



**PSET CLOUD**  
Innovation through collaboration

# CALL FOR PROPOSALS

The development of a foundational taxonomy, coding schema and opportunities matching framework for MVP2 of the PSET CLOUD

July 2022 - February 2023



**JET EDUCATION SERVICES**  
**THINK EDUCATION. THINK JET.**



**merSETA**  
MANUFACTURING, ENGINEERING  
AND RELATED SERVICES SETA

6 Blackwood Ave, Parktown, Johannesburg, 2193  
Tel: +27 11 403 6401 | Fax: +27 86 550 0115 | [info@psetcloud.org.za](mailto:info@psetcloud.org.za)  
[www.psetcloud.org.za](http://www.psetcloud.org.za)

# 1. Background

JET Education Services (JET), in collaboration with the Manufacturing, Engineering and Related Services Sector Education and Training Authority (merSETA), has initiated the PSET CLOUD (Post School Education and Training Collaboration and Learning Opportunities and Utilisation of Data), a programme that seeks to develop an integrated national digital ecosystem which is interoperable, can be used for effective skills planning and provisioning, and puts information in the hands of citizens and other stakeholders so they can make informed decisions about education and training.

The purpose of the programme is to enable the government and citizens to make informed decisions related to education, training and work. The programme seeks to ensure that data sets are interoperable, well synchronised and used effectively as sources of information for planning and improving efficiency in the PSET system as well as for individual decisions. In pursuit of this goal, the programme will establish a digital ecosystem that will strengthen, integrate, coordinate, improve efficiencies and solve challenges in the governance and management of the post-school education and training (PSET) system. The PSET CLOUD programme has different work streams including governance, partnership, advocacy, monitoring and evaluation and the development of an interactive platform, which all support the achievement of the PSET CLOUD goals. With this call for proposals, one more workstream is added to the PSET programme.

The terms of reference (ToR) outlined below pertain to the development of a foundational taxonomy, coding schema and opportunities matching framework for the PSET CLOUD, which will underpin the underlying logic of the MVP and form the basis for the technology stack used in the development of the PSET CLOUD.

Phase 1 of the programme has been completed and involved a situational analysis of the PSET sector, a mapping study, an international review of similar initiatives and a feasibility report. These research reports have been condensed and included in a publication titled [Unlocking the Power of Data: A review of the state of readiness of the Post-School Education and Training sector in South Africa for enhanced data interoperability](#), released in November 2020 and available on the JET website, [www.jet.org.za](http://www.jet.org.za). The international review, [Interoperable Data Ecosystems: An international review to inform a South African innovation](#) is also available to download.

Phase 2 of the programme commenced in 2020 with two focus areas:

- [Stakeholder engagement and scenario planning](#);
- Development of a business case for the PSET CLOUD platform for piloting/testing as the programme transitions into Phase 3 (report available upon request). ; and

Phase 3 of the programme commenced in 2021 with four focus areas:

- Development of a minimum viable product (MVP) and Self-Sovereign Identity (SSI) Solution for the PSET CLOUD;
- Development of a branding, communication and advocacy strategy as well as a website to update stakeholders on an ongoing basis as the platform is developed;
- Design of a suitable governance model for the PSET CLOUD;
- Ready the merSETA to be the early adopter of the PSET CLOUD.

The outcomes of the work specified in this ToR will inform the work to be carried out in the development of MVP2. Thus close alignment with the MVP2 development team will be needed in order to successfully implement the outcomes specified.

At a minimum, MVP2 is envisaged to connect key stakeholders of the PSET ecosystem, that is, learners, employers and education and training institutions. The emphasis is on improving data interoperability and networks to improve system performance and efficiencies, particularly in coordinating links between learners, education and work opportunities and related functions. This may, for example, require forging new connections between different available tools and data in the PSET sector, integrating innovative technologies to address systemic challenges and addressing 'pain points' etc. in the system.

