Survey Methodology

AppSec Practitioner's Landscape 2021 [Data Public Release]

About this Document

This is the documentation of the context, history, thought process, and rationale behind the survey design for public reference.

Release notes

2024-08-25 - Public Release of Data

We're releasing the RAW survey data collected via the Google Forms in 2021 to the community to enable the community to make use of the data we've collected; and hopefully do something useful with it before it becomes too stale. It's been too long.

We've included this document, to give you an idea of what we had in mind when we created the survey.

License

These are:

1. AppSec Practitioner's Landscape 2021 - Survey Methodology [Public Release][CC-BY].pdf (this document)
2. AppSec Practitioner's Landscape - Survey Form Archive [PublicRelease].pdf
3. AppSec Practitioner's Landscape RAW Responses [PublicRelease][CC-BY].xlsx

You may reference this body of work collectively as
`AppSec Practitioner's Landscape Survey 2021,`
`by the OWASP How to Get into AppSec Project`

The key contributors to and leaders of the How to Get into AppSec project are
`Jenn Janesko, Didar Gelici, and Daniel Ting`

Prepared by Daniel Ting and reviewed by Adrian Winckles - August 2024

Methodology

Goals

To establish necessary learning outcomes to be successful in Application Security, we must first know what is expected of an Application Security professional. This project aims to establish that, and maintain the current expectations of the profession as the occupation evolves.

Guiding Approach

The project is intent to inform what the occupation of an Application Security Professional is. To ensure easier future work and adoption into existing governance frameworks established for other occupations, this work attempts to align as closely as possible with existing Occupation Standards definition approach; including data collection and processing.

Ideally, this project will attempt to align the output guided by well-established approaches such as those that underpin, Professional Standards Council & ANZSCO Occupation Standard, NIST NICE, USA's Standard Occupational Classification (SOC), EU's ESCO (European Skills, Competences, Qualifications and Occupations), the International Standard Classification of Occupations (ISCO) and others.

Data Collection

Medium:
Survey, at most 10 questions with constraints to ensure completion <5 minutes.

Respondents/Channels:
OWASP Social Media channels (mainly Twitter), and OWASP membership email list.

Survey Design

Overview

Shaped by the goals mentioned above, led by other industries that has done this, as the base for this survey, we used the published methodology by the Australian Bureau of Statistics (ABS) approach as the inspiration for collecting these statistics. Further context were then influenced by publications from related organisations attempting to and have prior experiences establishing what an occupation does; such as those listed above in the "Guiding Approach" section.

Guided by the goal of publishing expectation guidance for jobs such as template job descriptions, and key capability indicators; which will guide the development of open curriculum; organisational environmental questions were included. As the Project committee is well-aware that the risk profile is contextual, the hypothesis that the job expectations varies based on the organisational context which will need to be accounted for.

Broadly, we capture 3 categories of information

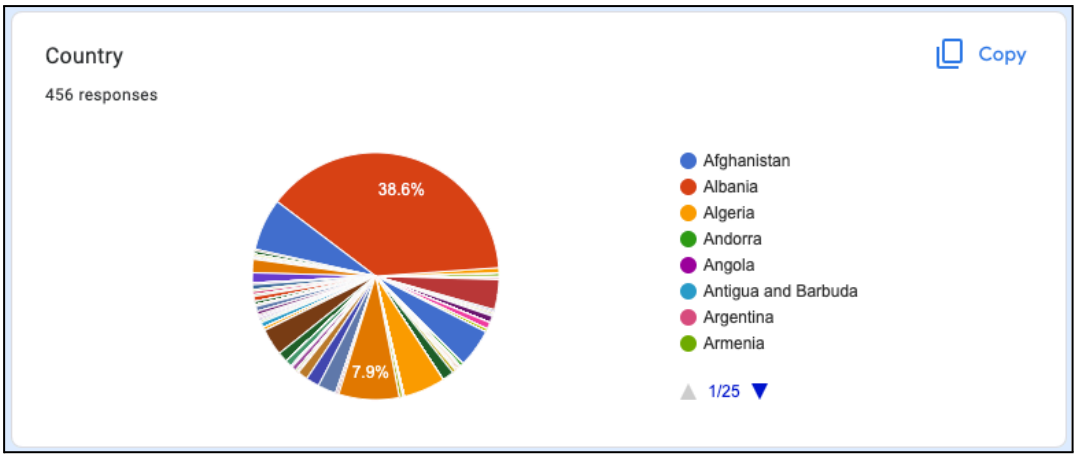
a) Environmental questions – demography and organisational context
b) Organisational maturity questions – AppSec practice maturity
c) Occupational questions – Individual's experience and activities

As this survey forms a pilot study to inform of the environmental landscape of the profession, to maximise response and minimise excise (burden on respondents), this survey is constrained to questions that are minimally necessary to help provide a guidance for further future study.

