



---

**To:** Finance & Resources Policy Board

**On:** 27 January 2016

---

**Report by:** Chief Executive

---

**Heading:** Initial Open Data Strategy for Renfrewshire

---

**1. Summary**

- 1.1 The purpose of this report is to bring to the attention of the F & R Policy Board the ongoing work by the council and its community partners towards publishing their non-personal and non-commercially sensitive data in an open electronic format for wider societal benefit.
- 

**2. Recommendations**

- 2.1 Members are asked to approve the implementation of the attached 'Initial Open Data Strategy for Renfrewshire' as it relates to Renfrewshire Council.
-

### 3. **Background**

- 3.1 Open Data is non-personal and non-commercially sensitive data which is accessible to anyone free of restriction on use and in a machine readable format. It is easily discoverable and can be freely used, reused and redistributed by anyone for any purpose including commercial without restriction i.e. it has an open license. It is made available, via the Internet, in an electronic machine readable format which supports its ready re-use for free. There are no royalties to be paid or patents to be adhered to.
- 3.2 In April 2014 the Scottish Data Management Board (DMB) published 'A Data Vision for Scotland' and set the overarching objective as being to champion and unleash across Scotland trustworthy uses of data for public benefit. The vision sets out where the Scottish Government wants to be in relation to open data by 2020. The DMB has created a strategic plan for effective and responsible collection, management and use of data across Scottish Public Services to the benefit of the Scottish public and in particular to achieving high quality public services and economic growth. The Scotland strategic plan has been developed with, and for, the Scottish public sector and its partners across the wider data community.
- 3.3 The DMB has also overseen the development of the Scottish Government's (SG) 'Open Data' strategy to help achieve the Data Vision and strategic plan. The SG's 'Open Data' strategy released in February 2015 seeks to create a Scotland where non-personal and non-commercially sensitive data from public services is recognised as a resource for wider societal use and as such is made open in an intelligent manner and available for re-use by others.
- 3.4 An 'Initial Open Data Strategy for Renfrewshire' has been created in accordance with SG guidance. It illustrates how the council with its community planning partners will deliver its open data vision including how data will be made open and the broad technical architecture required to publish.
- 3.5 The strategy for Renfrewshire contains a proposal to create an open data portal which will provide people with a convenient website to access data from the council and its community planning partners in one place. Its aim will be to increase the discoverability of data about Renfrewshire, its people, services, businesses, environment and economy. The portal will make it easier for people to find out what data is available from the council and its partners.

- 3.6 There will be no restriction on intended users of Renfrewshire's open data portal. The portal will be useful to the full range of current and potential data users, from normal citizens to web developers and commercial businesses in all sectors within and beyond Renfrewshire. Users will have the ability to download any or all data at no cost, other than their own reproduction cost, and with no patent restrictions. The download will be in an open, machine readable format to allow for the greatest re-use of data.
- 3.7 While the council has taken a lead role in complying with the Scottish Government's 'Open Data' strategy and initiating a strategy for Renfrewshire, it's believed the vision for open data in Renfrewshire should be holistic from a Community Planning Partnership perspective. To that end the attached background paper, 'Initial Open Data Strategy for Renfrewshire' has gone to and been approved by the CPP board.
- 

### Implications of the Report

1. **Financial** – There will be a financial implication but this is expected to be shared across the Community Planning Partnership and is expected to be absorbed in normal budgets.
2. **HR & Organisational Development** – The initial set up of processes, input to design and build, and the overseeing and maintenance of the open data portal are expected to be absorbed within the current Policy and Commissioning and ICT service areas. It is also hoped through the open data advocates network that there will be technical input from the partners particularly by Renfrewshire's educational institutions.
3. **Community Planning** – The Community Planning Partnership Board have approved the 'Initial Open Data Strategy for Renfrewshire'.
4. **Legal** – The council's Information Governance Team are fully aware of the open data agenda and support the 'Initial Open Data Strategy for Renfrewshire'.
5. **Property/Assets** - none.
6. **Information Technology** – As part of the open data strategy, a new portal platform will eventually be required. The ICT Applications Manager is fully aware of the impact the strategy will have to ICT. He has contributed to the open data strategy particularly in regard to how the broad IT architecture will work with the introduction of the open data portal.

7. **Equality & Human Rights** - The recommendations contained within this report have been assessed in relation to their impact on equalities and human rights. No negative impacts on equality groups or potential for infringement of individuals' human rights have been identified arising from the recommendations contained in the report because for example it is for noting only. If required following implementation, the actual impact of the recommendations and the mitigating actions will be reviewed and monitored, and the results of the assessment will be published on the Council's website.
8. **Health & Safety** – none
9. **Procurement** – Procurement of suitable platform software will be required in due course.
10. **Risk** - none
11. **Privacy Impact** - none

---

### List of Background Papers

Initial Open Data Strategy for Renfrewshire

The foregoing background paper will be retained within the Policy and Commissioning Service for inspection by the public for the prescribed period of four years from the date of the meeting. The contact officer within the service is Peter Rub, Data Analytics & Research Team.

Phone: 0141 618 7543, [peter.rub@renfrewshire.gcsx.gov.uk](mailto:peter.rub@renfrewshire.gcsx.gov.uk)

---

**Author:** Peter Rub, Extension 7543



# Initial Open Data Strategy

for

Renfrewshire

*How Renfrewshire Partners will make their data open to everyone.*

# CONTENTS

<b>CONTENTS</b>	<b>2</b>
<b>PURPOSE OF THE STRATEGY</b>	<b>4</b>
<b>WHAT IS OPEN DATA?</b>	<b>5</b>
<b>5 Star Open Data model</b>	<b>5</b>
<b>WHY MAKE DATA OPEN?</b>	<b>6</b>
<b>Legislation</b>	<b>6</b>
<b>Other Drivers</b>	<b>6</b>
Innovation	6
Improved public services and transparency	7
<b>Other Organisations</b>	<b>7</b>
<b>VISION FOR RENFREWSHIRE'S OPEN DATA</b>	<b>8</b>
<b>Open Data by Default</b>	<b>8</b>
<b>Quality and Quantity</b>	<b>8</b>
<b>Useable by All</b>	<b>9</b>
<b>Releasing Data for Improved Governance</b>	<b>9</b>
<b>Releasing Data for Innovation</b>	<b>9</b>
<b>Timeline for Vision</b>	<b>9</b>
<b>POTENTIAL DATA CONSUMERS</b>	<b>11</b>
<b>HOW WE'LL MAKE DATA OPEN</b>	<b>12</b>
<b>Openness philosophy</b>	<b>12</b>
Keep it simple	12
Completeness	12
Encourage participation and address common fears	12
Timeliness	12
Ease of electronic access	12
Machine readability	12

