

Open Data Industry Connections Activity Initiation Document (ICAID)

Version: 2.0, 13 Feb 2018

IC16-001-02 Approved by the IEEE-SASB 8 March 2018

Instructions

- Instructions on how to fill out this form are shown in red. It is recommended to leave the instructions in the final document and simply add the requested information where indicated.
- **Shaded Text** indicates a placeholder that should be replaced with information specific to this ICAID, and the shading removed.
- Completed forms, in Word format, or any questions should be sent to the IEEE Standards Association (IEEE-SA) Industry Connections Committee (ICCom) Administrator at the following address: industryconnections@ieee.org.
- The version number above, along with the date, may be used by the submitter to distinguish successive updates of this document. A separate, unique Industry Connections (IC) Activity Number will be assigned when the document is submitted to the ICCom Administrator.

1. Contact

Provide the name and contact information of the primary contact person for this IC activity. Affiliation is any entity that provides the person financial or other substantive support, for which the person may feel an obligation. If necessary, a second/alternate contact person's information may also be provided.

Name: Wesley Diphoko

Email Address: wesleydiphoko@gmail.com

Phone:

Employer: Sagarmatha 4th Industrial Revolution Centre

Affiliation:

2. Participation and Voting Model

Specify whether this activity will be entity-based (participants are entities, which may have multiple representatives, one-entity-one-vote), or individual-based (participants represent themselves, one-person-one-vote).

Entity-Based

3. **Purpose**

3.1. **Motivation and Goal**

Briefly explain the context and motivation for starting this IC activity, and the overall purpose or goal to be accomplished.

What is Open Data?

- Open Data is data that can be freely used, re-used and redistributed by anyone - subject only, at most, to the requirement to attribute and share alike.
- **Availability and Access:** Open Data must be available as a whole and at no more than a reasonable reproduction cost, preferably by downloading over the internet. The data must also be available in a convenient and modifiable form.
- **Re-use and Redistribution:** the data must be provided under terms that permit re-use and redistribution including the intermixing with other datasets.
- **Universal Participation:** everyone must be able to use, re-use and redistribute - there should be no discrimination against fields of endeavour or against persons or groups. For example, 'non-commercial' restrictions that would prevent 'commercial' use, or restrictions of use for certain purposes (e.g. only in education), are not allowed
- **Not all data should be "open":** the focus is on non-personal data (i.e., data which does not contain information about specific individuals) and data not subject to national security restrictions.
- **More than text:** Open data may include non-textual material such as maps, genomes chemical compounds, mathematical and scientific formulae, and medical data. Problems often arise because these are commercially valuable or can be aggregated into works of value and cost money to produce.

Please see references [1], [2] and [3] for a more thorough description of 'Open'. Getting the right definition of Open is a key activity in this preparatory phase. The convention for both public and private organizations is that any access to all data is controlled. Control may be through access restrictions, licenses, copyright, patents and charges are levied for access or re-use. Conversely, some advocates of Open Data may argue that any restrictions are against the communal good and that these data should be made available without restriction or fee.

In this proposal we do not confine ourselves to only public institutions - private institutions are frequently required (and are well-served) through the regular reporting of material data, such as shareholders updates, and so on.

The implications for organisations making data open will create significant business opportunities and make institutions more accountable.

Goals:

The goals of this proposal are:

Goal 1: Open Data (2018)

- Identification of value add information
- Centralization of existing open data sets in an accessible manner
- OS Pilot: development of tools and analytical engine to view the data in standardized method
- Creation of a sustainable mechanism to enable open data maintenance and management
- Development of supporting documentation and training
- Facilitate training of youth in Open Data (maintenance,management)

Goal 2: Open APIs (2018-19)

- Identification of projects to instantiate re: open data based APIs developed in open source
- Source code sets that already exist with appropriate license to streamline the efforts
- Develop new APIs that offer the ability for data calls to occur and enable innovation
- Development of supporting documentation and training

Goal 3: Open Algorithms (2018-19)

- Review and contextualize EAD for Capetown (See Link: http://standards.ieee.org/develop/indconn/ec/ead_v2.pdf)
- Provide Feedback into the EAD community
- Creation of specific intelligent algorithms that provide context around the Capetown open data sets with an ethical design perspective
- Utilize a lab like setting to encourage community development and growth inclusive of Robotics
- Set up open algorithm "marketplace" for certified algorithms

Goal 4. Open Learning (2018-2020)

- Training around each phase so that volunteers understand how to contribute to the project and the importance of it
- Provide sample sets of information and events to support multi--stream learning

Goal 5. Community Engagement (Years 2018-2020)

- Power Africa adjacent workshop
- Africom potential hackathon or workshop
- Ongoing events and workshops where volunteers can share the successes of their progress
- Champion an Industry event (Open Data)

Goal 6. Sustainability (2018-2020)

