CLEAR Info Progress Report
Document Control Sheet

Project Reference	LIFE10 ENV/UK/000175 CLEAR Info
Action number	5.2
Description	Report on the CLEAR Info hackathon
Version	FINAL
Contents	1.0 Introduction 1.1 Understanding the procurement route 1.2 Setting up the hackathon 1.3 Data 1.4 Outputs 1.5 Evaluation 1.6 Post event communication 1.7 Annexes
Written By	Katie Pullen
Project Manager Sign-off & Date	Suzanne Laidlaw
Project Executive Sign off & Date	Gillian Pratt

Report on EnvHack2- the CLEAR hack

1.0 Introduction

Part of the bid for the CLEAR Info project included an action to hold a workshop and produce practical guidance on how to integrate environmental data. In the midterm report submitted on 1 March 2014, the team proposed holding a hack day, with a broad cross section of data users.

A 'hack day' or 'hackathon' is an event in which computer programmers and others involved in software development, including graphic designers, interface designers and project managers, collaborate intensively on projects. As this was a relatively new and exciting approach for EU LIFE+ and the CLEAR Info project, the project team secured support from colleagues in the Environment Agency with experience in running this type of event.

The Environment Agency has previously used hackathons as a way to engage in collaborative computer programming to produce innovative ways of using Environment Agency data sets. Colleagues in the innovation team were responsible for organising the hackathons and provided the CLEAR Info team with expertise and guidance. The innovation team advised the CLEAR Info team of potential external suppliers to run the hackathon as this is an approach that was used previously.

1.1 Understanding the procurement route

The CLEAR Info team contacted the relevant procurement colleagues to understand the process for using an external supplier to run a hackathon and the rules around procurement for this type of work.

A colleague in the Corporate Information Services procurement team investigated which framework some known hackathon suppliers were on and advised the team that there were only two suppliers listed on the relevant framework that could supply this service.

The team were advised that as this work was expected to cost over £10,000 it needed to be advertised on the government's contract finder website (as part of the government's

transparency agenda).

In a series of discussions with colleagues in the innovation team it was decided that the CLEAR Info team working with Environment Agency colleagues had the resources and experience internally to organise and run the hackathon weekend, and could offer better value for money than procuring an external supplier.

1.2 Setting up the hackathon

An initial meeting was held on 22 January 2014 with colleagues in the innovation team to discuss:

→ marketing
→ preparation and how to arrange a hackathon
→ what works/what doesn't work
→ what we would need to provide
→ identifying participants and how to contact them

A follow up meeting was held on 7 March 2014 where a date for the hackathon was agreed and roles and responsibilities were decided and shared out amongst the team. The hackathon required tasks to be carried out in three main themes; practical, data and engagement. The team worked through the tasks identified initially and include additional tasks that the team felt would improve the hackathon such as creating 'hackathon packs' for attendees.

On 12 June a final preparation meeting was held, where the team ran through the timetable for the day and who would be responsible for the tasks during the weekend.

Practical

Internal colleagues in the innovation team arranged the venue and catering for the hackathon using contacts that had been built up from previous hackathon events.

The correct Environment Agency procedures for hiring the venue, providing catering and purchasing the prizes were completed in the run up to the event. The team ensured approval forms were signed off correctly and that all Environment Agency procurement rules were followed.

The innovation team recommended that the hackathon included prizes for products or ideas for certain categories. This recommendation came from previous experience running hackathons and is used as a way to encourage potential hackers to sign up for the event. The CLEAR Info team worked with the innovation team to establish the type of prizes needed, these were sustainability themed gadgets. Criteria were established for awarding prizes to hackers for:

→ best visualisation
→ best technical hack
→ best use of CLEAR data
→ best mash up using datasets from more than one source

→ best potential outcome for the environment

As the hackathon was held over a weekend it was important to ensure that enough staff were available to help coordinate the running of the event and with technical knowledge of the data. The CLEAR Info team drew up a list of staff needed and confirmed their availability for the hackathon. Staff availability regarding preparing the data and data licence also needed to be taken into account in the run up to the hackathon. This proved to a vital part of planning the hackathon as a large number of colleagues from around the Environment Agency were involved in preparing the data.