Non-discrimination .....	12
Licensing .....	12
<b>Open Data Management.....</b>	<b>12</b>
<b>Governance .....</b>	<b>13</b>
<b>Datasets and Sources .....</b>	<b>14</b>
<b>Metadata.....</b>	<b>14</b>
<b>Licensing.....</b>	<b>15</b>
<b>HOW WE INTEND TO PUBLISH DATA .....</b>	<b>16</b>
<b>Broad Technical Architecture .....</b>	<b>16</b>
Extract .....	16
Enrich/Transform.....	17
Load.....	17
<b>Platform Requirements .....</b>	<b>17</b>
Hosting .....	17
Administration.....	17
Consumer browse facilities .....	18
Consumer use .....	18
<b>Control and Data Quality.....</b>	<b>18</b>
Select.....	18
Create:.....	19
Edit: .....	21
Validate .....	21
Publish .....	21
Maintain .....	21
Roles.....	22
<b>Cost of Opening Data .....</b>	<b>22</b>
<b>Platform Options.....</b>	<b>23</b>
<b>Design and Build of Platform.....</b>	<b>23</b>
<b>TRAINING .....</b>	<b>24</b>
<b>INITIAL ACTION PLAN.....</b>	<b>24</b>

## **PURPOSE OF THE STRATEGY**

This strategy is being developed not only to comply with the Scottish Open Data Strategy but also to demonstrate the commitment by the council and its Community Planning Partners in making their data open to everyone. This strategy is the first edition of what will be an evolving strategy as the open data phenomenon takes hold and spreads throughout the Renfrewshire's Community Planning Partnership as it has throughout the world. This edition will explain why and how the council and partners will deliver their open data in a convenient and easy to access manner with the aim to increase the discoverability of data about Renfrewshire, its people, services, businesses, environment and economy.

Renfrewshire's open data strategy will continue to be developed and implemented using relevant guidelines and assistance issued by or obtained from Scottish, UK, EU and other relevant bodies. e.g.:

- The G8 Open Data Charter
- Scottish Government Open Data policies and strategies
- Open Knowledge Foundation
- Open Data Institute
- Other Local Authorities



## WHAT IS OPEN DATA?

For the purpose of this strategy, open data is non-personal and non-commercially sensitive data. It does not contain personal information relating to individuals or information which could be used to identify individuals. Additionally, information which could cause economic harm if released is not within the scope of open data. Discretion on the balance of public interest with transparency against the right to confidentiality will need to be made. The default position will be to release the information unless there is good reason not to. Open data is easily discoverable, accessible to anyone and can be freely used, reused and redistributed by anyone. It is made available, via the Internet, in an electronic machine readable format which supports its ready reuse for free. It can be used by consumers for any purpose including making money from it. There are no royalties to be paid or patents to be adhered to. Information data sets are not just published on their own. Other features must accompany the information for it to be considered as open data. These include:

- Available at no cost to the user
- Freely available to be used, redistributed and reused by anyone for any purpose including commercial, without restriction. i.e. have an open license
- Available online in machine readable non-proprietary formats;
- Easily discoverable through use of relevant metadata.

### 5 Star Open Data model

The Scottish Government has suggested a 5 star open data model should be what organisations are aspiring to achieve when making their data open. A summary of the model follows in the table below. Renfrewshire will be aspiring to release its data in a 3 star or above format.

★	Data available online with open license permitting re-use. Examples – Tables and charts in PDF document or scanned images
★★	Data available online in a machine readable format, with open license permitting re-use. Examples – Excel tables and charts
★★★	Data is available online, in non-proprietary machine readable format, with open license permitting re-use. Examples – Comma Separated Values (CSV) Extensible Mark-up Language (XML)
★★★★	Data is available online, in non-proprietary machine readable format, with open license permitting re-use. Data is described in a standard way and uses unique reference indicators, so that people can point to your data.
★★★★★	Data is available online, in non-proprietary machine readable format, with open license permitting re-use. Your data uses unique references and links to other data to provide context.

## WHY MAKE DATA OPEN?

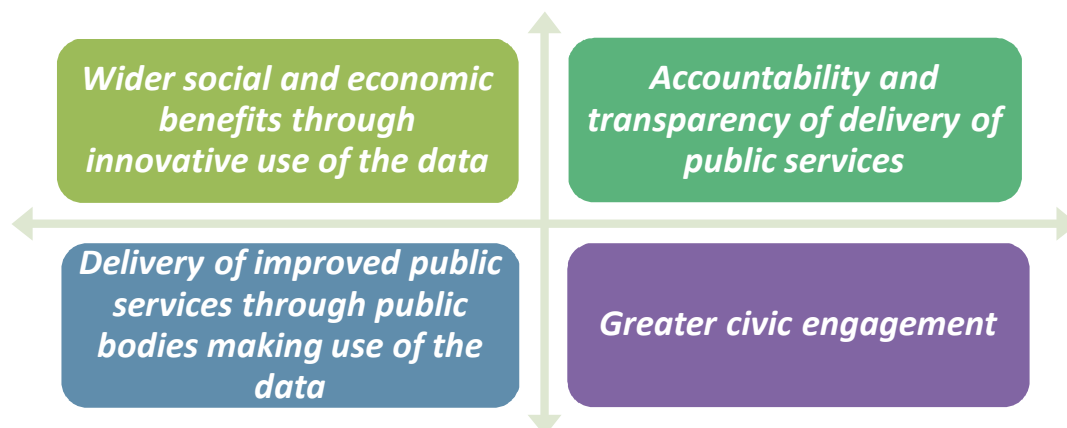
### Legislation

There are a number of legislative requirements that the implemented open data strategy will comply with by publishing its data. i.e.

- Freedom of Information (Scotland) Act 2002;
- Environmental Information (Scotland) Regulation 2009;
- Data Protection Act 1998;
- INSPIRE (Scotland) Regulations 2009;
- Re-Use of Public Sector Information Regulations 2015.

### Other Drivers

The council and its partners are large producers and collectors of information including non-personal and non-commercially sensitive data. This data is a valuable public resource which has been underused. Making the data open to the public will help to realise the full potential of the data and create many benefits the council, partner organisations, business and the community will receive. These include



### ***Innovation***

Publishing data in a reusable form empowers others to use the data for new and exciting purposes. Open data can be used in isolation or with other data to provide new insights or to develop apps which allow people to use the data in useful ways. Some examples of this include:

- London Schools Atlas: Projections show up to 4,000 extra primary classes across London are needed by the end of the decade. The pattern of which schools serve each area in London is extremely complicated but for the first time this app allows educational planners and parents to see the picture across London.
- Walkonomics: an app that rates and maps the pedestrian-friendliness of streets and urban areas.
- CollisionMap: A free service that displays and gives details of road accidents that involved an injury for the whole of Britain.
- Low Price Lessons: Driving lessons website helping learner drivers find their driving instructor and information about the areas they'll learn to drive in.

