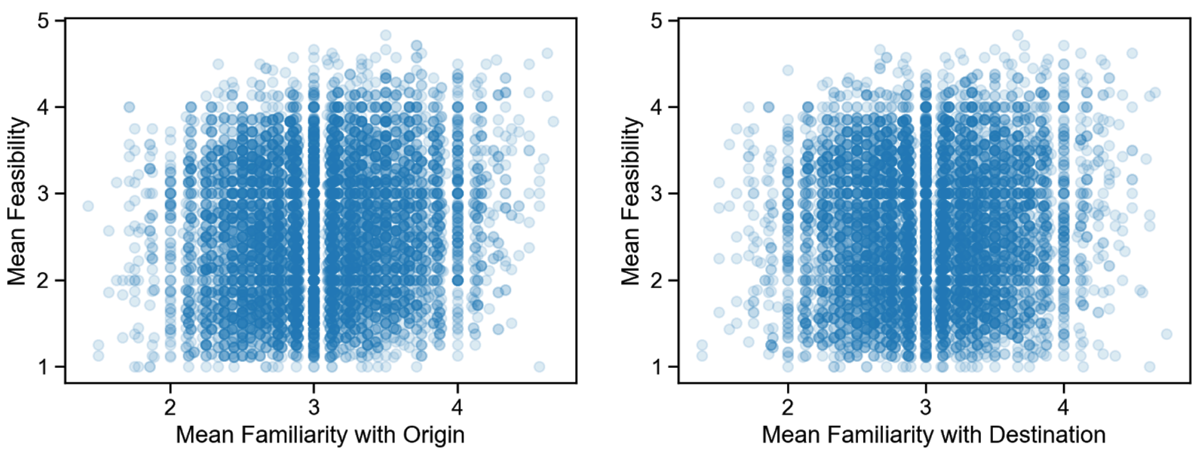

Figure 10. Relationship between mean feasibility and mean familiarity with origin (left) and destination (right) occupations across all tested transitions; Spearman correlation coefficients are 0.12 and 0.05, respectively (p-value = 0).

Taken together, these observations suggest that contributors' familiarity or unfamiliarity with the occupations did not exert a strong effect on their perceptions of the transition feasibility. This in turn suggests that the contributors were able to make consistent judgments based on the provided information about occupations and skills. Therefore, in further analysis and modelling, we have included all transition feasibility ratings regardless of the familiarity (while acknowledging that a more refined criteria for including the ratings could be used in future investigations).

Feasibility ratings

The feasibility of transitions was rated on an integer scale between 1 (not at all feasible) and 5 (very feasible). The distribution of feasibility ratings is positively skewed around the modal value of 2 (Figure 11). Assuming that feasibility of 3 is the threshold where transitions can be seen as rated 'feasible', we find that around half of the ratings consider the transition in question to be unfeasible. Among the most unfeasible transitions we find, for example, the transition from food analyst to aviation inspector, or from advanced physiotherapist to tax compliance officer (both have mean feasibility of 1.0). Conversely, among the most feasible we find the transition from shop assistant to bartender (with mean feasibility of 4.63) or from human resources officer to human resources manager (4.83).

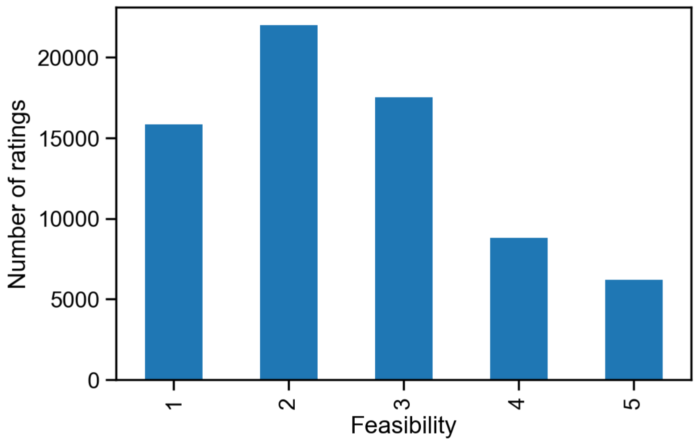

Figure 11. Histogram of all feasibility ratings (mean≈2.54)