Engagement

Following the innovation teams advice on using the established 'Envhack' brand, the hackathon was named Envhack2: the CLEARINFO hack. The team created a dedicated website (http://www.envhack.com/) and twitter account (@envhack2) for the hackathon. Twitter was used as the main communication route to encourage hackers to sign up to the hackathon and a series of messages were tweeted in the run up to the hackathon. A TweetReach report indicated that by 23 June 2014 the tweets had reached 37,114 accounts.

The team also used the Environment Agency weekly buzz to publicise the event, encouraging them to share with their networks and kept the CLEAR Info website up to date at all times (www.gov.uk/ea/clear-info).

Eventbrite allows event organizers to plan, set up ticket sales and promote events. The team set up an eventbrite page for the hackathon, allowing the project team to monitor how many tickets were requested and to adjust the promotion of the hackathon and the resources needed for the weekend (e.g. catering, staff and printed materials). Screenshots from the website, twitter account, Environment Agency weekly buzz and eventbrite are contained in Annex 1.6.1.

Using the innovation team's experience of previous hackathons, the CLEAR Info team were advised to created personas which would describe a fictional person working in a field related to the data available for the hack and a particular issue they wanted to resolve. The purpose of the personas was to guide the hackers to a practical solution using the CLEAR Info data and explore what could be done with the data in the context of CLEAR Info.

While developing the personas for the hackathon the CLEAR Info team recognised that a 'hack pack' would be a useful addition to the materials created for the hackathon. The hack pack consisted of a schedule for the weekend, key information about CLEAR Info, information about the Environment Agency and Open Data, contact numbers and the details of the personas. A hack pack was created for organisers which included the same information plus a series of evaluation questions to ask hackers during the weekend (Annex 1.6.2).

1.3 Data

One of the biggest challenges of the hackathon was organising the data needed for the event to run successfully and encourage attendees. This required a substantial amount of

effort from the team. Nine Environment Agency datasets needed to be agreed for use by the data custodians and prepared for the hackathon. Eight of the data sets were at site level and the CLEAR Info investor data set is at parent company level. The CLEAR Info team also indicated to the hackers additional open data sets that maybe useful, including air quality England data and local authority data.

Preparing the data

Working in conjunction with the technical working group, extracts of data were taken from the cube for use in the hackathon. Significant input was required from the project team designing which data sets and fields might be of interest and quality checking extracts.

Data Licence

A temporary data licence specifically for the hackathon was prepared; this was developed with the project team, legal department, the Data and Information team and several data custodians. The extensive work preparing the data licence enabled the maximum amount of data to be shared with the hackers in a controlled environment.

Data descriptions

Descriptions of the data had been compiled as part of the Approval for Access process, to use the data descriptions for the hackathon they needed to be re-worded into plain English so that hackers could understand the data. This required substantial effort as each data set required a simple description and each attribute name required an attribute description.

1.4 Outputs

On the first day of the hackathon, hackers looked at the data and started to formulate ideas based on the personas that were provided. The persona of 'Gregory Gaia' was looked at by one of the teams, this persona outlined the issues faced by a food manufacturing company that want to understand if their suppliers operated to the same responsible ethos that the company upholds. This required the hackers to use the CLEAR Info dataset to create an online tool for assessing the environmental risk of a business in a supply chain. This persona addressed the CLEAR objectives of improving compliance or other EU legislation or giving us learning about data integration. The hackers felt that by aggregating the metrics at all levels of the hierarchy, users can get an assessment of the company they are interested in without having to drill down. Data users or companies investigating their supply chain are often interested in a midlevel company rather then at site or parent company level. Having the full hierarchy to interrogate had investigations at individual levels / companies possible. Often Parent Companies might have very different assets in several different sectors and more diverse impacts / breaches then the company of concern.

In the afternoon, the hackers presented an outline of their ideas and formed into two teams to start to create products. Photographs from the event are contained in Annex 1.6.3.

With guidance from the CLEAR Info data specialists, the teams created multiple ideas and solutions including:

→ Application Programming Interface (API) to publish the CLEAR Info data. An API can be used to ease the work of programming graphical user interface components, to allow integration of new features into existing applications (a so-called "plug-in API"), or to share data between otherwise distinct applications
→ Query tool & visualisation for assessing the environmental risk of a business in a supply chain.
→ Heat map visualisation of permits.
→ Sewerage consents & incidents visualisations.
→ Start on a river network analysis to look at what's happening upstream.
→ Two concepts for exploring environmental risk - one for a local authority area and the other being "OPRAH does OPRA" with a riskometer.
→ An unfinished prototype for using Cucumber-style behavioural test logic to analyse environmental performance.
→ A code which translates data into different icons depending upon the values in a dataset, and when linked with grid reference, can be loaded into any mapping tool that supports KML files.