- Illustreets: an app useful when looking to buy or rent property putting deprivation, crime, education, transport, environment and census data on an interactive searchable map to compare locations while on the move.
- Check Business: provides risk and stability scores on millions of UK companies helping business to trade better with other UK firms.
- Informedtable: allows a search for restaurants near you and clearly shows their hygiene inspection ratings plus has an easy online booking service.

### ***Improved public services and transparency***

Making data held by public organisations open enables communities and individuals to understand more about public services, gain insight into their own community and contribute to future design and delivery. Studies by the EU have identified that people are most interested in gaining access to information about their local area. Making data more accessible can also help those delivering public services make better use of data themselves.

### **Other Organisations**

The adoption of open data is being embraced and adopted by many local authorities and public organisations across the UK and the world with varying degrees of success, quality and quantity. Examples include

Glasgow: <https://data.glasgow.gov.uk/>

Edinburgh: <http://opendata.cechosting.org/>

London: <http://data.london.gov.uk/>

Nottingham: <http://www.opendatanottingham.org.uk/>

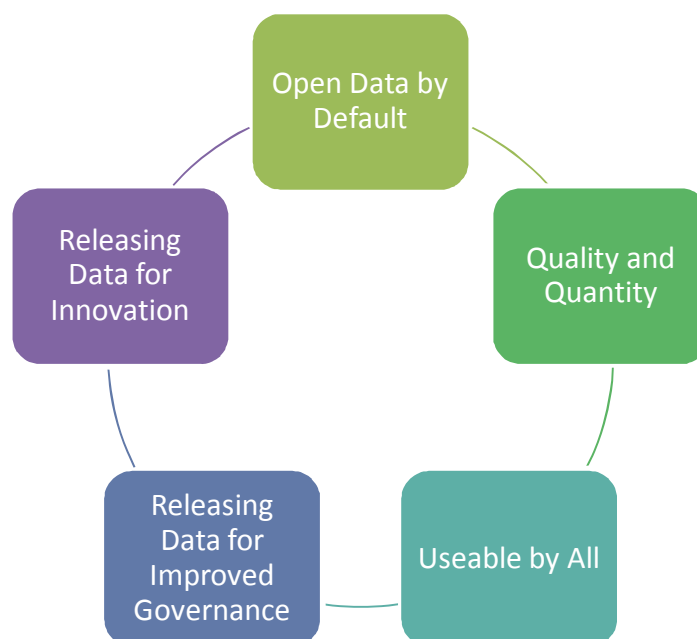
Paris: <http://opendata.paris.fr/page/home/>

New York City: <https://nycopendata.socrata.com/>

South Australia: <https://data.sa.gov.au/>

## VISION FOR RENFREWSHIRE'S OPEN DATA

The council and its partners gather many types of data to help inform their planning and decision making processes and deliver services in the most proficient and efficient manner. Making their data available will add value to work in communities, organisations and businesses. The council and its partners believe that open data will encourage innovation, develop new approaches to delivering services and develop new systems to support the quality of life in Renfrewshire. The vision is to introduce open data practices in line with the Scottish Government's Open Data Strategy which adopt the following G8 Open Data Principles. These practices will become embedded within the council and partner's normal information processing flows and complement their ICT, information governance, digital, records management and geographical information strategies and policies



### Open Data by Default

Those holding public data should make it open and available for others to re-use. Those collecting new data should ensure that the opening up of the data for re-use is built in to the process becoming part of the normal business of handling information.

### Quality and Quantity

The amount of public data owned by the council and partners is huge but the quality of that data will vary. Published data will be supported by metadata so that consumers of the data understand the data context and are aware of any limitations within it. Whenever possible, data will be released in a timely and frequent manner.

## Useable by All

Data should be published in a manner which supports both easy discovery and easy re-use of the data. This includes ensuring that the format it is published in supports re-use and that it is accompanied by an open licence. Data will be made available free, with defined exceptions.

## Releasing Data for Improved Governance

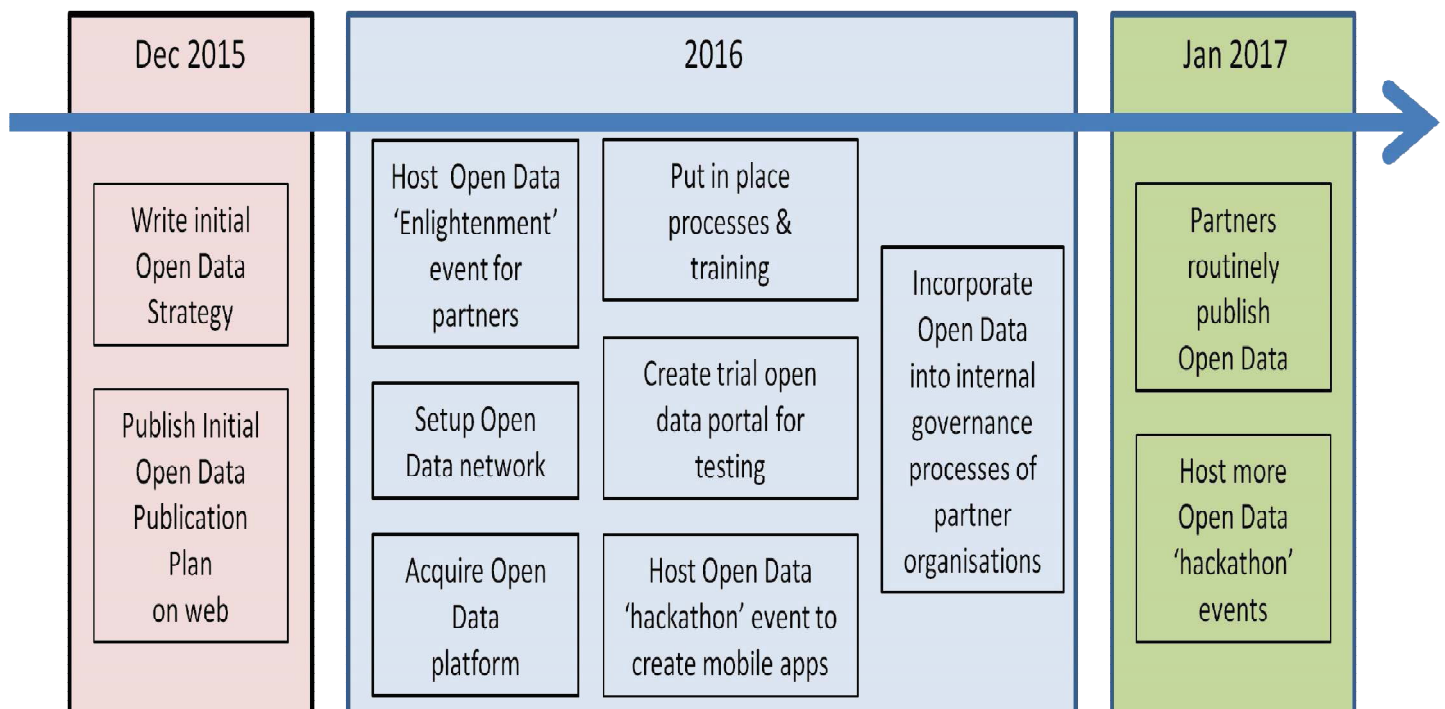
Data released will support delivery of better public services. Data will be used to improve services and policies that are delivered. The council and partners will seek through the release of data to better inform and engage with citizens.