Importantly, there are multiple feasibility ratings for each transition to be considered when estimating its overall validity. Figure 12 shows the relationship between the mean feasibility and the proportion of feasibility ratings above a certain threshold for each transition. Interestingly, there are many transitions with a low mean feasibility rating but at least some contributors giving a high rating. For example, two-thirds of the transitions with mean feasibility below 3 have 20% of their ratings greater or equal to 3. Alternatively, one-third of transitions with mean feasibility below 3 have 40% of their ratings greater or equal to 3.

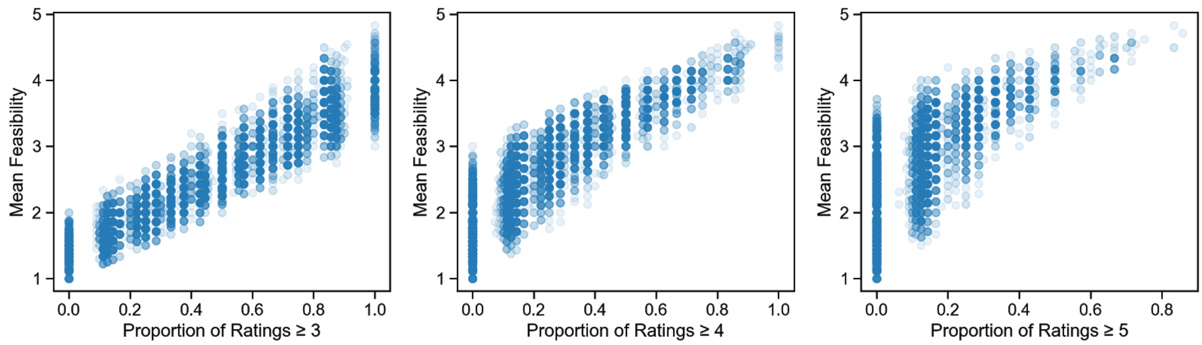

Figure 12. Relationship between the mean feasibility and the proportion of feasibility ratings greater than or equal to 3 (left), 4 (middle) or 5 (right).

The large variation across the ratings indicates a considerable disagreement among the contributors for some of the transitions. This raises the question of whether a transition should be considered feasible when it has a high average feasibility score, or if at least a certain proportion of respondents consider it highly feasible. Future investigations should also explore the nature of transitions that generate the largest disagreement between contributors.[4]

Correspondence between feasibility ratings and occupation similarities

As a first test of the validity of our career transition algorithm, we compared the estimates of occupation similarity derived by our algorithm with the crowdsourced feasibility ratings. Occupation similarity was measured by comparing the profiles of origin and destination occupations across four different facets: essential skills, optional skills, work activity types and work context.[5] These comparisons yield, for each transition, four similarity measures that we then combined using an equally weighted average.

Figure 13 shows that there is a good correspondence between the combined occupation similarity and crowdsourced feasibility. As a simple correspondence metric, we find the correlation coefficient around 0.70. This can be interpreted as the combined similarity measure explaining about half of the variance in feasibility. The variance of feasibility ratings appears largest around the similarity value of 0.30, which is to be expected as this value coincides with a previously calibrated similarity threshold for 'viable' transitions.[6] Moreover, the sampling of transitions for the

[4]Among the most contentious examples (with the highest variance across their ratings) were transitions from sales assistant to locker room attendant, from maternity support worker to radiographer, and from orthoptist to occupational therapist.

[5] All similarity measures are in the range between 0 and 1, with 0 indicating minimal similarity and 1 indicating maximal similarity. For more details on the specific methodology for comparing occupations, see the report [Mapping Career Causeways: Supporting workers at risk](#).