Realising a growing trend in the usage of foundational taxonomies, particularly for data driven ecosystems, the PSET CLOUD team undertook an exploratory analysis of what exists both locally and internationally, and how this could be incorporated into the development of foundational taxonomy and coding schema for the PSET CLOUD. The exploratory analysis was premised on the value of mapping, categorising and standardising a skills and occupations language so that users of the platform are easily able to identify/navigate which cluster of qualifications are linked to specific occupations and where skills are matched to jobs (through AI).

A skills and occupations taxonomy, very simply, is the science of naming, describing and organising skills and occupations into a database, and how this is done varies from method to method; however, the end goal always remains the same: to formulate an easy to comprehend infrastructure of skills and occupations (Van Echtelt, 2021). A well-curated skills taxonomy is a quick and effective way to ensure efficient matching of candidates to opportunities (Van Echtelt, 2021), and it will also help identify emerging or new skills as well as those which are being phased out.

We explored how existing skills taxonomies either arise from an expert-led process through a top down approach or are developed through a bottom up approach, where big data is aggregated from individual profiles. Data driven skills taxonomies aggregate unique skills mentioned in job adverts including specific knowledge, tasks and personal attributes. Machine learning is then applied to hierarchically cluster the skills. The taxonomy therefore captures the clusters of skills that are needed for existing jobs and their associated learning opportunities. Taxonomies that we explored included:

1. **US O\*NET-SOC Taxonomy** - A US database of occupational requirements and worker attributes. It describes occupations in terms of the skills and knowledge required, how the work is performed and typical work settings.
2. **ESCO** - European Skills, Competences, Qualifications and Occupations. The taxonomy identifies and categorises skills/competences, qualifications and occupations relevant for the EU labour market and education and training. The system provides occupational profiles showing the relationships between occupations, skills/competences and qualifications.
3. **UK Skills Taxonomy** - Aggregated from vacancy data, the taxonomy is a tree-like structure with three layers. The first layer contains 6 clusters of broad skills; these are split into 35 clusters, and these are in turn split to give 143 clusters of specific skills in the third layer.



4. **Emsi Open Skills Taxonomy** - An open-source taxonomy of 32,000+ skills gathered from hundreds of millions of online job postings, profiles, and resumes (Emsi, 2022).
5. **WEF Global Skills Taxonomy** - The interactive taxonomy shows clustering of skills at various levels. Granularity increases with each level. The end users (i.e. employers, governments, learning providers and individuals) are able to propose their own taxonomy at level 5, using level 4 to translate within the global taxonomy. Examples of potential level 5 skills are displayed for each level 4 skill.

What became clear is that all these skills and occupations taxonomies have some nuances in common, namely, that they rely on profiling skills, abilities and attributes in addition to job requirements. Additionally, as far as possible, there is use of atomic level data that is gathered through various methods, e.g. online scraping and national and industry data, which is then in turn used to define broader occupational groups. Subsequently, this allows mapping of skills, qualifications, experience and personal attributes to these occupational groups. Essentially, where a framework exists to map atomic-level data to aggregate occupational groups, it is possible to develop platforms that use a single unified taxonomy as the backbone of labour market matching and skills mapping capabilities.

In addition to this international review, a quick glance at what exists locally, led us towards the National Qualifications Framework (NQF) and the Organising Framework for Occupations (OFO). It became clear that the NQF together with its sub-frameworks uses a skills taxonomy that can enable mapping of:

- Job titles to occupations;
- Skills and skills levels to occupational categories;
- Qualification levels to occupations;
- Experience to skills level and job requirements;

Further review of what exists locally led to the realisation that, in fact, the Department of Higher Education and Training, through the National Career Advice Portal (NCAP) platform, had done much of the mapping between the NQF and the OFO. However, concern remains that a taxonomy solely based on the NQF and OFO remains static whereas the requirement for a data-driven platform such as the PSET CLOUD requires dynamism - particularly as it incorporates credentials earned through non-formal and informal education and training.