## Releasing Data for Innovation

The council and partners will encourage and empower others to make use of the data released to develop new products and services, for non-commercial and commercial use, which will create wider economic and societal benefits. They will encourage use of the data in education to increase awareness and participation and inspire a new generation of data users.

## Timeline for Vision

To achieve this vision will require enhanced provision and presentation of data as well as improvements in governance and data consumer engagement, supported by a refreshed technical architecture within the council and its partner organisations. The time line below shows the broad tasks to carry out and achieve the vision.



By the end of 2015 the initial strategy and publication plan of open data sets, as required by Scottish Government, will be published on the council's website. Early on in 2016 an event to inform partners and gather their support and input will be conducted. An open data network within the council and across willing partner organisations will be established to drive forward the strategy in a collaborative manner. A mutual decision on a suitable open data platform will be made and acquired. Key personnel in the council and partner organisations will undergo appropriate training and appropriate information flow processes and roles will be determined during the year. Work to build, design and use the new portal platform will be undertaken with a proposed beta version being created, trialled and developed until affirmed fit for purpose. A 'Hackathon' will be conducted with the aim of having participants using Renfrewshire data to create Renfrewshire specific mobile applications thus demonstrating one of the uses of open data. Any data 'open by default' processes and work practices will be identified and suggested implementation highlighted to partners and appropriate service areas within the council where necessary.

By the beginning of 2017 routine publishing of open data will have commenced with the council and partners publishing their open data in a 3 star or above format on a single Renfrewshire partnership open data portal. (3 star data is data made available online, with an open license, in an open non-proprietary machine readable format).

## POTENTIAL DATA CONSUMERS

Open data is a phenomenon happening across the globe not just in Scotland. Potential users can be located locally within Renfrewshire or further afield in the UK, Europe or anywhere in the world. Potential open data consumers include:

- Application developers
- Bloggers and journalists
- Business, including those who use the data themselves and those who add value to data and resell it
- Charities and other civil society organisations
- Council staff, chief officials and officers
- Customers and service users
- Data scientists
- Education establishments and students
- Elected members
- Front end web developers
- General public
- Other public organisations
- Researchers and academics
- Software developers
- Special interest groups
- Suppliers
- Visitors to Renfrewshire

Different types of data consumers will have different needs and skills for accessing and using different data formats. Consumers who are more visually orientated will need data presented in an easily accessible and digestible ready to use format. Those who have an interest in ICT but lack formal skills may try using any available tools and consumers who have specialist ICT training will have the skills to analyse and interpret the data. The following chart correlates the type of consumer with typical open data formats they might use.

Consumer	Browser			Download			Programmatic Interface				
	Report	Infographic	Applet	Spreadsheet	CSV	json	API	XML	REST	Linked data	SPARQL
<i>General public/Staff</i>	✓	✓	✓								
<i>Blogger/journalist</i>	✓	✓	✓	✓							
<i>Researcher/academic</i>	✓	✓	✓	✓	✓						
<i>Data analyst</i>			✓	✓	✓						
<i>Web developer</i>				✓	✓	✓	✓				
<i>App developer</i>					✓	✓	✓	✓			
<i>Web app developer</i>						✓	✓	✓	✓	✓	✓
<i>Data scientist</i>					✓	✓	✓	✓	✓	✓	✓



## HOW WE'LL MAKE DATA OPEN

### Openness philosophy

The council and partners' provision of open data will be based on the following philosophy:

**Keep it simple:** Start out small and simple. There is no requirement that every dataset must be made open on day one of the open data portal going live. The process will be iterative with a catalogue of datasets built up over time.

**Completeness:** In due course all public data should be made available. i.e. data that is not subject to valid privacy, security or privilege limitations, is non-personal and non-commercially sensitive is made open.

**Encourage participation and address common fears:** Get people engaged with the process and allay any fears arising out of people not fully understanding the requirement, reasons or usefulness of making public data open.

**Timeliness:** Data is made available as quickly as necessary to preserve the value of the data.

**Ease of electronic access:** Data is available to the widest range of users for the widest range of purposes.

**Machine readability:** Data is reasonably structured to allow automated processing.

**Non-discrimination:** Data is available to anyone, with no requirement of registration.

**Licensing:** Data is not subject to any copyright, patent, trademark or trade secret regulation.

### Open Data Management

Initially the Data Analytics and Research team within the council will be responsible for the stewardship of the whole Open Data process. This team will need to liaise and work with partners and all service areas within the council especially ICT, who will be responsible for the extraction of data from data storage within the council, and the Information Governance team, who will oversee that only data that is legally permitted is being released.

A network of open data advocates from willing partners and council service areas will be established. They will be people interested in making the open data project a success and be the point of contact for each partner organisation or service keeping the wider organisations updated on progress. They will take an active role in open data support by identifying suitable data from their organisation or service area and supplying corresponding metadata for data to be published. They will also be responsible for embedding the philosophy and processes for making data generated by their organisation and service area open by default while endeavouring to minimise the impact on their own service areas.