Learning was also captured using hackpad, this is a simple tool to capture, organize and share knowledge and is commonly used at hackathons to collaborate on projects.

1.5 Evaluation

The team evaluated the practical running of the hackathon and how the hackers used the data.

Practical Evaluation

Logistics

- Making the start time later in the day would be more realistic for an event happening on the weekend.
- Having a member of staff with programming experience worked well. It would be useful to have more regulatory specialists or produce a ‘cheat sheet’ on the data for other staff members helping the participant.
- The team was advised that most people who attend hackathons have had previous experience in this type of event. This was not the case and it may have been useful to have more guidance to the participants on what happens during a hack (including how to get into teams). It would also have been useful to have options for participants to choose if they are a coder, developer or ideas person when they sign up for a ticket.

Data Logistics

- The feedback from the hackers was that they used the personas and found them useful for shaping ideas and helped with the understanding of the data providing them with a framework of what to do with the data available.
- Data descriptions provided before the hackathon would have been useful.
- The way the hackpad was set up proved to

be difficult for the hackers to find and navigate around.

- Provide some information on how the datasets connect together to help the hackers link up data.
- Data experts that are available 'on call' would be useful if any questions come up that are not able to be answered by the staff in attendance.
- With more time we would have presented the data in a relational data store rather than .csv format.

Data Evaluation

Prior to the hackathon the CLEAR Info team developed a series of evaluation questions to help collate feedback from the hackers. These were included in the organiser hackpacks (Annex 1.6.2). The feedback from the hackers was varied and showed that it would have been useful to have the data in a different format (such as CSV or an API over the data) as processing the data took a long time. Overall, the hackers found the data to be useful but there were some issues with data either missing or not containing common content (e.g. northings and eastings), making it difficult to link across certain datasets. Further information is contained in Annex 1.6.5.

1.6 Post event communications

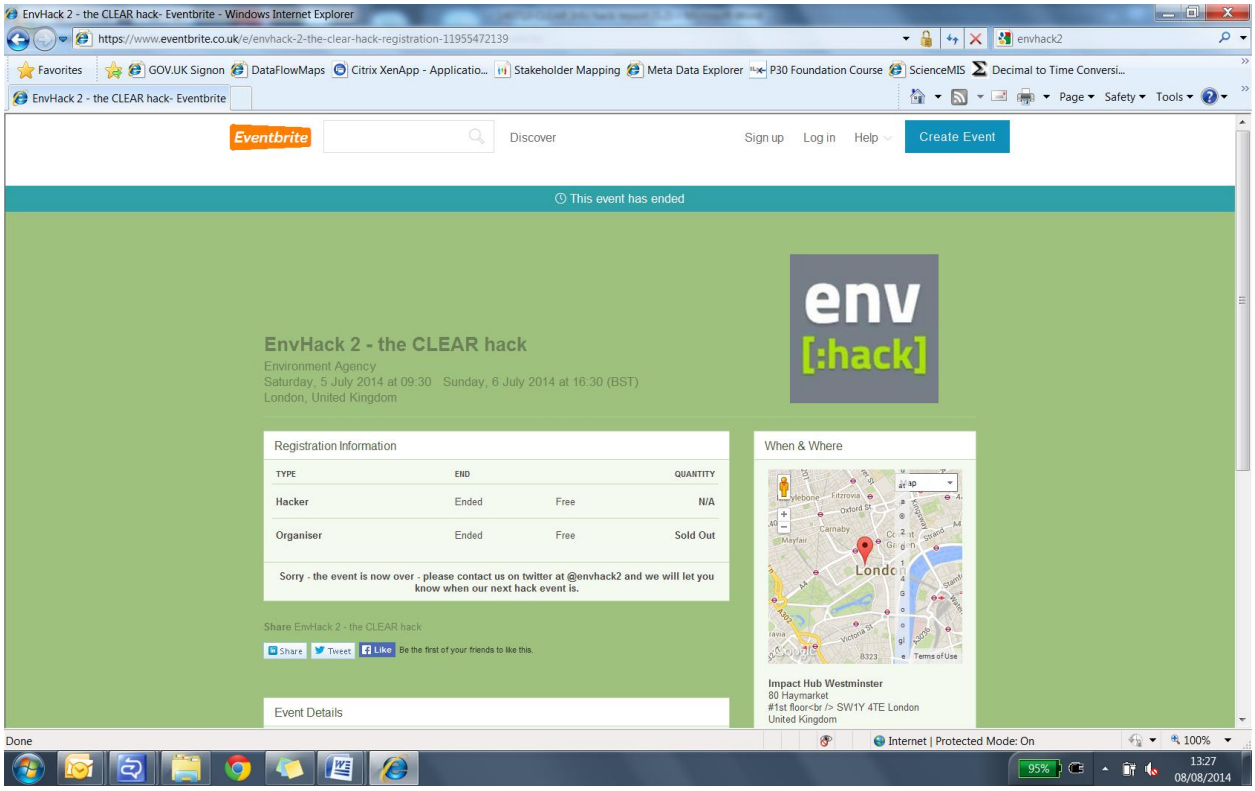
Following the event, the team updated the CLEAR Info webpages (www.gov.uk/ea/clear-info) and issued communications through weekly buzz within the Environment Agency (see Annex 1.7.5). The team felt that the valuable lessons learnt from organising a hackathon event were worth sharing with colleagues, and organised a livemeeting on XXXX which XX colleagues attended. Copies of the weekly buzz article and presentation delivered are contained in annex 1.7.5.