With understanding, and expected that the data collected may be used for secondary investigations and analysis by the public; we also consider minimisation of any personally identifiable information in this survey. The questionnaire and rationale behind these are below.

Environmental Questions

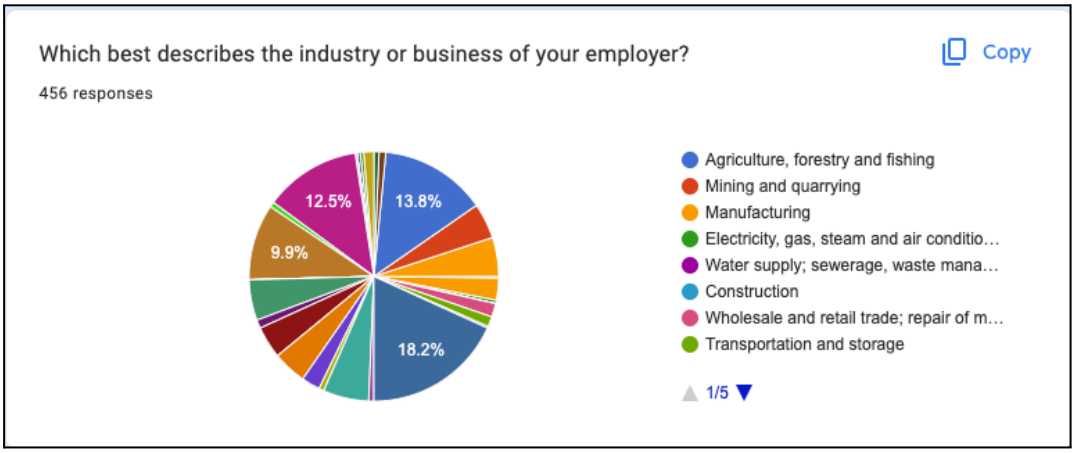
1. **Title**: Country
   Prompt: The main country of which the organisation you work for is based in



   Thought Process:
   Affordances, culture, and expectations vary based on locale. Capturing this demography so that the data can be normalised, and bias informed.

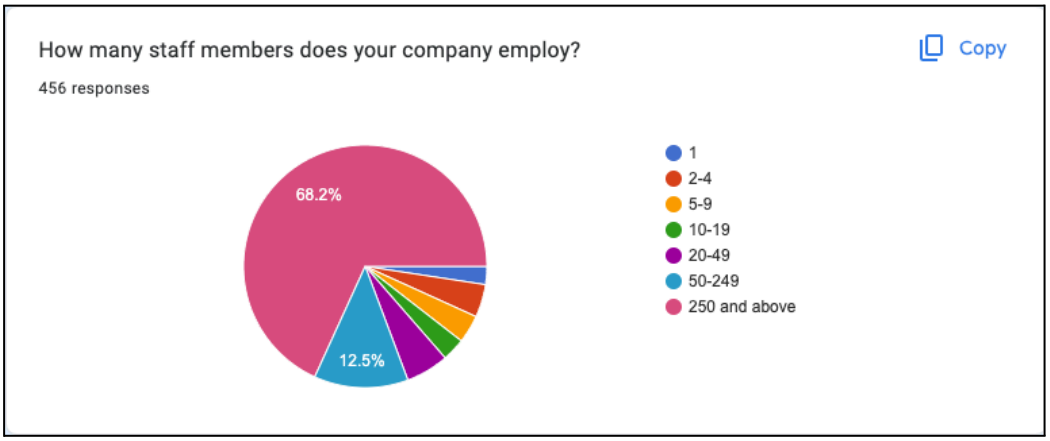
2. **Title**:
   Which best describes the industry or business of your employer?

   Prompt:
   Select the best that apply. If you're interested in the formal definitions of these categories they are defined at https://unstats.un.org/unsd/publication/seriesM/seriesm_4rev4e.pdf



   Thought Process:
   We selected the United Nations definitions so that it is more universally applicable, and with the hope that mappings exist between these industry definitions, and local definitions. We additionally made the decision to only use Tier 1 classifications, with the exception for the category of “Information and communications” because we assume that due to the nature of Application Security being heavily represented by this category, it would be worthwhile to obtain additional granularity. As such, for the “Information and communications” classification, we included options for its subcategories.