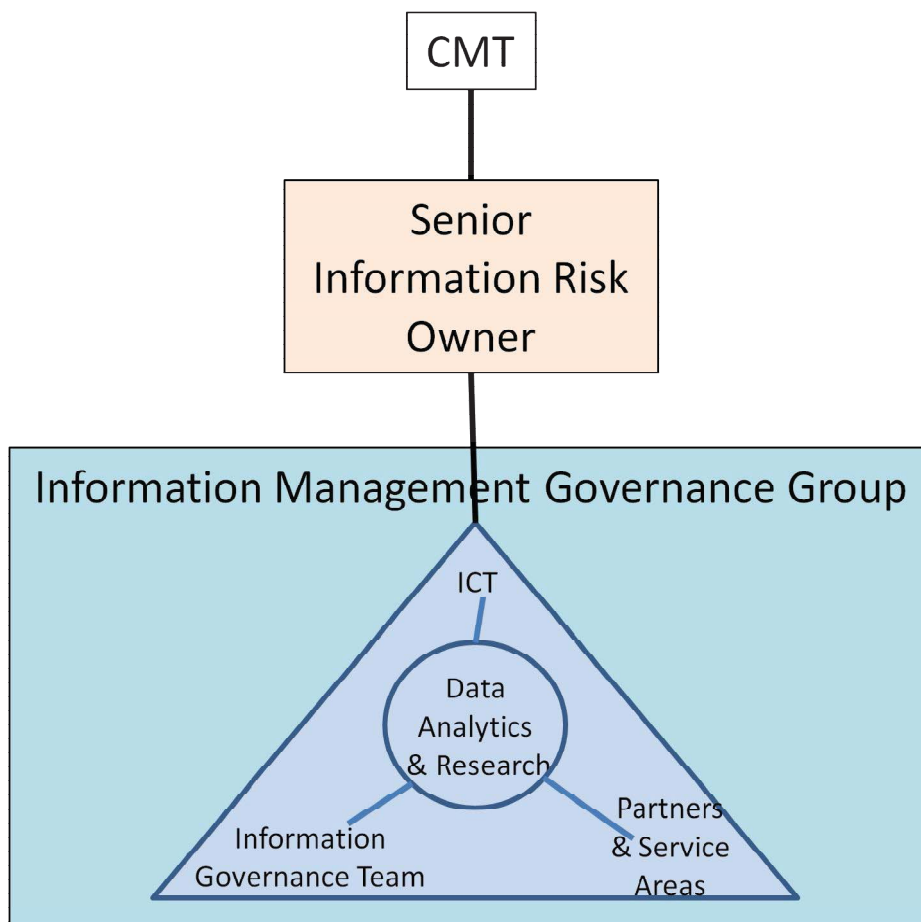


When deciding if a dataset should be published or not, discretion and balancing the public interest of transparency against the right to confidentiality will be made. The default position will be to release the information without attempting to prevent its release unless there is good reason.

## **Governance**

The following initial governance arrangements of the open data strategy will be put in place acknowledging that as the open data is deployed and partner organisations become more involved the arrangements may need to be reviewed to ensure they remain robust and appropriate.

Initially the management of the open data process will be undertaken by the Policy and Commissioning service's Data Analytics and Research team under guidance of the Information Management Governance Group (IMGG) and reporting to the Senior Information Risk Owner who will in turn be reporting to the Corporate Management Team. Partner organisations will have their own governance arrangements for making their datasets available to be published. These may be included in later versions of this strategy.



## **Datasets and Sources**

Datasets will originate either from within the council, from partner organisations or from external sources. The datasets can be about anything as long as they are non-personal and non-commercially sensitive. Data held by the council are stored in a myriad of places and across multiple platforms, e.g. databases, spreadsheets, folders, documents and websites. Large data volumes of council and partner data are likely to correspond with key or critical business systems while low volume data will be more isolated in spreadsheets or similar scattered in service area folders on the council's corporate network.

There is a significant amount of Renfrewshire related data publically available on the Internet from various public, private and third sector organisations. Most public sector organisations like National Records Scotland will publish good quality data, metadata and have robust download procedures, while other organisations will perhaps have less well organised or less formal data sharing facilities. The council will undertake to scrutinise and publish data that is of interest to Renfrewshire open data consumers by working with partners to enhance, if necessary, the quality of source data and arrange the most appropriate data handling procedures. This could include the council extracting and holding other source data or simply employing links to partner or other external sourced data.

## **Metadata**

Data can only be used effectively if metadata is also provided. Metadata is descriptive information about the data. Metadata assists in the recovery of the data, gives context about the data, ensures the data is properly understood and allows better interoperability with other datasets. Metadata will accompany every data set that is published. Many features about the data will be described such as the content, format, currency, frequency of updates, coverage and limitations. To some degree datasets created within the council will already have some corresponding metadata available. The challenge will be to introduce a uniform metadata standard for council and partners that provides a minimum amount of descriptive information for all datasets.

As the portal and datasets mature, work towards having all metadata conforming to Scottish Governments recommended intermediate metadata standard, Dublin Core, will be ongoing with the aim of progressing to the Data Catalog Vocabulary (DCAT) standard which is used to describe all public data in Europe. Geospatial data metadata will conform to INSPIRE standards.

## **Licensing**

Applying a license to the data explicitly indicates to the users what they can and cannot do with the data. It provides clarity and certainty. Applying an open licence allows people and organisations to re-use, modify and share the content in any way they wish. However open licences can have two restrictions attached to them:

- Attribution: Users must acknowledge the source of the data; and
- Share-alike: users must publish any derived data under the same licence.

Therefore, open licenses can have no restrictions (all rights waived), attribution or attribution and share-alike. Requiring attribution will help promote Renfrewshire's open data initiative as users have to link back to the original work. Share-alike restrictions will require users of the data to publish their work openly which may deter commercial businesses and people who want to make profit from their use of the data, resulting in reduced innovation and use.

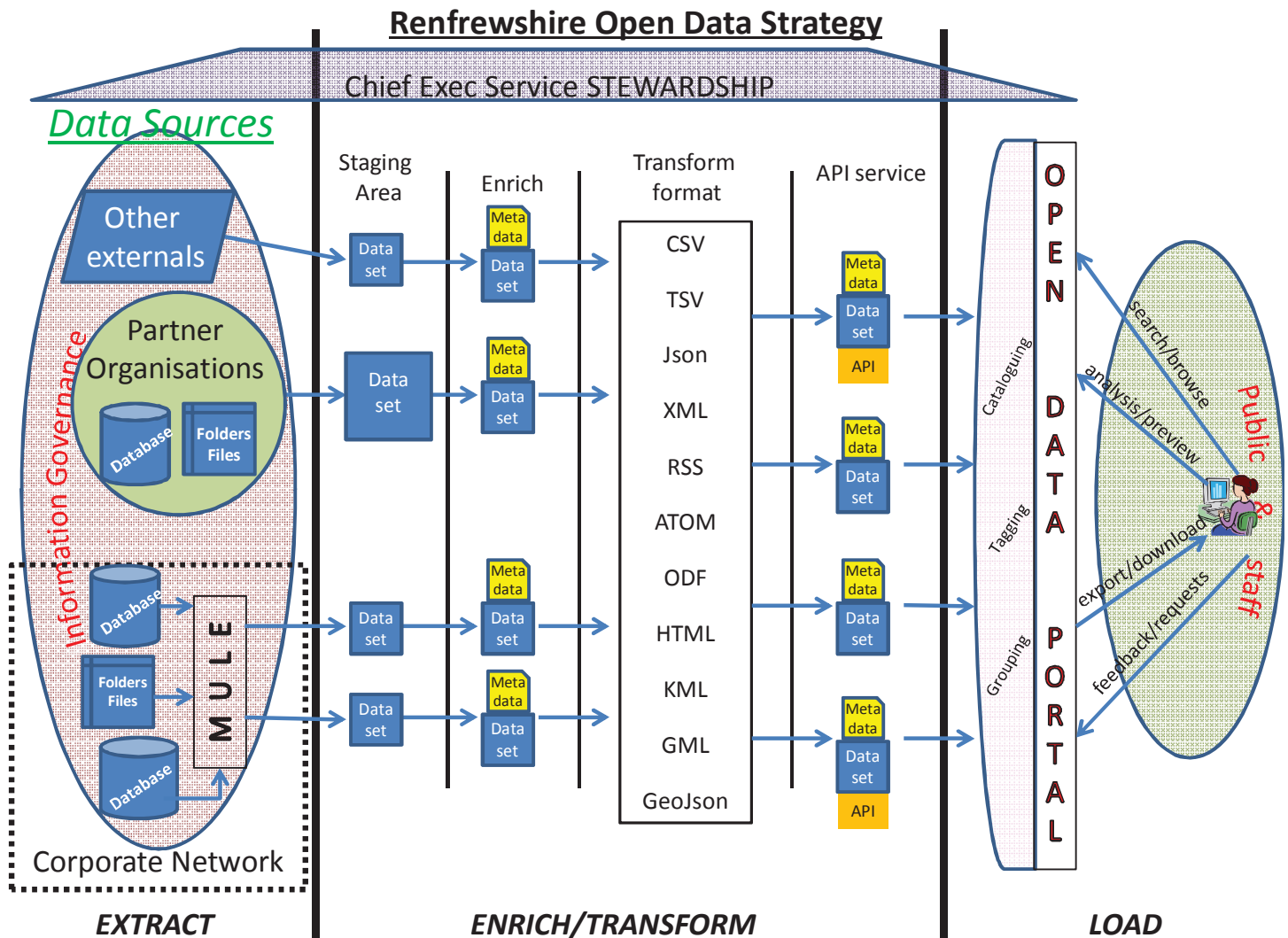
Unless otherwise indicated, Renfrewshire's data will be published using