1.7 Annex

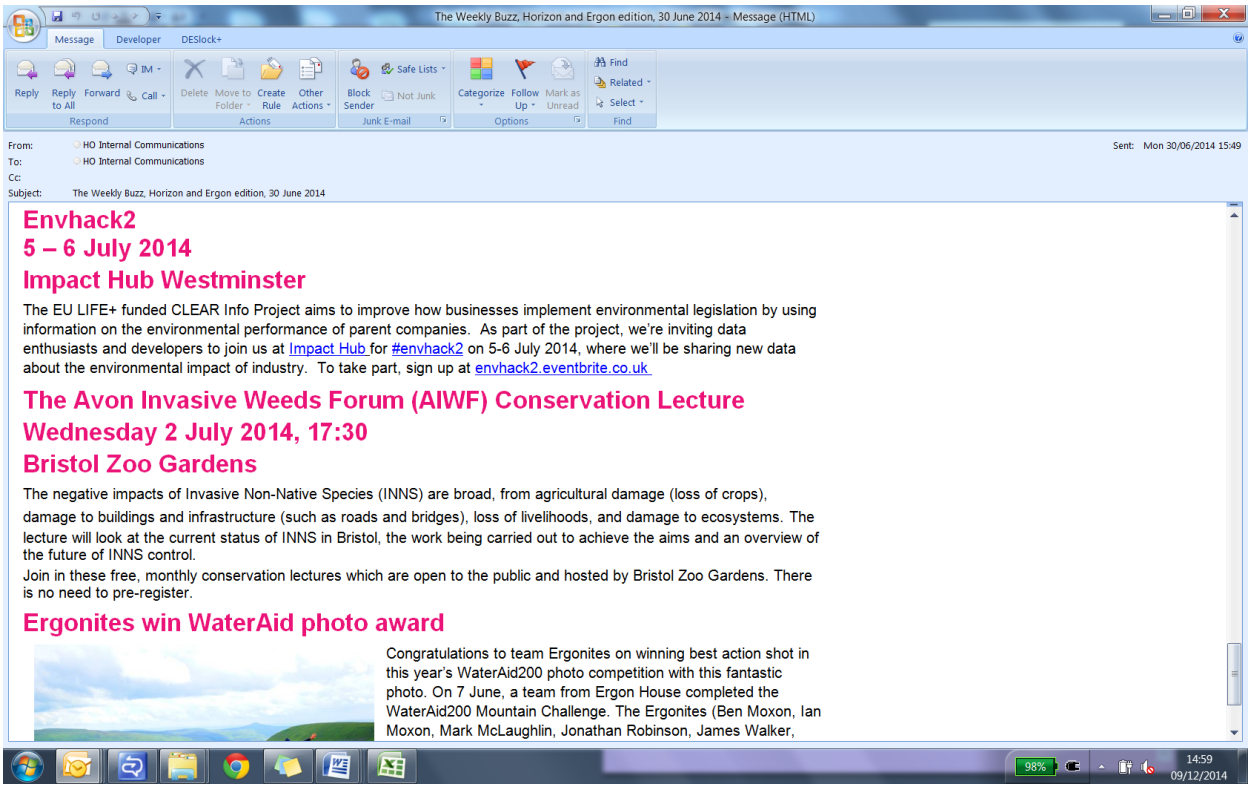
1.7.1 Publicity materials

Screenshot of www.envhack.com

Screenshot of twitter @envhack2



Screenshot of Eventbrite page



Screenshot of weekly buzz article

1.7.2 Hackpack

ENVHACK 2 - THE CLEARINFO HACK

5th-6th July 2014

Explore new data about the environmental impact of industry with the Environment Agency

Impact Hub Westminster

Central London

HACK PACK– ORGANISER

Contents

Schedule for the weekend

SATURDAY

9.30am Arrive & Coffee

10.00am Intro to the event

11.00am Idea time

11.30am Pitching & formation of groups

12:00pm Lunch

12.30pm Start to hack

6.00pm Dinner

8.00pm We're knocking off but you're welcome to carry on!

SUNDAY

9.00am Breakfast

9.30am Back to Hacking! (with lunch on the run)

13:00pm Hackers present ideas

3.00pm Presentations & Prizes

4.00pm We all go home happy

CLEAR Info

CLEAR Info is an EU funded project involving leaders across Europe in sustainable investment, business information and environmental data and regulation.

We've been joining regulatory data with business information to identify environmental risks and opportunities. By transforming disparate site and subsidiary data into influential company information, we can demonstrate how data can be used as a powerful tool to encourage businesses to improve their environmental compliance.

To date, we've been working with regulators, investors and regulated businesses. Your involvement at Env:hack2 will help us:

→ Prototype solutions that use the CLEAR data, environmental regulation and other datasets

→ Explore innovative uses of the data.

→ Gather feedback on the quality, format and access to the data.

→ Identify measures which could be taken across Europe to increase data sharing.

The findings from this weekend will be used to inform the open data agenda in Government, and will be shared with the Environment Agency, other environmental regulators and the European Commission, to influence and inform how we share data across Europe.

The Environment Agency's approach to 'open data'

The Environment Agency is on an open data journey. In practice, this will take time as each dataset needs to be cleaned, checked for Intellectual Property Rights & published in an open format. This means the Environment Agency will be issuing open datasets progressively over the next 2-3 years.

So, some of the data we're presenting at this event is really "open" and some is on the list to be opened.

We're looking for individuals and organisations to join our new "open data steering group" which will set the priorities for releasing Environment Agency data.

If you'd like to join us on this journey, speak to one of the team.