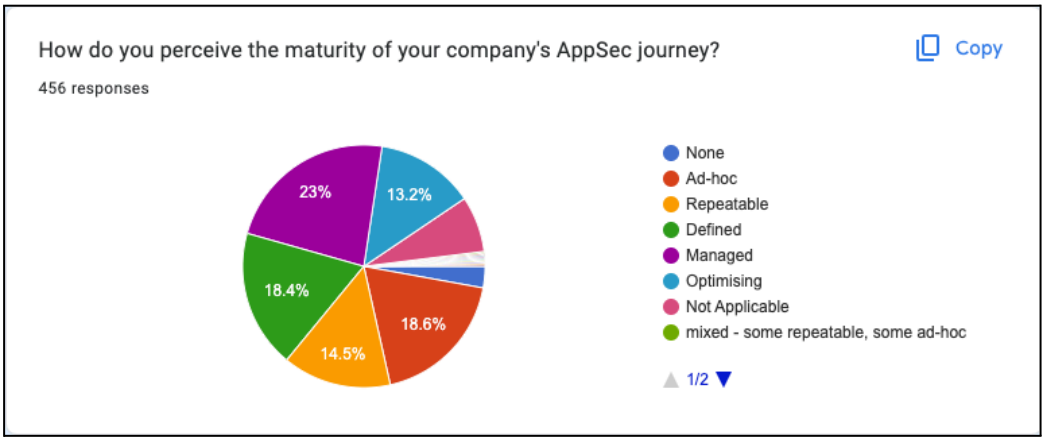
3. **Title**: How many staff members does your company employ?
   Prompt: None.



   Thought Process:
   As the roles distribution and needs of Small Business, and Large Enterprises differ, we reason that the size of the organisation would be important to know as it affects the roles and deliverables of individuals. The hypothesis is that smaller organisations will likely have a single person doing more tasks, and larger organisations performing more specialised tasks. The size ranges are [a mesh-up of various global definitions of organisation] sizes to be usable by as many as possible globally.

Organisational Maturity Questions

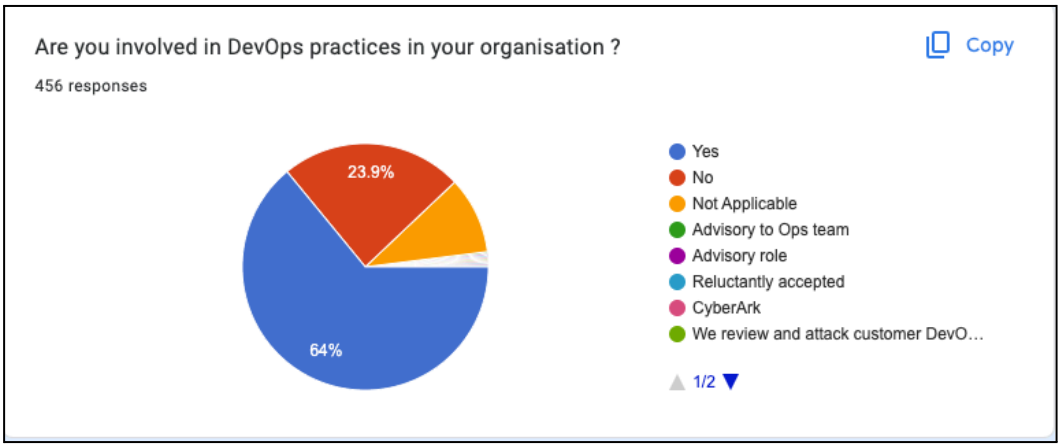
4. **Title**: How do you perceive the maturity of your company's AppSec journey?
   Prompt:
   Select "Not Applicable" if this does not apply to your type of organisation. Eg. An advisory consultancy. For guidance on what these mean, see https://en.wikipedia.org/wiki/Capability_Maturity_Model#Levels



   Thought Process: An objective assessment of maturity is challenging, as it would require the respondents to perform an assessment of its maturity against a standard. The process of doing so will imply some maturity. In addition to that, performing and verifying these assessments against any known standards (including SAMM) would be challenging and beyond our time budget for respondents to complete this survey.

As such, we decided to base this on the respondent's perceived maturity of their organisation, with full understanding of the perception bias that might have. We simply picked a well-known capability maturity model framework that is commonly used by Cyber Security frameworks and other capability maturity assessments.