[6] We calibrated a data-driven viability threshold by making the assumption that transitions between ESCO occupations that belong to the same ISCO four-digit group should typically be viable. This yielded a combined similarity threshold of 0.30; in addition to this, we also ensured that the expected levels of education and experience in both occupations are comparable (see the report for details).

crowdsourcing exercise was also biased towards testing transitions close to this threshold.

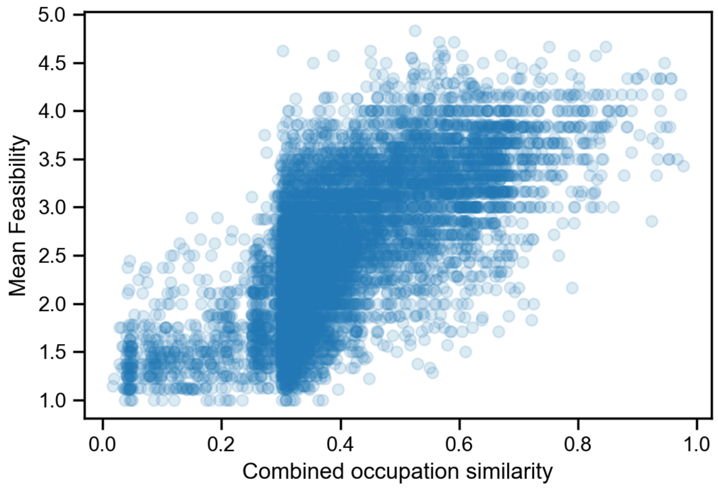

Figure 13. Correspondence between the combined occupation similarity measure and crowdsourced feasibility ratings (each circle is a transition); Spearman correlation coefficient = 0.71, Pearson correlation coefficient = 0.70 (p-values = 0).

While all four similarity measures that build up the combined measure were positively correlated with feasibility ratings, essential and optional skills similarity exhibited the strongest correlations (Figure 14). This is unsurprising as the contributors were shown job descriptions and skills, and they were not explicitly stimulated to consider similarities across other occupational aspects. However, the rather weak correlations for the coarser similarity measures for work activities and work context does raise the question about how important are these features when jobseekers are considering career moves.[7]

[7] The initial motivation for including the coarser similarity measures based on activities and work context was for the algorithm to capture broad similarities across different jobs and facilitate the broadening of workers' career horizons.