Contact Numbers

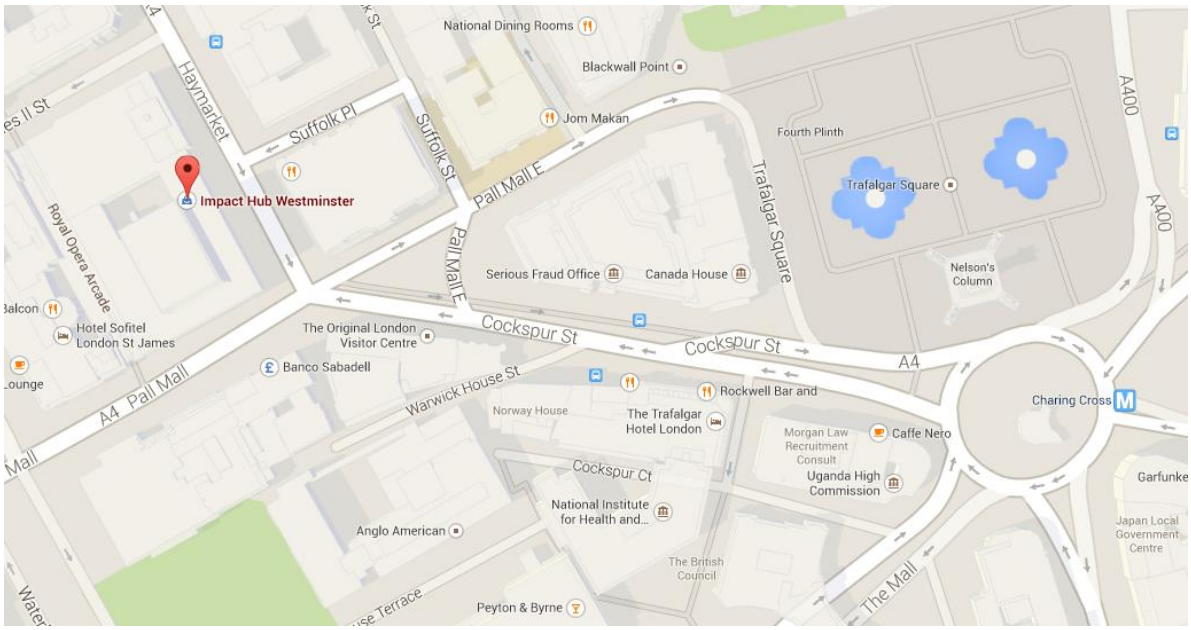
Who?	Sat	Sun	Name	Number
Impact Hub Westminster Event Organiser	X	X	Scherri	
The Peoples Supermarket (Catering)	x	x	Jackie	
EA CLEAR Info Team	X	X	Suzanne Laidlaw	
EA CLEAR Info Team	X	X	Katie Hughes	
EA CLEAR Info Team	X	X	Katie Pullen	
EA Data Team	X	X	Adam Redpath	
EA Data Team	X		Neil Gillan	
EA Data Team	X	X	Alan Cruikshanks	
EA Data Team			James Procter	
EA Hack team	X	X	Simon Redding	
EA Hack team	X	X	Paula Nickson	
Trucost	X	X	David Hobson	
Southampton University		X	Ian Williams	

Other useful information

Venue: Impact Hub Westminster

Address: 1st floor, New Zealand House, 80 Haymarket, London SW1Y 4TE

Phone: 020 7148 6720



The keynote speaker is Chris Taggart. Co-founder & CEO of OpenCorporates: The Open Database Of The Corporate World. Also founded OpenlyLocal, making local public data open and accessible

The prizes are:

- Comic con London weekend priority tickets
- Google chromecast! turn your TV into a smart TV, stream Netflix and more
- Bam-boom speaker! Amazing sustainable gadget, no power needed to amplify your phone!
- Powermonkey! solar power for your gadgets

There may also be an additional incubator membership scheme prize from Impact Hub Westminster if we get a really good hack

Evaluation Questions – please use these questions as prompts and record any key learning or discussion points. Bring the information to 'mission control' for collation.

Hack Team Name/ Project:
Questions for Hacker teams
What personas have you considered?
What have you done with the CLEAR Info data?
What uses have you created for the CLEAR Info data?
What level of CLEAR Info data have you used– Parent Co Level data or site level data, or both?
What benefits have you seen to having the Company Level data?

What CLEAR Info data sets have been most useful?
Which fields have you used?
What limitations have you found to using the CLEAR Info data?
How have you found the format of the data? How could it be improved?
Is there any other data you would have liked to have access to, to use in your solution?
Has the quality of the data met your needs? If not, how would it need to change?

What other data sets have you integrated with the CLEAR Info data?
What has worked well?
What problems have you faced with integrating the data
How could we make the CLEAR Info data easier to integrate with other data sets?
How have you visualised the data?

Personas

1. Brian Breeze, Air Quality Officer, Local Authority

Brian works in the Environment Department of a local authority. He is in his late 40's, always wears a suit, and is a reserved, quiet character. He lives on his own, keeps exotic house plants and enjoys a good social life through a massively multiplayer online role-playing game (MMORPG) World of Warcraft. Brian has a background in engineering and is good at processing information for pre-determined/defined tasks.

Background:

Brian is responsible for producing an Air Quality Action Plan for the local area. The plan identifies areas of high air pollution, the priority areas for action to reduce air pollution, and the monitoring strategy for recording air quality and tracking improvement.