5. **Title**: Are you involved in DevOps practices in your organisation ?
   Prompt: Select "Not Applicable" if this does not apply to your type of organisation. Eg. An advisory consultancy.

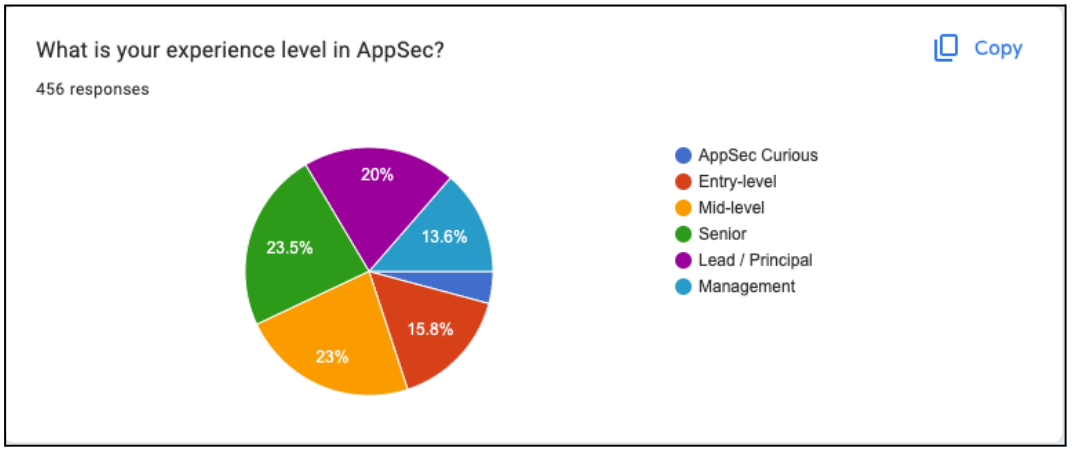


Thought Process:
This was included to identify if DevOps practices are linked to AppSec practices, or an indicator of maturity. There has been many discussions that AppSec and DevOps significantly overlap, and that building DevSecOps pipelines are what AppSec professionals do. This question hopes to identify these and additional questions for future studies around DevOps and AppSec.

Occupational questions

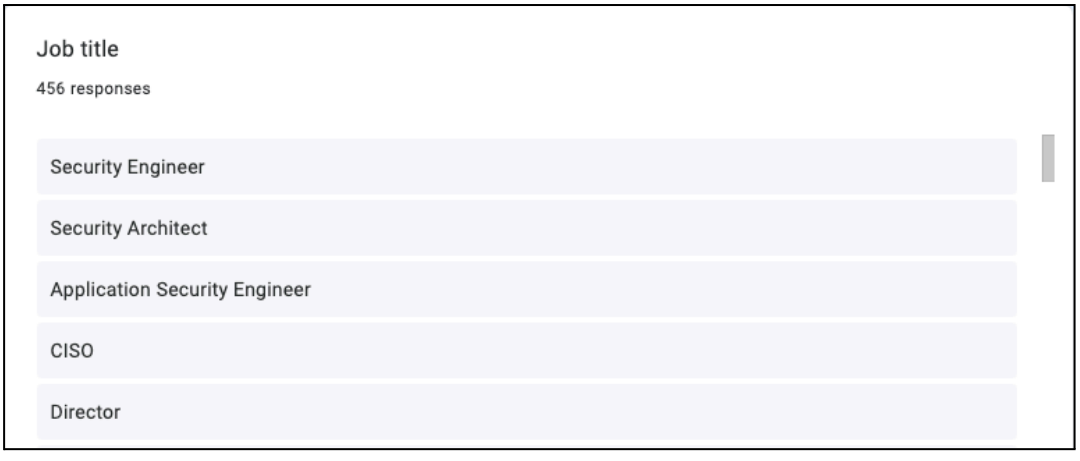
These questions are heavily influenced by Professional Standards frameworks such as SFIA, Professional Standards Council and Occupational Standards definitions. Overall, these questions are intended so that we can use these collected survey data to explore certifications that are congruent with international standards and expectations necessary for practical adoption.

6. **Title**: What is your experience level in AppSec?
   Prompt: None.



   Thought Process: Roles at different seniority levels and experience perform different tasks, and have different outcomes. This is to help identify and segment respondent's based on their experience level in AppSec for the career/profession analysis.

7. **Title**: Job Title
   Prompt:
   If you have multiple job titles, what's your primary job title.

