MyData Dictionary (Draft)

Intellectual Property Statement - see Operators paper, needs to be open and free to use

The MyData Dictionary is an articulation of the key data fields an individual would wish/ expect to see in a data-set/ database that is designed to empower them, and that they control. **The MyData Dictionary is looking and built from the perspective of the individual**, not from that of organisations.

To bring that to life. There are some 7.8 billion humans on our planet; subject to some known and very specific anomalies, all have:

- A name
- A gender
- A date, time and place of birth
- Biological Parents (with names, times and places of birth)
- A home/ contact address

A sub-set of the above will have some very common add-ons:

- Phone number(s)
- Employer(s)

This list of fields does not claim or wish to be based on any one particular technology or existing standard. A first attempt is shown in Appendix 1. Clearly there will be anomalies to take into account but we should build for the simpler majority at this stage; trying to deal with every complexity is what takes one down into the weeds, never to return.

A core data dictionary might be useful in the following scenarios:

- As an individual, I might expect the many online forms that I need to fill in could become informed of and work with the core human schema. *Therefore, if the services the individual engages with utilise the core schema, form filling can become more automated and less onerous.*
- As an individual, I want to be able to pick up my data from one personal data service (locker etc.) and drop it into another with minimal fuss. *Therefore, if these services share, or at least understand (can map to and from) a commonly understood reference point then that will happen more easily.*
- As an individual, I want to be able to have applications run on my personal data, even when it exists in multiple different services. *Therefore, if these applications can all utilise a common, shared reference point, or at least map to and from such a thing, then that will happen more easily.*
- As an individual, I want the apps I had running on my personal data in one locker service, to work when I move my data to another one. *Therefore, if these applications share, or at least understand (can map to and from) a commonly understood reference point then that will happen more easily.*
- As an application developer, I want the apps I build to run, with minimal overhead, across multiple personal data services. *Therefore, I can build with the assumption*

that these data services will share a common model, or at least a common reference I can map to and from, then that will work more easily.

- As an organisation willing to receive and respond to MyData style data feeds; I don't want to have to set up different mechanisms for each different MyData operator. *Therefore, I can confidently build the API's I will need to help me co-manage data with my customers.*
- As an organisation willing to provide/ return data to customers as part of our proposition to them; I want to be able to make this data available in the minimum number of ways that meet the users needs, not a different format for each personal data service provider. *Therefore, I can build my data portability mechanisms in the knowledge that they will work for many different ways in which my customers may wish to port their data.*
- As an organisation willing to provide/ return data to customers as part of our proposition to them; I don't want to have a different button/ connection mechanism for each personal data service provider on the 'customers signed in' page of my web site. *Therefore, I can use some standardisable components to support subject access and data portability requests from my customers.*
- As a new MyData Operator, I don't want to have to re-invent the core database design from scratch; i'd like to build to an agreed model. *Therefore, I can get up and running more quickly, and have a degree of interoperability from day one.*

It is important that we do not focus on any particular technology stack for using this schema; but it is clear that the more modern ways of expressing schema (e.g. JSON LD) are a good start point. In turn, the emerging Overlays Capture Architecture (OCA) could be very useful in many ways, not least scaling the concept globally in language terms. In OCA terms, the MyData Dictionary would act as the base schema, example here.

It is worth noting that the above list of standardisable fields is only a start point; there are many hundreds of additional data attributes for which the individual is undoubtedly the best originator.

Our working assumption in the sub group is that, at least in the short term, we should steer clear of data attributes and sub-schema that require agreement from any type of organisation.

We have a github repo (https://github.com/mydataglobal) that we can use.

Questions/ Suggestions

What process do we need to make it credible? Review before publishing.

Should this form a recommendation made to all operators?

What should be the review period for all operators / proto-operators / community to give feedback

This group has the decision to publish once review period ends and all comments have been addressed

We will create guidance notes for operators on how to use the schema

We will maintain release numbering

Note - as human-centric data is, by definition, life-long, a general principle applied would be that each entry would have a validity period as meta data. That is to say ‘my employer’ has a start date and potentially an end data. A related point is that, by definition, an individual can have many variants of the same field (e.g. my email address), each with a validity period, and each ideally with tags to help describe/ further delineate.

Note 2: Field definition and descriptions must be precise and carry sufficient detail and pre-work to be unambiguous; i.e. be well defined.

Note 3: at this stage we are not dealing with derived or concatenated fields; this should be the raw data, low level individuals fields (so ‘address lines individually rather than combined, and no derivations such as ‘age’ from ‘birth date’.