- A sustainable, economic approach to the open data, apis, and algorithms such that its value is not full absorbed only by industry.
- Output will be a whitepaper that will offer suggestions for practical approaches to ensuring a thriving ecosystem

Connection points with following existing projects:

- Future City as a Lab: The Connected Person & Immersive City Industry Connections Program
- Smart City Compliance Indicators Industry Connections Program
- Global Initiative on AI & Autonomous Systems Industry Connections Program
- Internet of Things Interest Group Industry Connections Program
- India Robotics Roadmap Industry Connections Program
- PHD Cybersecurity Industry Connections Program
- Tie back to our Standards & Startup ecosystem and open source platform on Git

1.2. Related Work

Provide a brief comparison of this activity to existing, related efforts or standards of which you are aware (industry associations, consortia, standardization activities, etc.).

There is a considerable amount of attention given to data in general at the moment [4]. Sir Tim Berners-Lee, inventor of the World Wide Web, believes that in 2016 an Open Data Charter will be created [5].

Three prominent initiatives are parallel to and overlap Open Data: these are Open Government, Big Data and Smart Cities. We discuss these now.

- In 2009 the USA made all data collected by government “open by default” (except for personal information). Since the IEEE Open Data initiative will focus on local government, interaction with Open Government initiatives will be critical. We seek to engage participants who are actively using and creating this data (rather than macro policies).
- We see Open Data as being an initiative that focuses on smaller data sets than Big Data. Currently, Open Data sets that are made available by individual organisations are often small in size, static and uniform. The proposed Open Data standard would help define the process leading to the roll-out of this data, as well as address the lifecycle of this data.
- Smart Cities is an initiative that spans multiple disciplines, including Open Data.

1.3. Previously Published Material

Provide a list of any known previously published material intended for inclusion in the proposed deliverables of this activity.

There are extensive references with regard to Open Data. We have some of them noted in the Reference section at the end. We will be investigating and expanding these and additional published material for future inclusion.

1.4. Potential Markets Served

Indicate the main beneficiaries of this work, and what the potential impact might be.

The beneficiaries of an Open Data standard will be:

1. entrepreneurs who will get sustainable access to data to create useful service
1. big companies who will be able to facilitate local government's data needs
2. local governments who will be able to enhance their services
3. local companies who will be able to find better resources through better access
4. local government who will get taxable revenue streams
5. society as a whole who will benefit from openness

The above list of benefits is not exhaustive. The benefits of a universal Open Data are significant and profound. We would intent as part of this activity is to clarify benefits in different markets.

4. Estimated Timeframe

Indicate approximately how long you expect this activity to operate to achieve its proposed results (e.g., time to completion of all deliverables).

Expected Completion Date: 12/2020

IC activities are chartered for two years at a time. Activities are eligible for extension upon request and review by ICCOM and the IEEE-SA Standards Board. Should an extension be required, please notify the ICCOM Administrator prior to the two-year mark.

5. Proposed Deliverables

Outline the anticipated deliverables and output from this IC activity, such as documents (e.g., white papers, reports), proposals for standards, conferences and workshops, databases, computer code, etc., and indicate the expected timeframe for each.

The following deliverables are proposed:

- a. Identify and recruit a diverse set of participants from local governments and institutions to discuss their challenges and needs with regard to managing their data.
- b. The participants and current focus will be in South Africa and based on a Use Case of SA local government, with the intention to expand the project at the earliest opportunity engaging with other communities within IEEE to standardize platforms

- c. A report that summarizes the related initiatives and proposes ways to unite and harmonize these initiatives.
- d. A proposal (PAR) for one or more Open Data standards.

The definition of Open Data is being debated. Wikipedia uses a definition from OpenDefinition.org: "A piece of data is open if anyone is free to use, reuse, and redistribute it — subject only, at most, to the requirement to attribute and/or share-alike." However, Wikipedia goes on to acknowledge problems around the meaning. The proposed standard would attempt to address this and other issues.

Proposed standards would need to be consistent with the many data processes and models, various IEEE data standards and related standards.

There are a multitude of software tools for Data Analysis. The proposed standard should help users choose the correct ones.

- d. Host an international forum for interested parties to broaden global participation building on the South African use case
- e. A proposal for a Proof of Concept as a reference implementation for a local government Open Data standard.
- f. A proposal for a permanent organization structure for the initiative to continue after the incubation period in IEEE-SA Industry Connections ends.

6. Funding Requirements

Outline any contracted services or other expenses that are currently anticipated, beyond the basic support services provided to all IC activities. Indicate how those funds are expected to be obtained (e.g., through participant fees, sponsorships, government or other grants, etc.). Activities needing substantial funding may require additional reviews and approvals beyond ICom.

Initially the activity would consist of volunteer organisations and individuals giving of their time and resources to establish a common understanding and shared objectives. The basic Industry Connections support provided by IEEE-SA is viewed as sufficient to accomplish the initial developmental goals of the group.