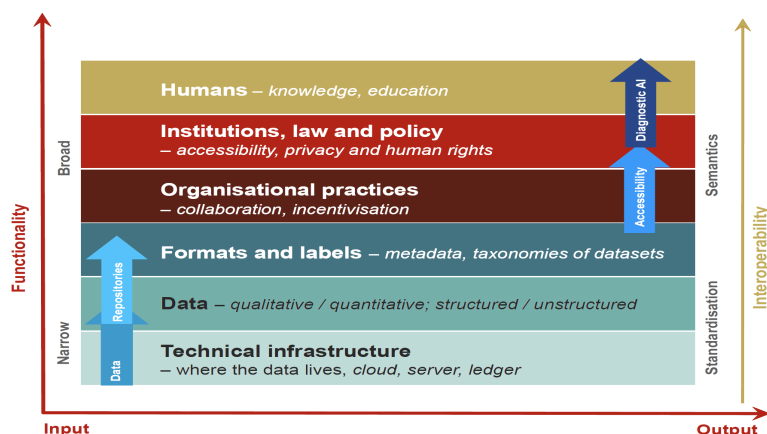
Summarily, the exploratory assessment made the following recommendations for consideration:

- *Option 1:* Adopt the NQF and OFO taxonomy in current format, thereby allowing for interoperability with existing systems across the local PSET landscape.
- *Option 2:* Adopt nascent international taxonomies developed through big data rather than policy experts and promote uptake locally. This requires enquiry into portability of the lingua franca between the local worlds of PSET and the labour market.

It is against this background that this call is being issued.

# 1. Scope of work

The scope of work is twofold. Component one looks at the establishment of a foundational taxonomy and coding schema for the PSET CLOUD system. Component two looks at the development of a



framework that will make use of the adopted taxonomy to establish the logic in the opportunities matching and skills mapping that will be carried out by the system through the use of artificial intelligence (AI). Looking at the Data Commons Interoperability framework, the framework of choice for the PSET

CLOUD programme, taxonomies are found in layer 3 from the bottom. This layer deals with formats and labels used to sort information as set out by metadata and taxonomies of datasets (O'Malley et al., 2018), and it falls under the narrow data commons which focuses on standardisation, which plays a key role in interoperable information systems.

In component one the service provider is expected to:

1. **Develop a foundational taxonomy for the PSET CLOUD.** To do this, the service provider will need to review existing skills taxonomies locally and internationally and explore how these could be merged with the skills taxonomy that is based on the NQF and OFO frameworks. Annexure A contains a table with a list of key local and international frameworks and taxonomies that are to be considered when carrying out this work.
2. The service provider will need to **harmonise** these taxonomies so that a common understanding of the different methodological approaches and vocabularies is created. This will require an extensive **mapping exercise** that will map different frameworks/taxonomies within the PSET space and the labour market. Furthermore, the service provider will need to consider the work being done by the World Economic Forum as it is in the process of establishing a global skills taxonomy.
3. Key to the development of the taxonomy is the inclusion of credentials earned through non-formal and informal education and training and micro-credentials as these sit outside of the NQF and are unaccounted for; however they are prevalent in the PSET space and the labour market. The provider thus needs to explore validation pathways outside of the NQF in

order to factor these credentials into the development of the foundational taxonomy for the PSET CLOUD. Currently the PSET CLOUD system allows users to self-attest skills that they have acquired through work they have completed informally. This is a powerful feature that aims to recognise skills acquired outside of a formal environment. The goal for the PSET CLOUD is to have these skills recognised through means of formal verified credentials from the appropriate professional bodies. Ideally what the system could do is the following:

- a. Accept self-attested skills from users that have been acquired informally.
  - b. In the event that the user wants to formalise their skills, the system could compare the set of skills listed by the user with the set of skills listed on the OFO for that particular occupation.
  - c. If the self-attested skills listed by the user match those listed on the OFO for that particular occupation, then the user needs to be issued an open badge.
  - d. If the user with an open badge requires their skills to be recognised formally by a professional body, they will then use the open badge as verification to confirm that their skills match those listed on the OFO, therefore qualifying for an assessment with a professional body within their industry.
  - e. The professional body will then assess the user and issue a formal verified credential.
  - f. In the event that the user does not have a complete list of skills that are listed on the OFO for their particular occupation, the system will recommend what courses to study in order to have the relevant skills as per the OFO.
4. Align with the development work carried out in MVP2 as this is crucial for the development and implementation of the foundational taxonomy for the PSET CLOUD.
  5. Lastly, produce a **coding schema** for the newly developed foundational taxonomy for the PSET CLOUD.

The success of the PSET CLOUD lies in the successful harmonisation of the various taxonomies within the different information systems and frameworks in the PSET ecosystem and labour market. It is therefore crucial that development of the PSET CLOUD is embedded within a common language that will allow users to search, discover and compare qualifications and credentials that meet their needs and also allow employers to find both the workers and education and training programmes that meet their needs. The PSET CLOUD should also allow everyone to understand how qualifications, credentials and careers connect to best prepare for the current and emerging workplace.

In component two the service provider is expected to:

1. Use the adopted taxonomy to create a framework that will underpin the logic of the job matching that will be carried by the system through the use of AI. The framework needs to enable a matching functionality that utilises both formal and informal education and training qualifications and credentials, together with skills, knowledge, competencies gleaned from



user profiles to ensure that there is a clear path that can be taken within the system by citizens with informal/non-formal education and training.

## Inception

---

The service provider will need to produce an inception report with a detailed work plan that spells out the approach, timelines, deliverables, risks and mitigation strategies, budget and key resources.

### Key deliverables

1. A detailed inception report.

## Component one - Foundational Taxonomy

---

An example of what is expected from the service provider in terms of developing a foundational taxonomy for the PSET CLOUD can be seen [here](#) & [here](#). This must be cross referenced with the work that is being carried out by the World Economic Forum around [global skills taxonomies](#).

### Key deliverable

1. A harmonised foundational taxonomy that has mapped key frameworks and taxonomies within the PSET space, such as the NQF, OFO, EMSI, with a key focus on non-formal education and training and micro-credentials.
2. A high level outline of the coding schema i.e database structure for the foundational taxonomy.

## Component two - Framework

---

The PSET CLOUD system will use AI and ML to match users with employment and learning opportunities. This matching needs to be based on a logic that encompasses the NQF and OFO but also has a key focus on non-formal education and training opportunities along with micro-credentials. The service provider will need to produce a framework that will serve as the underlying logic that will enable the AI to carry out its matching tasks. The framework will detail the data requirements, relationships, rules and procedures for occupational and qualification matching.

### Key deliverables

1. A framework that will underpin the matching of jobs and learning opportunities in the PSET CLOUD system.



## 2. Key competencies & team compositions

- Knowledge and understanding of the PSET landscape, in particular research relating to the codification of skills, knowledge, competencies and other relevant domains;
- In-depth understanding of international and local occupation classification frameworks like O\*NET, ESCO and OFO;
- Sound knowledge of applicable legislation such as the Skills Development Act, 1998 (Act No. 97 of 1998) and the National Qualifications Framework Act, 2008 (Act No. 67 of 2008) and other relevant legislation aimed at skills development, post schooling, training and development;
- Knowledge of mapping specific jobs and competencies to OFO codes and qualifications;
- Demonstrated skills and competence in educational and training research;
- A proven track record in conducting successful similar projects relating to occupational and skills taxonomies;
- A basic understanding of data requirements for PSET related technology platforms platforms.