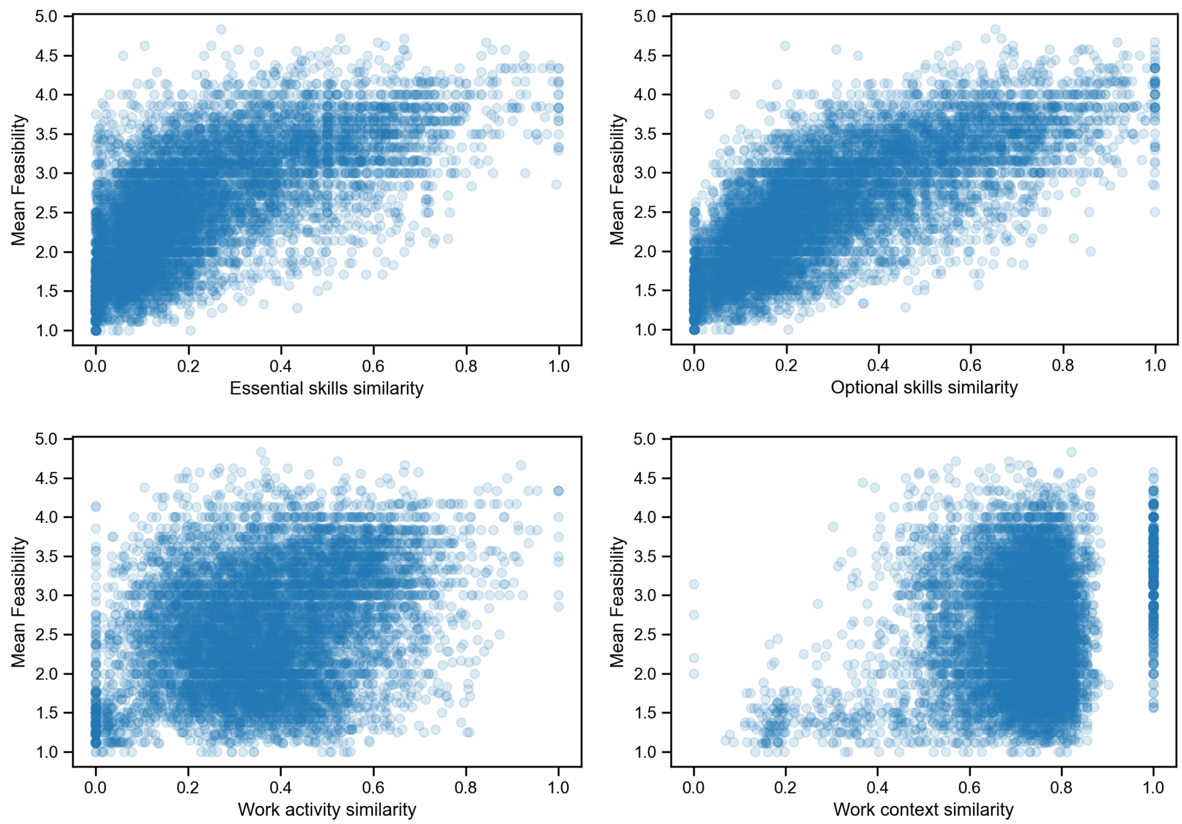

Figure 14. Correspondence between the four different similarity measures and crowdsourced feasibility ratings (each circle is a transition). All Spearman correlation coefficients are positive and statistically significant: 0.70 (essential skills), 0.77 (optional skills), 0.30 (work activities) and 0.14 (work context); p-values≈0.

Taken together, the positive, significant correlations between occupation similarity measures and feasibility are encouraging and support the overall validity of our approach. Importantly, the main observations are also robust with respect to the minority of core contributors who contributed most of the ratings (see the section Robustness tests in the Appendix). However, at the same time, a number of transitions that would be considered viable by our algorithm (with the combined similarity above 0.30) are rated with low feasibility. In the next section, we combine our occupation profiles, similarity measures and the crowdsourced feasibility ratings to predict the feasibility ratings for all possible transitions between any two ESCO occupations, and derive a more refined transition recommendation threshold.

Predictive modelling of feasibility ratings

To extend the crowd feasibility ratings to all recommended transitions, a predictive model was trained on the mean feasibility ratings awarded by the crowd contributors. Several different features were engineered and tested in the construction of the model, however ultimately it was found that the most predictive features were largely those which represented the information available in the crowd survey.

These features included the work context similarity and NLP adjusted skills overlap exactly as they are used in the transition algorithm. In addition, we created features that described the distribution skill overlap and the similarity of the job descriptions. The first set was created by calculating the cosine similarities for the matching skills between the two occupations, and calculating the 10th and 90th percentiles and the mean values. If no match was found for a skill in the destination occupation, zero was used. For the final feature, the cosine similarity between embeddings of the ESCO occupation descriptions was calculated. A Sentence-BERT[8] model was used to create the embeddings. Other features, such as the difference in job zone between the transitions and annual earnings ratios were trialled, but were found to add no predictive power. Figure 15 shows correspondence between the features and the mean feasibility rating.