- the Open Government Licence (OGL) which must be used by Crown organisations that have gathered or created data; or the
- the same licence as the original data if publishing data derived from data published under a share-alike licence.

There will be no charge for data licensed under the OGL and users will be signposted to the licence by using both human-readable and machine-readable descriptions.

# HOW WE INTEND TO PUBLISH DATA

## Broad Technical Architecture



A technical architecture similar to the broad outline in the diagram above will be employed. While partner organisations and individual service areas remain responsible for the veracity of their data, the Data Analytics and Research team will initially be responsible for the stewardship of the whole Open Data process from co-ordinating the initial identification of native data to making the data open and available to consumers. There after partners and service areas will be encouraged to promote and put forward their datasets for publication. There will be three basic steps to making the data available.

### ***Extract***

The Mule application is already in use within the council and can already harvest data i.e. ICT can already pull datasets from different databases and data storage areas and send to wherever the data is required. In this case it will be sent to a staging area. Data sets from partner organisations and external sources will also be required to be sent to the staging area by a process yet to be determined.

### ***Enrich/Transform***

In the staging area each data set will be enriched by creating and combining corresponding metadata so consumers have a complete understanding of the dataset. The quality and format of the dataset will be checked, edited and validated then transformed into as many of the most suitable formats as deemed appropriate for consumers. This may include transforming into Application Program Interface (APIs) as well.

### ***Load***

In the third stage the data will be uploaded with suitable cataloguing and tagging and grouped into a theme or category then made available on the open data portal for consumers to download and reuse.

To accomplish the Enrich/Transform and Load steps of this process a suitable cost effective proven open data platform with a broad range of capabilities will be required.

### **Platform Requirements**

The open data portal should allow data to be searchable using a variety of methods, selectable, have some limited analytic capability, be available in a variety of formats to download or print and present in the most user friendly, dynamic and visually pleasing way. This includes integrating charting and GIS mapping capabilities to give maximum visual impact. Users should be able to create their own requests for information filtering linked data in or out. The open data platform will be required to have the following functions to facilitate a streamlined process of publishing data for data administrators and finding, sharing and using data for data consumers:

### ***Hosting***

The council's IT service area have indicated a preference for the open data platform supplier to host the platform rather than have it hosted within the council IT infrastructure.

### ***Administration***

- Accept multiple data types and formats;
- Supports uploading and downloading of data in bulk;
- Data management system with catalogue of data set information;
- Monitor data harvesting and data management,
- Edit datasets;
- Create and or incorporate chosen metadata standard;
- Data review;
- Data set revision history;
- Approve datasets to be public or private;
- Statistics on dataset usage and downloads;
- Data storage of raw data in any format and metadata with hosting either on or off council premises;
- Tools to design, create and change the look and feel of the portal and its outputs

### **Consumer browse facilities**

- Fuzzy matching search and filter facility by keywords, tags, formats, licences, themes, metadata, sources or other selectable dataset indicators;
- Search via selectable geospatial area;
- Search via API;
- Visual data analytics with interactive tables, graphs, images and maps;
- Make user feedback comments;
- Make requests for datasets not already available

### **Consumer use**

- Download datasets in a variety of machine readable formats for free with no patent restrictions;
- Print capability

### **Control and Data Quality**

A number of robust publication processes, checks and roles will need to be defined, agreed and implemented before any open data publication takes place. In general terms these processes will revolve around the following data flow from selection of data to ongoing dataset maintenance after publication.














**Select:** Selecting data for publication will be based on a number of methods and criteria including:

- Review of council and partner organisation's public facing data, websites, announcements, online and print publications.
- Highlight data used in response to Freedom of Information requests as a priority for release.
- Use examples of open data sets already released by other public organisations as a guide for identifying similar Renfrewshire datasets to be made open.
- Review the results from the Scottish Government's Open Data dialogue app which was created to engage with the public and promote discussion via a number of different channels on what types of public sector data and format the public would like to see released as open data.
- Examine the reuse of other source public sector information that relates to Renfrewshire.
- Consult partners and other organisations in the local authority as to their participation as a data source to enhance the overall value of open data for Renfrewshire and as a data consumer.
- Engage with consumers to understand their needs.
- Review consumer dataset usage, feedback and requests.
- Recognise and assess a dataset's accessibility, completeness of metadata, quality, privacy and usefulness to consumers.