The Air Quality Action Plan is due to be updated. Over recent years a number of members of the public have referred to the plan for information on the causes of local air pollution and the hot spots for poor air quality. The AQAP has limited information and people often request more information. Brian wants to expand the information available to support the next revision.

The Local Authority are responsible for identifying and monitoring air quality in the worst areas in the district. They publish the monitoring regularly. These hotspots for poor air quality are usually due to transport emissions.

The Environment Agency regulates larger industrial processes, and manages these by setting emission limits in site environmental permits. They also receive data from industrial sites reporting their emissions.

Usage scenario:

Brian would like to provide information about the emissions from industry in the district, alongside the transport emissions information in the Air Quality Action Plan, to provide a wider overview of the issues and pollutants. He wants to visualise the data to make it easy for people to interpret and draw conclusions from.

Example Problem Definition:

- Identify industrial sites with a Local Authority Boundary, and their emissions of air pollutants
- Identify which Industrial Sectors are key emitters in the area.
- Identify sources of Local Authority air quality monitoring data. One example is given below.
- Identify the key air pollutants in a district, both from transport, and from Industry.
- Develop a solution for 1 local authority, or an approach that can work for any Local Authority
- Identify recent trends in air quality in the Local Authority area.

EA Data sets:

- Industrial and Waste Permits – This contains site addresses, Local Authority, postcodes and grid references.
- Pollution Inventory– This contains emissions and waste data for industrial sites.
- Carbon Reduction Commitment - Carbon emissions for whole or partial companies.
- Company Hierarchy data – Shows the parent company who owns industrial sites.
- Environmental Pollution Incidents, contains incidents reported by the public
- Permit Breaches - Contains breaches or permit conditions recorded at permitted companies.

Other datasets

- Air Quality (data) from Local Authority monitoring e.g. www.bristol.airqualitydata.com
- Air Quality Management Areas (maps)
- Description of the impacts of different pollutants

How would this help CLEAR Info:

Explore Geospatial analysis of regulatory data.

Demonstrate linking EA data with other data sets such as Local Authority.

Generate new ways to visualise and interpret the EA data

2. Helen Home – Board Member of Heritage Healthcare Trust

Helen is a board member for an NHS healthcare trust. Helen is very friendly, homely and nurturing but also has a shrewd eye for business. In her early 40's she has a very busy home life and a passion for adventure sports. She is conscientious and needs to have confidence in data she is using. She is relatively new to the board with previous experience in both media and business management.

Background:

Healthcare Trusts can include a variety of sites:

- Doctors surgery
- Care homes
- Drop in clinics
- Local hospitals
- Clinical Incinerators

They are not companies, and collect data in different ways depending upon the size and complexity of their sites. Each site currently reports to the various environmental regulators on an individual basis but this information isn't collected centrally by HHT estates. It is difficult to get an overview of both environmental impact.

Key goals:

Another Health Care Trust recently caused a pollution incident and claims were made in the media that this caused a health hazard to the local population. In light of this recent scandal Helen is worried about the reputation of Heritage Healthcare Trust. She wants to check what information is available on the environmental performance of the Trusts' sites. Environmental impacts may include air emissions, waste, discharges to the water environment. Any breaches of environmental permit conditions, and risk rating done by the regulator would also be of interest.

A usage scenario:

Helen needs an authoritative source of information on the environmental impact of the Healthcare Trust. She knows some parts of the business are regulated by the Environment Agency. However, she does not have a definitive list of permits, licences and registrations they hold.

Example Problem Definition:

- Identify sites belonging to a Health Care Trust in the data sets, or look at the example given in the supporting materials. (They should not appear in Company hierarchy data because they do not have to register as a company).
- Identify permits, licences, and monitoring data belonging to Health Care Trusts, from the EA data
- Generate an overview of a Health Care Trust environmental performance, and highlight areas that need attention, or areas with potential health impacts.
- Bring in other data sets to give a wider range of environmental data, or to bring a health or another context to the data.

EA Data sets:

- Industrial & Waste Permits. This contains business sector, postcodes and NGRs etc Hospital Incinerators with permits can be identified from this data
- Pollution Inventory data. Contains waste transferred from site with an Industruial Installation permit, and air emissions.
- Permit Breaches – Record of breaches of environmental permit conditions
- Carbon Reduction Commitment – Carbon emissions
- Water Abstractions.
- Water Discharge Permits – Permits to discharge to the water environment (e.g. septic tanks and small sewage treatment plant in areas with no mains drainage)

Other materials

- A list of the sites and activities in the HHT estate (see example given in the supporting materials)

How would this help CLEAR Info:

How to link data in absence of an ownership hierarchy, and where names are different and there is no alternative identifier (i.e. not a company or no company number).