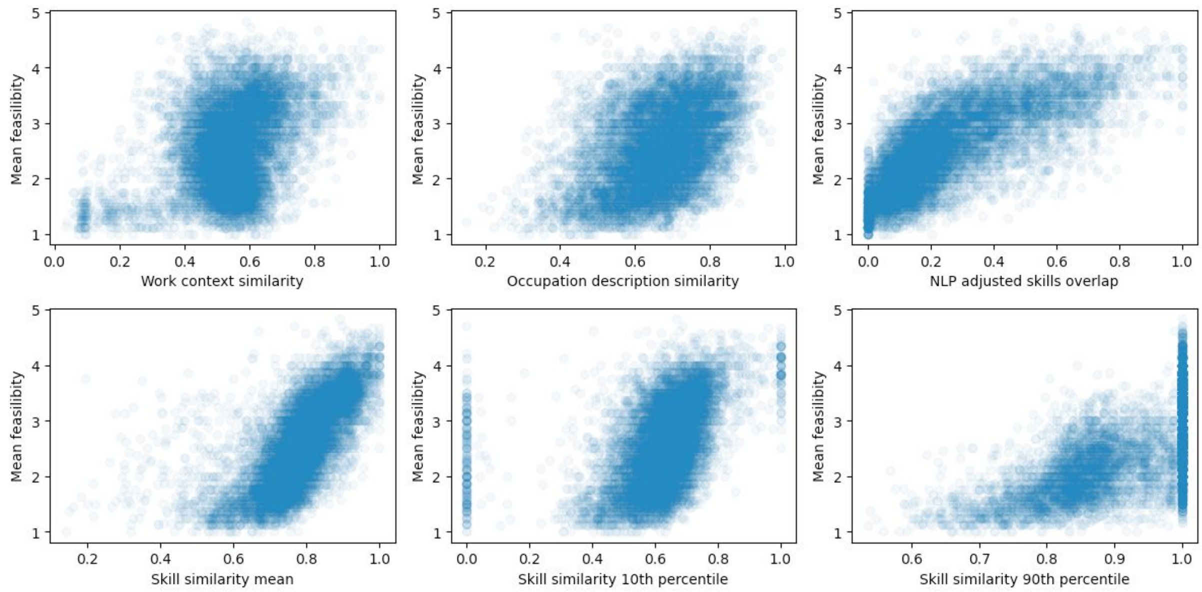

Figure 15. Correspondence between the model features and the mean crowdsourced feasibility rating.

The model type used was a support vector regression with a radial basis function kernel. The crowd labelled data was split into train and test sets with an 80:20 ratio. Cross validation for hyperparameter optimisation was carried out using a randomised search with and 5-fold splitting. The resultant model had an explained variance of R^2 = 0.64 and a mean squared error of 0.21 on the test set (74% of predictions on the test set were within 0.5 feasibility points of the true value).

[8] Reimers & Gurevych (2019) 'Sentence-BERT: Sentence Embeddings Using Siamese BERT-Networks'; we used 'bert-basenli- mean-tokens' model, which is part of the sentence-transformers Python package.

Identifying minimum viable transitions

For real world settings, such as digital careers services, an algorithm such as the one presented here needs to generate sufficient trust to be useful to those that it is trying to serve. In this case, trust may be reduced if a user is exposed to too many transitions that do not appear to meet a certain threshold of feasibility or reasonableness. We therefore attempted to flag and remove such transitions by establishing a minimum required feasibility threshold based on the predicted values from the model described above.

While establishing any such threshold is a subjective matter, we triangulated between three methods to determine a value that would strike the balance between excluding transitions which would be judged to have a low feasibility, while retaining those that a sufficient number of people would think were reasonable. First, we looked for the mean predicted feasibility rating for the transitions which at least 50% of crowd contributors gave ratings of 3 or more. Second, the team that worked on this project hand labelled almost 400 transitions with a simple binary judgement on their feasibility. We then looked for the predicted feasibility value at which at least 50% of transitions judged by the team were deemed acceptable. Using both of these approaches, we settled on a minimum predicted feasibility threshold of greater than or equal to 2.5. This also corresponds well with the scale used for feasibility ratings in the crowdsourcing exercise; a rating of 2 is below the middle rating, so can be taken to mean that the contributor thinks that the transition is mostly infeasible. Therefore, transitions with a mean rating of 2.5 can be considered to be in a space of disagreement, but where a sufficient proportion of members of the public deem the transition to be acceptable.