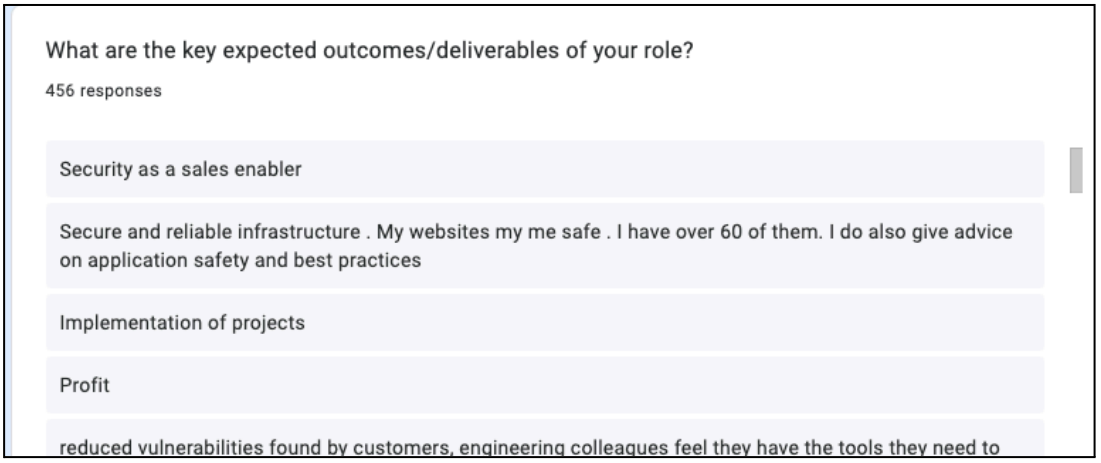


   Thought Process:
   This is a free text area for the respondent to provide their Job Title. This is helpful in understanding the titles that are associated with the work performed, and capabilities presented.

8. **Title**: What 3 key activities do you perform as part of your role?
   Prompt:
   Think about your day-to-day work, describe what are the top 3 activities that you spend majority of your time performing

   Thought Process:
   What does the respondent actually do? These are to capture the main tasks the respondents perform as part of their work. We limit this to 3 main tasks, so that it's an intentional constraint on the respondent to reflect and prioritise what they actually do. Irrespective if they are AppSec related. This is intentional as to understand if "AppSec Professionals" main jobs are elsewhere and "AppSec" is simply part of another role in practice. This helps us discover what a self-declared "AppSec Practitioner" actually does.

9. **Title**: What are the key expected outcomes/deliverables of your role?
   Prompt:
   Describe what key indicators or results are checked to see that you are meeting expectations set for your role.



   Thought Process:
   How does the employer judge success in the role? The tasks performed, and the KPIs of a person may not always align. This question is used to determine how the person's performance will be measured by their employer. This may be useful to understand for example, what the person does versus what their employer expects or uses as performance indicators.

Data Preparation

> 🚨 **NOTE**
> This section is included in the release for historical posterity, and to help provide context. These are our initial discussion draft of our data preparation methodology and may not be adequate for prioriate analysis rigor.

Before meaningful analysis can be performed, we must first prepare the data for analysis. Additionally, as the questionnaire contained qualitative free-text responses, these needs to be qualified and normalised.

From a high-level, the data preparation, extraction, transformation and loading (ETL) process, will be conducted as follow:

a) Cleaning & Distillation – removal of invalid survey entries and distilling text submissions.
b) Transformation & Normalisation – of the text inputs (using frequency analysis).
c) Categorising & Mapping – Clustering/categorising job titles, activities, and outcomes; via mapping job titles to activities and outcomes.

These steps should prepare the data for analysis, being able to map common “buckets” of activities and “KPIs” expected for job title clusters.

a) Cleaning & Distillation

The aim of this stage is to ensure that the data we’re working with is valid submissions. For each entry, we qualify and eliminate entries from future stages of the process should it not pass casual inspection for validity.

b) Transformation & Normalisation

Job Titles

1. An initial sweep is performed to “normalise” and clean up entries.
   This means, expanding words (eg. Snr → Senior), and capitalisation (we adopt Capitalised Words).
2. Where there are multiple spelling options for a role, we pick the current top "Worldwide" trend’s spelling based on Google Trends.
   An example of this is “Cyber Security” vs “Cybersecurity”, where the prior is selected.
3. Where there are multiple titles/roles submitted, the role perceived to be the best match to the key activities performed is selected.
4. Then, we perform a keyword/keyphrase frequency analysis on the phrase/job title, to select the most representative version of that title.
5. Where there is still ambiguity, we will rely on our industry experiences to make a decision. If a judgement call is made, it will be noted.

