Local Government Open Data Breakthrough Projects Evaluation Report 2013/14

Project title: *Cambridgeshire Insight Open Data: Stimulating Economic Growth and Encouraging Innovation*

Lead organisation: *Cambridgeshire County Council*

Submission date: *29/08/14*

Type of project: (Data release, standards, platform/tool development)

- *Community engagement, data scoping and release of data;*
- *Platform and tool development (DKAN)*
- *Data Standardisation and development of schemas*

Total grant: *£65,000*

Summary of the project:
Cambridgeshire's position as the fastest growing county in England and Wales was reconfirmed by the 2011 census. Against this backdrop of rapid growth there were unique local datasets, pertaining to both the community and economy, that were available to select government organizations but are not currently being published in an open and transparent manner.

Cambridgeshire Insight is an established partnership initiative (hosted by Cambridgeshire County Council) that has already created a single access point for all written research material about the county. The proposed 'Open Data' project builds on this success by developing a new, open methodology that will unlock these hitherto unpublished datasets.

The release of this data will enable local companies, community groups and individuals to plan and then respond to the challenges created by rapid growth alongside local government.

Challenge/opportunity:
- *The Cambridgeshire Insight partnership is committed to being a transparent, accountable partnership and since inception one of our main goals has been to increase the amount of data and insights published for reuse by others.*
- *Stimulate economic growth and encourage innovation. The emphasis is that Cambridgeshire is the fastest growing county in the country and with this growth there needs to be efficient planned services and development of ideas to allow this growth to become sustainable and universally beneficial. By releasing data we increase the area's capacity for analysis at a time when resources for this work are being cut back within government organisations.*
- *Making previously unavailable data open and adding extra value to process methodology for local specialised data enabling greater use and accessibility for better decision making.*
- *Integration with other local and national data providers (making data available through data.gov.uk).*
- *Breaking through traditional barriers to development by building a progressive partnership to develop these products and promote local useable and understandable open data.*

- *Promote open, meaningful relationships and linkages between local public sector organisations, local neighbourhoods, local economic partnerships and business leaders in order to promote Cambridgeshire, to increase Cambridgeshire's potential and to improve the local economy for the people of Cambridgeshire.*

Data published and uses:
Over the course of December 2013 – February 2014 we held a series of consultation exercises including:
- *Online Survey*
  (The initial online survey formed the 'root' for the follow on exercises.)
- *Focus groups and workshops*
- *Attendance at local data groups*
- *Online seminars*
- *Discussion groups online such as LinkedIn and Knowledge Hub*
- *E-mail and telephone discussions with key contacts, for example with OKN (Open Knowledge Foundation) members, Connecting Cambridgeshire and others.*

The consultation process was used to prioritise the datasets that we released:
- *Local Housing Allowance (LHA) Changes*
- *Employment in the hi-tech community: Cambridgeshire & Peterborough 2012*
- *Cambridgeshire Population and Dwelling Stock Estimates 2012*
- *Cambridgeshire Population Forecasts 2012*
- *Cambridgeshire Dwelling Stock Forecasts 2012*
- *Affordable Housing Completions*
- *Housing Completions*
- *Traffic count data (2009 to 2013)*
- *East of England Forecasting Model 2013 Baseline for Cambridgeshire*
- *Housing need register applicants parish preference, social rented lettings and bidding behaviour.*
- *Count of Crime in Cambridgeshire by financial year*
(In some cases this involved a release of multiple tables associated with each dataset)

More detail is include in the attached spreadsheet and the datasets are also listed at http://data.gov.uk/data/search?q=cambridgeshire&publisher=cambridgeshire-insight

Approach to publishing data
- What approach is used to publish data:
  For the project we developed a specific release platform http://opendata.cambridgeshireinsight.org.uk/ (The platform is in a DKAN format). Data is published under five main headings 'Transport', Population/ Demographics, Planning, Economy and Housing. These were defined and prioritised during consultation with stake-holders. Data was also published / linked on data.gov.uk

- How data users assess the availability and use of data:
  (How have users found out about the service?, How has the authority promoted the publication of the data?, Have users found it easy to download datasets and use the datasets?)
  User have found out about the data initially through the consultation process. Since that point the data has been publicised via Cambridgeshire County Council's and Cambridgeshire Insight's social media channels. We have also

begun a series of Blogs highlighting the data and its use. We've also made specific links with some organisations that might like to use the data produced e.g. Cambridges University in regard to the hi-tech dataset release.