Discussion

This crowdsourcing study has provided a reference dataset for the public perception of career transition feasibility, which we have used to further refine our career transition recommendation algorithm. To the best of our knowledge, this is the first open dataset of its kind, using the ESCO occupational framework, and we invite other researchers and practitioners to use and build upon it. The potential research value of this dataset is further increased by the crosswalk between ESCO and US Occupational Information Network (O*NET) frameworks that was also developed as part of the Mapping Career Causeways project.

It is important to note that the results from this study have been generated from the subjective opinions of a relatively small crowd of the UK public. As such, they cannot be treated as any kind of "gold standard" for the real-world feasibility of career transitions (as opposed to something like a randomised control trial looking at job placements in the real world).

We are only able to say that the judgements represent observers' perceptions of the transition feasibility. The survey design does not place the contributor in any specific role when they answer the questions, however the language of the survey implies that it is the worker's perspective that is important to us. In addition, it is unlikely that contributors with senior roles in an organisation who might be more concerned with an employer's perspective would be highly represented in the survey due to the piecework nature of the task. Even if employers were present in the contributors, it is unlikely that they would have broad familiarity with all of the occupations included and their industrial sectors.

While the majority of transitions are likely to receive a layperson's judgement, we have included a question that assesses the contributor's familiarity with each occupation. This can be used to select insights only from the presumably most knowledgeable participants.

It is also important to acknowledge that in the process of refining our career transition recommendation algorithm, we have framed transition "validity" in terms of the "common sense" judgement of the general public. This is acceptable for achieving our present goal of filtering out transition recommendations that could be considered clearly impractical by the general population. However, validity of the algorithm's suggestions could also be alternatively framed in relation to real-world career transitions undertaken by workers (for which granular datasets are scarce and proprietary), or in terms of whether workers are encouraged to broaden their career horizons and make transitions that they would otherwise not have considered.

One final challenge with interpreting the results is that in theory, it is highly unlikely that any transition is completely infeasible. Given enough time, resources and support most workers could transition from one occupation to another. While we do ask contributors to give reasons for their feasibility ratings, we do not specifically ask them what constraints they are applying to their judgement.

Appendix

Survey design

Introduction to the Task

We have created a tool which can recommend possible new careers to people, based on the job that they are currently in. We call these ‘career transitions’. This is particularly important for people whose jobs might be at risk from issues such as COVID-19 or automation.

There are thousands of possible transitions that the tool could suggest, and we need to test whether they are possible and appropriate. In each question you will be presented with two jobs (A and B) and will be asked to judge whether someone could feasibly transition from job A to job B. To make the decision you will be provided with:

- The title and description of job A
- The title and description of job B
- Skills that are shared between job A and job B
- Skills that are similar between job A and job B
- Skills gaps - i.e. the skills that are required by job B, but not found in job A[9]

We would like you to use these pieces of information, along with your own knowledge and experience, to make a decision on whether the transition is feasible within 6 months to 1 year. If you believe that there are any factors that mean the transition might take longer than 1 year, then you should flag the transition as unfeasible.

If you flag a job as having low feasibility, we will ask you why you think this is. Reasons we have included are groups into categories:

- Skills - reasons relating to the skills required in job B compared to job A, such as poor skills overlap, different required skills or essential missing skills.
- Job market - in some cases the recommendation might suggest a very niche industry with few jobs or one that is declining, or job B might be much more highly competitive than job A.
- Personal - these are reasons relating to aspects of job B that are different to job A and might affect or be affected by personal circumstances. For example, job B might require a change in working hours or needing to work away from home compared to job A.
- Training - in some cases the lack of overlap in skills could be tackled through training or education. However, in some cases this might not be easy

[9] At maximum, 10 skills gaps (with the lowest skills similarity) and 10 shared and similar skills (with the highest skills similarity) were shown for each transition.

because of the time or cost required for training, or because training is not available for this job.
- Conditions - it might be undesirable to move to a job with lower compensation or to a job that involves more risky activities.
- Other - all other reasons

Pre-Qualifying Questions

These questions are here to help us to understand who is verifying the occupation transitions. It will also be used to help inform further research to make our recommendation tool more effective.