## 3. Finance

The service provider's financial proposal should provide an explicit budget with a detailed breakdown by level of cost and must contain itemised costs for the following broader deliverables: an inception report and a harmonised foundational skills taxonomy, coding schema and PSET CLOUD matching framework. Payments will be made upon satisfactory completion and acceptance of deliverables by the JET/merSETA team. Work is considered complete when it has been signed off by the project manager. As a result, payments will be based on what has been signed off.

All costs associated with the development, preparation, production and/or delivery of goods and/or services incurred without an executed contract copy signed by all parties will be for the account of the bidding company or organisation. Neither JET nor the merSETA will pay for any costs associated with the development, preparation, production and/or delivery of goods and/or services connected to these terms of reference.

## 4. Estimated time frames

Activity	Date
----------	------





Clarification questions deadline	03 August 2022
Submission of proposals deadline	17 August 2022
Evaluation of submissions	22 - 26 August 2022
Shortlisted service provider presentation date	31 August 2022
Commencement of work	01 September 2022
Completion	24 February 2023

## 5. Contact details

All queries should be directed to Boitumelo Mancini and must be submitted via email to [boitumelo@jet.org.za](mailto:boitumelo@jet.org.za).

Responses to queries and notification of shortlisting, Interviews and award of tender will be provided via email. Proposals should be submitted to [tenders@jet.org.za](mailto:tenders@jet.org.za). Technical and financial proposals may be combined.

## 6. Annexure A

Local	
Name	Link
OFO	<a href="https://www.ofomapping.co.za/Account/Login?ReturnUrl=%2fTasks">https://www.ofomapping.co.za/Account/Login?ReturnUrl=%2fTasks</a>
NQF	<a href="https://www.saqqa.org.za/faq/what-national-qualifications-framework-nqf?language_content_entity=en">https://www.saqqa.org.za/faq/what-national-qualifications-framework-nqf?language_content_entity=en</a>
Occupational Atlas	<a href="http://18.130.52.211/">http://18.130.52.211/</a>
Harambee	<a href="https://digital.harambee.co.za/">https://digital.harambee.co.za/</a>
Khetha	<a href="http://ncap.careerhelp.org.za/">http://ncap.careerhelp.org.za/</a>
ARPL	<a href="https://www.dhet.gov.za/Part%20E%20%20Guidelines/6.Criteria%20and%20Guidelines%20for%20the%20implementation%20of%20Artisan%20Recognition%20of%20Prior%20Learning.pdf">https://www.dhet.gov.za/Part%20E%20%20Guidelines/6.Criteria%20and%20Guidelines%20for%20the%20implementation%20of%20Artisan%20Recognition%20of%20Prior%20Learning.pdf</a>
Career Pathing in Manufacturing Industry	<a href="https://www.merseta.org.za/wp-content/uploads/2021/07/2019-May-New-Tyre-Chamber-Research-Report.pdf">https://www.merseta.org.za/wp-content/uploads/2021/07/2019-May-New-Tyre-Chamber-Research-Report.pdf</a>
International	

Name	Link
US O*NET-SOC	<a href="https://www.onetcenter.org/taxonomy.html">https://www.onetcenter.org/taxonomy.html</a>
ESCO	<a href="https://ec.europa.eu/social/main.jsp?catId=1326&amp;langId=en">https://ec.europa.eu/social/main.jsp?catId=1326&amp;langId=en</a>
UK Skills Taxonomy	<a href="https://data-viz.nesta.org.uk/skills-taxonomy/index.html">https://data-viz.nesta.org.uk/skills-taxonomy/index.html</a>
Emsi Open Skills Taxonomy	<a href="https://skills.emsidata.com/">https://skills.emsidata.com/</a>
WEF Global Skills taxonomy	<a href="https://www.reskillingrevolution2030.org/reskillingrevolution/insights/skills-taxonomy/index.html">https://www.reskillingrevolution2030.org/reskillingrevolution/insights/skills-taxonomy/index.html</a>
Credential Registry	<a href="https://credentialfinder.org/">https://credentialfinder.org/</a>