Appendix 1 - Field Listing (all fields remains a work in progress at this stage)

Name	**Field Label**	**Expected Type**	**Description**	**Data Sensitivity Level**	**Example**	**Reference/ Mapping Equivalent(s).** (Examples, not a definitive listing)
my.email	My Email	String (or email)	An email address		alice.mcperson2000@gmail.com	Contact Field = Email in Salesforce.com contact object; schema.org email (https://schema.org/email)
my.firstname	My First Name	Text	A person's first (given) name		Alice	
my.lastname	My Last Name	Text	A person's surname		McPerson	
my.salutation	My Salutation	Text	A person's desired salutation		Ms	
my.birthdate	My Birth Date	String	A person's birthdate and optionally birthtime in RFC 3339 format			
my.gender	My Gender	Text	A person's gender			
my.mailingstreet	My Street (postal address)	Text	A street name and number or other address description for postal purposes			
my.mailingcity	My City (postal address)	Text	A city name for postal purposes			
my.mailingstate	My State/ Province/ County (postal address)	Text	A state, province, county or other area designation for postal purpose			
my.mailingpostcode	My Postcode (postal address)	Text	An area postal code			
my.mailingcountry		Text	A country name or code designation for postal purposes			
my.homephonenumber	My Home Landline Phone Number	String	A landline phone number			
my.mobilephonenumber	My Mobile Phone Number	String	A mobile phone number			

my.faxnumber	My Fax Number	String	A fax phone number			
my.businessname		Text	A business name			
my.businessindustry		Text	An industry name associated with a business			
my.businessstreet		Text	A person's business title associated with a business			
my.businesscity		Text	A business street name and number or other address description for postal purposes			
my.businessstate		Text	A business state, province, county or other area designation for postal purposes			
my.businesspostcode		Text	A business area postal code			
my.businesscountry		Text	A business country name or code designation for postal purposes			
my.businessphone		String	A business phone number			
my.currentemployer			To be discussed			
my.currentposition			To be discussed			

currentposition The current position of the person

Long List of Potential 'My' Data Attributes (developed in conference open space brainstorming)

As we wish the MyData Dictionary to evolve considerably and quickly over time, we wish to build a long list of candidate data attributes that could be evolved and enriched to then become part of the dictionary.

Suggestion: In building the long list we should consider only attributes that sit nicely with the form 'my….xxxxx'. We should avoid listing attributes that are clearly relating to another data object other than a person. So, 'my car' is a good choice, 'my car's VIN number' is not (it could be at a stretch, but realistically' VIN number is an attribute of the vehicle, not the person'. Critically, we should be using language in developing this long list that would not be a big technical stretch for a mass, adult audience. What we are working towards is a data model/ schema that might be incredibly technical and complex under the hood, but which should surface in ways that would be readily explainable to a mass audience.

Examples, showing different data areas

My Height
My Weight
My Waist Size
My Shoe Size
My Blood Type
My Blood Pressure
My Cholesterol Count
My Mobile Phone Type
My Mobile Phone Service Provider
My MyData Membership Expiry/ Renewal Date
My Favourite Colour
My Marital Status
My Qualifications
My Location
My Number of Dependents
My Life Stage
My Native Language
My Communication Channel Preferences
My Maiden Name
My Married Name
My Net Worth
My Covid Status
My Covid Risk
My Regular Foods
My Habits
My Friends List
My Customer Number
My Gender
My Assets, Products and Services
My DNA

My Credit Rating
My Next of Kin
My Emergency Contact Person (ICE)
My IP Address
My Mac Address
My Device ID
My Bank Transactions
My Credit Card Transactions
My Product/ Services Reviews
My Supplier Ratings (net promoter)
My Permissions
My Notifications
My Allergies
My Favourite Brands
My Favourite Coffee
My Apps
My User Names
My Passwords
My Medical Prescriptions
My Eyes/ Sight Prescription
My Fingerprints
My Iris Scan
My Face
My Photos
My Videos
My Movies
My DVDs
My Sports
My Exercises
My Electoral Roll Registration
My Votes
My Religious Beliefs
My Sports Teams
My Regular Shops
My Regular Cafes
My Electronic Documents
My Pensions
My Investments
My Insurances
My Savings
My Experiences
My Algorithms
My Profiles
My Supplier/ Product Complaints
My Supply Needs
My Advertising Preferences
My Sustainability Preferences
My Risks

My Trips
My Journeys
My Digital Interactions (in and out)
My Family Members
My Receipts
My Taxes
My Life Events
My Bereavements
My Persona(e)
My Roles
My Carbon Footprint
My Nutritional Intake
My Daily Footsteps
My Education
My School
My Education Title
My Graduation Day
My Graduation Score
My Education History
My Badges
My Education Certifications
My Professional Certifications
My Professional Skills
My Soft Skills
My Residence History
My Wearables
My Social profiles
My Banks List
My Sleep Duration
My Heart Rate (spot rating)
My Heart Rate (string of values)
My SPo2 level
My Calories Burnt Today
My Body Mass Index
My Stress Level
My Popularity (likes)
My Dates
My Sexual Preferences
My Metabolism
My Momentum
My Social Class
My Caste
My User Names
My Health Operations/ Medical / Aesthetic Procedures
My Music Collection
My Computer Settings
My Books
My Crimes/ Criminal Record

My Driving Record/ History
My Flights
My Hotel Stays
My Vaccinations
My Covid Tests
My Fluid Intake
My Medical Symptoms
My Private Keys
My political activities
My aesthetic preferences
My Retail Invoices
My Values
My Consent Receipts
My Browsing History
My library records
My Invoices in Arrears
My Tasks
My Icons
My period cycle
My Mood
My Energy Level
My Measurements
My Suppliers

Add Data Privacy Guide listing