- Year of birth (dropdown select)
- Gender (multiple-choice)
    - Female
    - Male
    - Prefer to self-describe
    - Prefer not to say
- Any sectors you have worked in: (multiple-choice, multiple-select)
    - Animal care and husbandry
    - Armed forces
    - Arts and media
    - Business and administration
    - Construction, transport and security
    - Education
    - Engineering and research
    - Food, drink and tobacco
    - Healthcare
    - ICT
    - Logistics
    - Manufacturing
    - Charity and non-profit
    - Public services and government
    - Sales and services
    - Social work
    - Textiles
    - Other
- Highest level of qualification (or equivalent): (multiple-choice)
    - GCSE
    - A-Level
    - Higher National Certificate
    - Higher National Diploma
    - Bachelor's degree
    - Master's degree
    - Doctorate
- Region (multiple-choice)

- England
- Scotland
- Wales
- Northern Ireland

Transition Questions

For each transition the contributors were presented with the following questions:

- How familiar are you with job A?
  - Range from 1 (not at all familiar) to 5 (very familiar)
- How familiar are you with job B?
  - Range from 1 (not at all familiar) to 5 (very familiar)
- How feasible do you think it would be for someone to transition from job A to job B, based on the description, skills, and your own knowledge?
  - Range from 1 (not at all feasible) to 5 (very feasibly)
- What are your reasons for this choice?
  - Skills (e.g. poor skills overlap, missing essential skills, …)
  - Job market (e.g. industry of job B in decline, job B too competitive, ...)
  - Personal (e.g. big change in working hours, working away from home, undesirable transition, …)
  - Training (e.g. training would take too long, training not available, …)
  - Conditions (e.g. compensation of job B too low, perceived lower quality of work in job B, ...)
  - Other

Robustness tests

We tested that the observed positive association between the feasibility ratings and occupation similarity measures is robust with respect to the small minority of the most active contributors. In Figure 16 we specifically highlight the eight most active core contributors who provided over half of all the ratings.

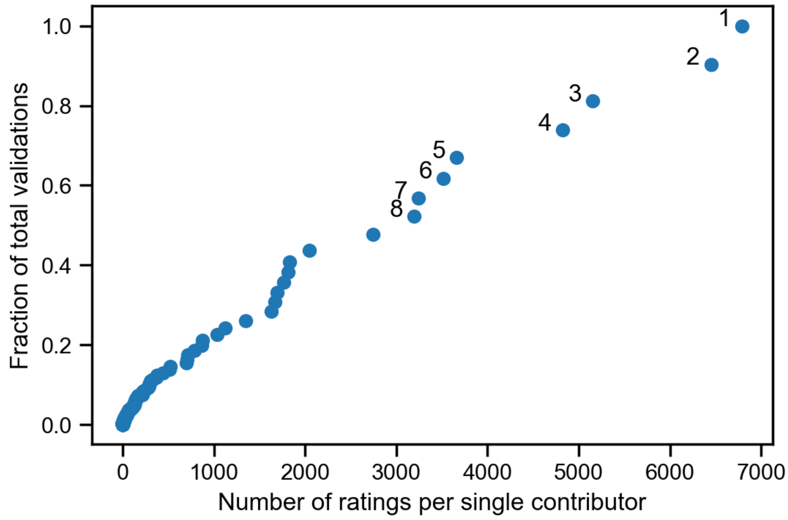

Figure 16. Cumulative distribution of the number of ratings per contributor (each circle is one contributor). Integer labels indicate the top eight most active contributors that together contributed more than 50% of the ratings.