The Cambridgeshire Insight Open Data interface has been tested by users and they have found the download process to be straightforward. Data is available in a .csv format as well as JSON and RDF. In addition to downloads the site also enables the user to explore the data. The data store can also be accessed via web APIs to allow for integration across different sites / applications.

- Is the contextual information such as metadata provided with the data of use? Whether data systems are sufficiently robust and sustainable. Are there any plans for further monitoring, developing and improving the system?

  Contextual data is provided with data releases such as the metadata and the taxonomy of the dataset. The comprehensiveness of this provision is variable depending on the dataset released. The data APIs are available to allow for remote linking / integration across different sites / applications.

  We have had discussions about the standards that we would wish to implement for metadata / taxonomy. There is a demand from our publishing community to have a 'simple' publishing procedure through to a more complex methodology where the dataset demands it.

Immediate Benefits and Long Term Impact

The project has brought about some immediate benefits for Cambridgeshire County Council and also for other partner organisations involved in the data release programme.

- A step change in the standard of data released and the methodology for data release has been achieved. Through the process of setting minimum standards for Open Data, and working through initial data release for the 10 agreed datasets we've established a methodology through which we are now able to improving pre-existing data e.g. the statutory transparency datasets. With proved that releasing data to a good standard can be straight forward. We've also been able to demystify some of the terminology around the approach.

- We have released datasets for the first time. For example we have collected data on the Cambridge Phenomenon (the cluster of hi-tech / bio-tech firms in and around Cambridge) for a significant period of time. Up until now this has been used to generate a bi-annual report and has been only shared with Cambridge University for economic research purposes. We've now shared an open dataset based on this work for the first time.

- We've raised the profile for Open Data in the County. The award and the activity around the project has served to considerably improve awareness and has led to fruitful discussion around a shared partnership agenda for data release that has attracted funding from the Cabinet Office 'release of data' fund.

- There has been an improvement in the knowledge skills that a number of people possess as additional people have been brought into the project to

complete it. This knowledge / skill is now being shared around the rest of the organisation.

Longer term impacts on the publication of data cannot yet be fully measured as data was released towards the end of the project. Initially releases of data have been largely linked to the theme of improving the County's economy and over the next six months we will continue to monitor the use of the data that has been released and identify case studies where its use by third parties has proved beneficial.

Lesson Learnt

There have been some fair straight forward lessons learnt during the project.

- The County Council's existing approach to Open Data was poor. One of the benefits of the project has been a step change in our knowledge and practice. This in turn has cast our approach to the release of statutory transparency datasets in a poor light. These are currently released in a non-standardised way, not beyond one star standard with a number of broken links on both the main County Council pages and on data.gov.uk. Now this has been highlighted work is underway to improve our approach across the board.

- We've learnt that within our local authority the ''leadership' for open data was vague. Centrally the Information Governance (IG) team have ownership for ensuring that the authority complies with legislation to publish certain transparency datasets but we've struggled to move other services beyond these to think more widely about the opportunities that publishing open data can provide. The establishment of this project alongside other initiatives has provided us with the opportunities to evaluate where we are and now the renewal of an Open Data Strategy is underway and a more robust Open Data Strategy Group is meeting regularly.

- The current toolset available for open data development within our particular setting was very limited and limiting to the project. We operate in an MS Office driven corporate IT environment. There are an extensive range of tools available https://github.com/theodi/open-data-tech-review/wiki to support open data work but these aren't generally available to users without a deal of organisational compliance regulations etc. that need to be overcome. Examples tools that would prove useful would be format editors for xml, JSON and GeoJSON. We've since initiated a 'gateway' review to corporately identify the most suitable (and compliant toolset).

Further information:

Contact Michael.Soper@Cambridgeshire.Gov.uk
01223 715312
http://opendata.cambridgeshireinsight.org.uk/