**Create:** The selected data will need to be organised so it can be made available for download, in bulk if possible, and in open and machine readable formats. These formats allow the data to be used and edited easily and allows for interoperability between different datasets. For example a PDF publication severely limits the consumer's ability to reuse the information. Open formats are non-proprietary and platform independent making them accessible by anyone without the requirement to access licensed software. E.g. Microsoft formats are not open as they use proprietary software. Machine readable formats allow a computer to read the data, is structured and easy to query using code. Examples of open data formats include:

Format Name	Definition	Type of data to use this for
<b>Comma Separated Values (CSV)</b> 	Comma Separated Values (CSV) is a great way of storing large amounts of data with just commas separating the data values. Often the CSV file will contain a header with names describing what data is populating the file.	Tabular data e.g. Use instead of Excel
<b>Tab-Separated Values (TSV)</b> 	TSV is a very common form of text file format for sharing tabular data and is highly machine readable.	Tabular data Use instead of Excel
<b>JavaScript Object Notation (JSON)</b> 	JSON uses human-readable text to transmit data objects consisting of attribute–value pairs. It is used primarily to transmit data between a server and web application, as an alternative to XML. The file size will be more compact or smaller than XML.	Complex structured data Multidimensional data Tabular
<b>Extensible Markup Language (XML)</b> 	XML is a widely known markup language that defines a set of rules for encoding documents in a format that is both humanreadable and machine-readable. Users create and define their own tags.	Complex Structured data Multidimensional data Tabular data e.g. database extract metadata
<b>Rich Site Summary (RSS)</b> 	RSS (originally RDF Site Summary), often dubbed Really Simple Syndication, uses a family of standard web feed formats to publish frequently updated information: blog entries, news headlines, audio, video. An RSS document (called "feed", "web feed" or "channel") includes full or summarised text, and metadata, like publishing date and author's name.	Use for announcements or events e.g. on websites
<b>ATOM</b> 	The Atom Syndication Format is an XML language used for web feeds. The Atom format was developed as an alternative to RSS. Note RSS is the preferred standard.	Use for announcements or events e.g. on websites
<b>Open Document Format for Office Applications (ODF)</b> 	The Open Document Format for Office Applications (ODF), also known as Open Document, is an XML-based file format for spreadsheets, charts, presentations and word processing documents. It was developed with the aim of providing an open XML-based file format specification for office applications.	Non-system generated metadata or additional information you release with your dataset. (replaces Excel, Word, PDF)
<b>HTML</b> 	Used for formatting information on the web	Non-system generated metadata or additional information you release (replaces PDF, Word)

<p>Keyhole Markup Language (KML)</p> 	<p>KML is an XML language focused on geographic visualization, including annotation of maps and images.</p>	<p>Spatial/location data</p>
<p>Geography Markup Language (GML)</p> 	<p>GML is the XML grammar defined by the Open Geospatial Consortium (OGC) to express geographical features. GML serves as a modelling language for geographic systems as well as an open interchange format for geographic transactions on the internet.</p>	<p>Spatial/location data</p>
<p>GeoJson</p> 	<p>GeoJSON is an open standard format for encoding collections of simple geographical features along with their nonspatial attributes using JavaScript Object Notation.</p>	<p>Spatial/location data</p>

Creating a dataset will also include capturing metadata for every dataset. This is compulsory as no dataset will be published without metadata. An minimum metadata standard will be introduced at first with the intention to progress to a standard metadata format as the open data portal and the whole open data process matures. Those responsible for the data will also be responsible for ensuring corresponding metadata is current. Liaison between data administrators and the open data portal administrators will be ongoing to ensure metadata accompanying datasets is kept up to date. Information to comply with the Dublin Core metadata standard is shown below:

Dublin Core Requirement	Description
<i>Contributor</i>	An entity responsible for making contributions to the resource
<i>Coverage</i>	The spatial or temporal topic of the resource, the spatial applicability of resource, or the jurisdiction under which the resource is relevant
<i>Creator</i>	An entity primarily responsible for making the resource
<i>Date</i>	A point or period of time associated with an event in the lifecycle of the resource
<i>Description</i>	An account of resource
<i>Format</i>	The file format, physical medium, or dimensions of the resource
<i>Identifier</i>	An unambiguous reference to the resource within a given context
<i>Language</i>	A language of the resource
<i>Publisher</i>	An entity responsible for making the resource available
<i>Relation</i>	A related source
<i>Rights</i>	Information about rights held in and over the resource
<i>Source</i>	A related source from which the described resource is derived
<i>Subject</i>	The topic of the resource
<i>Title</i>	A name given to the resource
<i>Type</i>	The nature or genre of the resource



Another key requirement to creating a dataset will be the application of an open license. This is another compulsory requirement before a dataset is published. When applying an open licence thought must be given on what is hoped to be achieved by releasing the dataset. Requiring attribution will help promote Renfrewshire's open data initiative as consumers will have to link back to the original work. Share a-like restrictions will require users of the data to publish their work openly. This may deter commercial businesses and people who want to profit from their use of the data, resulting in reduced innovation and use. However, if a dataset about Renfrewshire has been derived elsewhere under a Share a-like licence than that dataset must be published under the same licence as the original data. Whilst it is possible to create a unique licence, guidance from Scottish Government is to use a standard re-usable licence as they provide greater recognition amongst users, increased interoperability due to the use of standard terms and increased user compliance.

**Edit:** When checking the quality of data and making any edits it must be remembered there is no such thing as perfect data and imperfections should not deter data release. Any limitations or caveats about the data should be made explicit. These will help any re-user to understand the limitations of the data. Re-users will have the opportunity to provide feedback on data quality and mistakes which will help improve the quality of the data.

**Validate:** The open data team will define and implement a robust set of processes or checks for all datasets to go through before they are published. These checks will include:

- Check legality and information governance rules of the data
- Verify source
- Identify relevant data owner
- Confirm responsibility for ongoing maintenance of dataset
- Check contextual metadata is completed to an acceptable standard
- Prepare as many open formats of the data as possible
- Check completeness and quality of the data
- Consider expected rate of change
- Validate machine readability
- Confirm any spatial elements are INSPIRE compliant
- Approval to publish

**Publish:** Validated datasets will be released and made available to the public on the acquired open data platform ensuring any limitations or caveats on the quality of data accompany the dataset. Appropriate data tagging, topic linking and updating of the information management catalogue will be applied. A check of the success of the release of datasets will be made and if and when required the release of data will be publicised.

**Maintain:** Keeping data current on the open data portal will be a challenge particularly as the number of published datasets grows. However there are some options that depending on the dataset and source will be considered and used including:

- Automated extract and upload when data changes at source – ideal solution.
- Schedule dataset extracts driven by the expected frequency of change at the source
- Notifications of data changes triggers a dataset extract.

**Roles:** The Data Analytics and Research team has been tasked with creating and implementing the council and partner's open data strategy in line with Scottish Government directives. Specific individual roles within this team will need to be defined after a portal platform is acquired and the process becomes more defined, but it is expected that those in the Data Analytics and Research team will require training and be involved with most of the data flow processes in the transformation and loading of datasets.

As already mentioned in the [Open Data Management](#) and [Broad Technical Architecture](#), the ICT and Information Governance teams as well as a proactive network of Open Data advocates representing partners and all service areas will be critical to the success of the strategy. Specific duties are still to be identified but it's envisaged the ICT team will be extracting datasets and the Information Management team will have more of an advisory/governance role while the advocates will be the conduit between partners, service area dataset owners, the Data Analytics and Research, ICT and Information Governance teams.

### **Cost of Opening Data**

The bulk of open data about Renfrewshire will come from existing internal data held by the council and partners. Therefore apart from staff resources involved with preparation, releasing, updating and maintenance of data, costs are expected to be limited to generic costs such as:

- Platform software;
- Design and or build of platform;
- Web hosting of the portal;
- Training of staff;
- Additional visualisation tools;
- Ongoing support, maintenance and software updating of the portal; and
- Advertising and promotion of the portal e.g. conducting hackathon events.