Thereafter, any deliverables benefitting from funding support (e.g. forum on Open Data and any other identified) would be subject to developing a plan and gaining acceptance based on equal support from businesses, government and the IEEE. Some of this support will be in-kind, rather than financial.

7. Management and Procedures

7.1. IEEE Sponsoring Committee

Indicate whether an IEEE sponsoring committee of some form (e.g., an IEEE Standards Sponsor) has agreed to oversee this activity and its procedures.

Has an IEEE sponsoring committee agreed to oversee this activity?: No

If yes, indicate the sponsoring committee's name and its chair's contact information.

Sponsoring Committee Name: Committee Name

Chair's Name: Full Name

Chair's Email Address: who@where

Chair's Phone: Number, including country code

Additional sponsoring committee information, if any.

7.2. Activity Management

If no IEEE sponsoring committee has been identified in 7.1 above, indicate how this activity will manage itself on a day-to-day basis (e.g., executive committee, officers, etc).

The activity will be managed by an Executive Committee as defined in the baseline *Industry Connections Activity Policies and Procedures*.

7.3. Procedures

Indicate what documented procedures will be used to guide the operations of this activity; either a) modified baseline *Industry Connections Activity Policies and Procedures*, or b) Sponsor or Working Group policies and procedures accepted by the IEEE-SA Standards Board. The chosen policies and procedures must be reviewed by ICom

The baseline *Industry Connections Activity Policies and Procedures*

8. Participants

8.1. Stakeholder Communities

Indicate the stakeholder communities (the types of companies or other entities, or the different groups of individuals) that are expected to be interested in this IC activity, and will be invited to participate.

There are many organizations looking at data and open from one or another point of view. For example:

- Open Knowledge Foundation
- Sunlight Foundation
- IBM's Open Data for Africa
- Code4Africa

The City of Cape Town is a leader in Open Data [6] and the Provincial Government of the Western Cape likewise.

South Africa has been taking a lead in related Open Government initiatives [7]. In addition Statistics South Africa, National Treasury and the South African Bureau of Standards are collaborating in this regard [8].

We would start with a local South African community consisting of City of Cape Town and Provincial Government of the Western Cape. However, we would like to

rapidly include other local and regional councils in South Africa, Africa and beyond (e.g. India) [9].

Secondly, we would like to start with a number of local companies such as Telkom South Africa and Vodacom (part of Vodafone), but rapidly include other national and international companies such as Microsoft and IBM.

Thirdly, we would like to engage with research bodies and universities.

8.2. Expected Number of Participants

Indicate the approximate number of entities (if entity-based) or individuals (if individual-based) expected to be actively involved in this activity.

We would start with 3-7 entities and grow the participating members to 15 or more including participating entities from other open data initiatives outside South Africa.

8.3. Initial Participants

Provide a list of the entities or individuals that will be participating from the outset. It is recommended there be at least three initial participants for an entity-based activity, or five initial participants (each with a different affiliation) for an individual-based activity.

Use the following table for an entity-based activity:

We have begun testing the water with several of the stakeholders identified in section 8.1.

Entity	Primary Contact	Additional Representatives
Western Cape Economic Development Partnership	Jodi Allemeier jodi@wcedp.co.za	
Silicon Cape	Lianne Du Toit lianne@siliconcape.com 021 832 0200	
City of Cape Town	Andre Stelzner andre.Stelzner@capetown.gov.za	Ronelle.Rudman@capetown.gov.za Neil.Hoorn@capetown.gov.za
Provincial Government of the Western Cape	Olyvia Dyers olyvia Olivia.Dyers@westerncape.gov.za	
Cape Innovation & Technology Initiative	awaiting response	

Vodacom	Awaiting response	
Telkom	Awaiting response	
Stats South Africa	Awaiting response	

References

- [1] *Definition of Open* <http://opendefinition.org/od/2.1/en/> and the *normative reference*: https://en.wikipedia.org/wiki/Open_data
- [2] *Draft Open Data Policy*: <https://drafts.dc.gov/docs/draft-open-data-policy>
- [3] *Open data Handbook* <http://opendatahandbook.org/guide/en/what-is-open-data/>
- [4] *Out of the box*, The Economist, 21 November 2015, page 55
- [5] <http://www.theworldin.com/article/10648?fsrc=scn/tw/te/bl/ed/thewiderworldweb>
- [6] http://www.capetown.gov.za/en/PublicParticipation/Documents/HYS_Open_Data_Draft_Policy_version_21_February_%202014.pdf
- [7] <http://www.opengovpartnership.org/blog/dumisani-nkwamba/2014/08/08/south-africa-co-chair-global-open-government-partnership>
- [8] *City portals incubate innovation*, Stephen Timm, Business Day, 7 Dec 2015, <http://www.bdlive.co.za/business/innovation/2015/12/07/city-portals-incubate-innovation>
- [9] *Global Open Data Index*: <http://index.okfn.org/place/south-africa/>