This should put the “Job Titles” data in a state for further analysis, and mapping to role activities.

3 Key Activities

1. An initial sweep to identify any entries that did not appropriately complete the task as requested. Examples include those that submitted a paragraph, or a single task, or more than 3 key activities. Any that may require too much logical leaps/imagination are removed from the data analysis.

2. The activities listed are broken down into columns/"bags of terms" associated with the submission/role.

3. Interpretation and summarisation will be performed by the analyst where necessary to convey the main activity performed. Especially in cases where responses have been particularly verbose.

4. **Mapping Synonyms**
   As a global industry we may use different terms to often describe functionally the same thing, and is generally understood by practitioners to be interchangeable. As an example, grouping "coding", "development", "build", and "writing code" to be represented by a single term – "coding". "Managing People", "Management", "Team Management" are represented as "Management"; and "CI/CD", "DevOps", "Build Pipelines" are presented as "DevOps".

   In efforts to make the outcome meaningfully actionable, we need to generalise/simplify some of the data we've collected. This is the usual granularity trade-off, where we're attempting to get the right balance between resolution, and meaningful insights for our purposes.

   At this stage, the focus is on data preparation, not analysis. The goal is that the effort will be minimally invasive to the data. As such, this is the process we take for normalising the "Activities Performed" terms:

   1. An initial sweep is performed to "normalise" and clean up entries.
      This means, expanding words (eg. Snr → Senior), and capitalisation (we adopt Capitalised Words).

   2. Where there are multiple spelling options for an activity, we pick the current top "Worldwide" trend's spelling based on Google Trends.

      An example of this is "Cyber Security" vs "Cybersecurity", where the prior is selected.

   3. Where multiple similar activity terms are submitted, the activity perceived to be the best match to the key activities performed is selected by the analyst. Synonyms, and functionally similar activities will be mapped and represented as a single term.

      Example: "CI/CD", "DevOps", "Build Pipelines" are presented as "DevOps".

4. Where terms can be mapped to an OWASP SAMM activity, it will be. A judgement call will be made by the analyst on best fit. Consideration will be given to the context of the practitioner's submission.

5. Where terms/activities submitted do not map to an activity within OWASP SAMM; we perform a keyword/keyphrase frequency analysis on the "activity", to select the most representative version of that phrase/term. This is so that activities performed by practitioners but are not captured by OWASP SAMM can be discovered.

   Example: BizDev, Sales, and similar are not part of the SDLC.

6. Where there is still ambiguity, we will rely on our industry experiences to make a decision. If a judgement call is made, it will be noted.

As a final step once this has been completed, a spell check will be performed on the normalised terms to catch any mistakes.

The original to normalised terms are mapped in the Spreadsheet Tab "Activites->NormalisedActivityTerms".

For the survey responses to "What are the key expected outcomes/deliverables of your role?", as these are qualitative responses, these will be analysed and processed as part of the Data Analysis process.

References Used

- https://www3.weforum.org/docs/WEF_Skills_Taxonomy_2021.pdf
- https://www.reskillingrevolution2030.org/reskillingrevolution/initiatives/forum-led/skills-consortium/index.html
- https://www.weforum.org/topics/education/
- https://www3.weforum.org/docs/WEF_Skills_Taxonomy_2021.pdf
- https://niccs.cisa.gov/workforce-development/nice-framework
- https://www.abs.gov.au/statistics/standards/occupation-standard/2018
- https://www.psc.gov.au/professional-standards-schemes/why-apply
- https://www.psc.gov.au/legislation
- https://www.abs.gov.au/statistics/standards/occupation-standard/2018
- https://www.ilo.org/public/english/bureau/stat/isco/
- https://www.ilo.org/public/english/bureau/stat/isco/isco08/index.htm
- https://en.wikipedia.org/wiki/International_Standard_Classification_of_Occupations
- https://www.weforum.org/reports/building-a-common-language-for-skills-at-work-a-global-taxonomy
- https://www.bls.gov/soc/2018/soc_2018_manual.pdf - See FAQs
- https://www.bls.gov/soc/2018/soc_2018_class_prin_cod_guide.pdf
- https://unevoc.unesco.org/e-forum/A_Framework_for_Defining_Training_Standards.pdf