The principal of open data is data which is available for free. This allows equal access to the data and allows it to be widely used and re-used. Any data which requires a fee to access cannot be considered true open data. Therefore recuperating costs for providing open data will not be an option.

However having an open data portal does provide a single access point to make other data available which may not necessarily be considered open data though in the vast majority of cases the data was captured or created using public funds and should be made accessible to all for re-use. Nevertheless, if considering charging for a specific dataset, entitlement to charge should be checked against existing access to information legislation. Whether the request for data is fulfilled will depend on the nature of the data and the request. This could possibly provide some limited revenue.

## **Platform Options**

The open data portal will be a website, accessed independently or via the existing council or partner websites, where datasets can be uploaded and updated and the public can search for and download datasets. To ensure the long term sustainability of any chosen platform it will need to be consistent with the council and partner's IT and Digital Strategies and conform with the following principles:

- **Ease of use** – provide a simple, accessible and responsive interface to allow access from various devices such as PCs, tablets and mobile phones;
- **Secure** – be compliant with the Council's web security policies;
- **Employs open standards** – this removes access barriers and guarantees interoperability and integration to facilitate the reuse of data;
- **Scalable** – to future proof for increases in data volumes;
- **Cost effective** – using an open source system and software coupled with virtual machines reduces the licensing costs and increases the cost effectiveness of the platform.

There are essentially three options to obtain the platform:

1. Build own platform in house
2. Buy an existing platform
3. Use an existing open source platform

Option 1, build a platform in house is not the most efficient way to spend council or partner money and use resources particularly when there are several very good existing platforms available and in use by many UK government organisations. Buying an existing platform, option 2, is not the most cost effective way of procuring an open data platform and therefore option 3, using an existing open source platform, is the preferred option to procure the open data platform.

Some examples of candidate open source platforms used by other local authority areas are listed below and should be appraised for suitability:

- CKAN
- InstantAtlas
- Swirrl

## **Design and Build of Platform**

While selecting the open data platform, consideration of the design and build will also need to be considered. Design and build could possibly be undertaken by partners, depending on the skill and expertise of those getting involved, or within the council, or by also working with one of the many service professional organisations listed in the government's digital market place website: [www.digitalmarketplace.service.gov.uk/g-cloud](http://www.digitalmarketplace.service.gov.uk/g-cloud)

These service professional organisations have a successful track record of constructing, introducing and mentoring open data portals for local government authorities using a variety of open source data platforms.

## TRAINING

Scottish Government has procured a supplier of open data courses which will be delivered across Scotland over 12 months from November 2015 at no cost to attendees. Unfortunately at this stage this training is only for public service organisations. The training consists of a 1 day introduction course and a 2 day extended course which includes all of the content from the 1 day course but extends the learning to help organisations implement the principles and practice of open data in line with the Scottish Government strategy. Dates and bookings to attend courses are at: [www.scotopendata.eventbrite.co.uk](http://www.scotopendata.eventbrite.co.uk)

1 Day Introduction Course	2 Day Extended Course
Open Data – Introduction, context and commitment	All topics covered during 1 day intro course
Scottish Government Open Data Strategy 2015	Open Data Toolkit <ul style="list-style-type: none"> <li>• Open Data manifesto &amp; prioritisation strategy</li> <li>• Barrier identification * data compliance</li> <li>• Target setting &amp; realising the benefits</li> </ul>
The opportunity for the public sector in Scotland	Open Data Engagement <ul style="list-style-type: none"> <li>• Creating compelling, shareable data stories</li> <li>• Introduction to data visualisations</li> <li>• Creating data communities &amp; building digital literacy</li> </ul>
Help, support and the Open Data Toolkit	Open Data Technologies looking at different options for the publication of Open Data
The business case: making the most of the benefits	
Engagement and Open Data	
Leading others into open data – leadership and cultural change	

There will of course be bespoke training or at least awareness sessions required for those participating or who have a role to play within the open data strategy to operate the open data portal acquired for Renfrewshire and on the data flow process from the data source to publishing as an open data set.

## INITIAL ACTION PLAN

A table outlining the initial general actions to start the open data vision for Renfrewshire is shown below. No time limiting dates have been set other than the overall timeframe of publishing an initial open data publication plan by the end of 2015 and start publishing open data by January 2017. Some tasks will be ongoing past the publishing start date and some tasks that have no affect on the start may begin after as they are more pertinent to the ongoing use of the open data portal. Of particular note is an open data ‘Enlightenment’ event to be held in January 2016. Partners will be invited to find out more on the open data phenomenon and how they can get involved in Renfrewshire’s open data activities and be part of an open data network for Renfrewshire. The initial plan includes building of a suitable application using Renfrewshire open data which will be used to demonstrate the potential benefits of open data to other partners and service areas within the council.

INITIAL ACTION PLAN		
Now - 2015	2016	2017 and future
<ul style="list-style-type: none"> <li>• Get approval from CMT to progress OD in Renfrewshire</li> <li>• Raise OD awareness</li> <li>• Write initial OD strategy</li> <li>• Publish initial OD dataset publication plan as per SG guidance</li> <li>• Present initial OD strategy to CPP &amp; CMT</li> </ul>	<ul style="list-style-type: none"> <li>• Conduct OD 'Enlightenment' event</li> <li>• Establish Partnership OD advocates network</li> <li>• Identify potential skill &amp; expertise contributors</li> <li>• Establish consumer requirements</li> <li>• Identify potential dataset providers</li> <li>• Establish roles for data flow process</li> <li>• Identify platform candidates</li> <li>• Agree &amp; acquire platform</li> <li>• Identify training needs                             <ul style="list-style-type: none"> <li>○ Dataset owners</li> <li>○ Advocates</li> <li>○ Data Analytics &amp; Research team</li> <li>○ ICT</li> </ul> </li> <li>• Agree look &amp; feel and OD themes</li> <li>• Create &amp; trial beta portal</li> <li>• Conduct Renfrewshire OD 'Hackathon' event                             <ul style="list-style-type: none"> <li>○ For web developers &amp; programmers for apps</li> </ul> </li> <li>• Identify &amp; build a suitable app using open data</li> <li>• Review dataset publishing priorities</li> <li>• Review governance</li> <li>• Set future dataset targets</li> <li>• Release portal to the public</li> </ul>	<ul style="list-style-type: none"> <li>• Regular publishing of OD</li> <li>• Embed 'open by default' culture</li> <li>• Monitor consumer usage</li> <li>• Monitor consumer feedback</li> <li>• Review dataset priorities &amp; targets</li> <li>• Refine OD processes &amp; governance</li> <li>• Enhance portal to include other functions e.g. create bespoke user dashboards, access to apps</li> </ul>