We removed an increasing number of top *n* (with *n* between 0 and 8) most active contributors from the data, and recalculated the correlation coefficients between the feasibility ratings and our five occupation similarity measures (Table 1). Importantly, the correlations remain positive and significant as an increasingly larger number of core contributors is removed, and the mean transition feasibility across all transitions remains practically unchanged. Interestingly, the correlation strengths are attenuated with the removal of the core contributors, which suggests that their ratings were more consistent with the similarity measures compared to the rest of the population. This raises the question of whether the core contributors paid more attention to the provided information about jobs, or whether perhaps they "mastered" the task and adapted their ratings over the duration of the study.

Table 1. Spearman correlations (denoted by *ρ*) between the crowdsourced mean feasibility ratings and occupation similarity measures, when an increasingly larger number *n* of the most active 'super-users' are removed from the dataset. All correlations are statistically significant (p-value ≈ 0).

n, Number of most active users removed	Number of remaining ratings	Number of transitions with ratings	Mean feasibility across all transitions	*ρ* (combined similarity)	*ρ* (essential skills)	*ρ* (optional skills)	*ρ* (work activities)	*ρ* (work context)
0	70359	9113	2.56	0.71	0.70	0.77	0.30	0.14
1	63573	9113	2.50	0.62	0.61	0.69	0.24	0.14
2	57122	9113	2.57	0.60	0.60	0.67	0.23	0.14
3	51966	9113	2.55	0.58	0.58	0.65	0.22	0.14
4	47140	9113	2.54	0.58	0.57	0.65	0.21	0.14
5	43480	9113	2.58	0.57	0.56	0.63	0.21	0.14
6	39964	9113	2.60	0.57	0.56	0.63	0.21	0.14
7	36725	9112	2.60	0.51	0.50	0.57	0.18	0.13
8	33528	9080	2.59	0.48	0.48	0.54	0.16	0.12

Crowdsourcing dataset schema

Column	Description	Type
classification_id	Unique integer identifier for each feasibility rating	integer
subject_ids	Integer identifier for each unique transition	integer
origin_id	Unique integer identifier of the origin occupation	integer
origin_label	Preferred label of the origin occupation.	string
origin_description	Description of the origin occupation.	string
destination_id	Unique integer identifier of the destination occupation	integer
destination_label	Preferred label of the destination occupation.	string
destination_description	Description of the destination occupation.	string
perfectly_matched_skills	Semicolon separated list of skills that are shared by both origin and destination occupations	string
partially_matched_skills	Semicolon separated list of skills that are similar between both occupations (more specifically, destination occupation's skills that can be partially matched with skills from the origin occupation's skills set)	string
unmatched_skills	Semicolon separated list of skills gaps (more specifically, destination occupation's skills that cannot be matched with the origin occupation's skills set)	string
familiarity_with_origin_1-5	Contributor's familiarity with the origin occupation, on an integer scale from 1 to 5 (0=not at all familiar, 5=very familiar)	int
familiarity_with_destination_1-5	Contributor's familiarity with the destination occupation, on an integer scale from 1 to 5 (0=not at all familiar, 5=very familiar)	int
feasibility_1-5	Contributor's feasibility rating for the transition, on an integer scale from 1 to 5 (0=not at all feasible, 1=very feasible)	int
Reasons	Reason for choosing the particular feasibility rating (see the section on Survey design in the Appendix)	string
coder_id	Unique identifier of the contributor	string
questionable_respondant_flag	True if the contributor's behaviour differed substantially from the rest of the population	boolean
gender	Gender of the contributor	string
background	Sectors where the contributor has worked in	list of string
highest_qualification	Contributor's highest level of qualification	string

time_stamp	Time when the transition feasibility rating was made	datetime
age_group	Age group of the contributor	string