Supporting Information

Hackers could use Cambridge University Hospitals NHS Foundation Trust as an example for a Heath Care Trust.

Including the following estate:

1. Addenbrookes Hospital, Cambridge Uni Hospital NHS Foundation Trust
2. George Eliot Hospital
3. QMC Nottingham CHP Plan
4. Linden House Care Home
5. River Head House, Care Home
6. The Laurels drop in clinic
7. Huntingdon Road Surgery – (GP)

(NB this is a fictitious estate list, not all sites are really within Cambridgeshire. Most, but not all the sites listed appear in the data sets provided. Hackers may choose alternative examples to work with.)

Data on these sites can be found in

- Industrial & Waste Permits
- Permit Breaches
- Pollution Inventory
- Water Abstractions
- Water Discharge Permits
- Carbon Reduction Commitment

3. Gregory Gaia: Environment Manager, Earth Foods Manufacturing

Gregory leads a team at a food manufacturing company, focusing on their corporate responsibility for environmental, social and governance arrangements. In his spare time he is a caller for a Ceilidh band and runs a small festival in Somerset. Gregory is a big character, energetic, passionate and inspiring to those around him. He likes quick answers, and delegates detailed analysis to others in his team.

Background:

Earth Foods have a good history of compliance with their environmental permit conditions from Environment Agency and have a good reputation they want to maintain. They have held international accreditation (ISO14001) for their environmental management system for 5 years and (ISO9001) for their quality management systems for 6 years.

Earth Foods manufacture several different ready meals that specialise in being organic and nutritious. They have a good reputation because they apply a strong ethical approach to their food and everything about the way the business is run.

The company has a range of suppliers from primary food producers to packaging and logistics. Their suppliers range from SME's to big companies. (See example list of suppliers in the supporting information).

A usage scenario:

Gregory wants know if Earth Food's suppliers operate to the same responsible ethos that the company upholds. A high profile environmental accident by one of their suppliers could damage Earth Foods image with consumers.

Example Problem Definition:

- Use the example supplier list given in the supporting information, or create your own list of suppliers.
- Identify what EA environmental data is available about the suppliers
- Identify risks from bigger corporate supplier companies, and small sole traders.
- Consider alternative suppliers

- Bring in environmental data about the companies from other sources to give a more comprehensive picture.
- Identify sources of Social or Governance data, to give a rounded picture of sustainability in the supplier companies
- Set criteria, or design a methodology to identify suppliers that may be a reputational risk

EA Data sets:

- CLEAR Info data – for a range of environmental data collated to Company level
- OPRA Waste Ops (Operational Risk Appraisal) – This provides a risk score for sites with permits
- OPRA Installations (Operational Risk Appraisal) – This provides a risk score for sites with permits
- Permit Breaches– This show breaches of the environmental conditions on permits, e.g. emissions above the agreed threshold
- Environmental Pollution Incidents

Other materials:

- Example list of suppliers – given in supporting information
- Corporate responsibility reports from company websites or other sources

How would this help CLEAR Info:

Show how the cube extract and other publically available data can be used to analyse big companies.
Explore new potential uses of the data
Explore options for analysing and presenting the data.

Supporting Information

Earth Food Manufacturing

Example List of Suppliers

1. Alexander & Angell Farm Ltd – Pig Farmers
2. Balingour Limited – Poultry products
3. 2 Sisters Food Group/ Boporan Hold Co – Food Ingredient suppliers
4. Arla Foods – Food Ingredient supplier
5. Hay Farm Produce LLP – Food Ingredient suppliers
6. Sun Valley Foods Limited – Food Ingredient Suppliers
7. Merseyside Waste Disposal Authority – Waste management contractors
8. A W Chapman Ltd – Packaging suppliers
9. Zwanenberg Food UK Ltd – Packaging Suppliers

(NB this is a fictitious supplier list. Most, but not all supplier can be found in the data sets provided. Hacker may wish to choose alternative examples to work with).

Data on these sites can be found in

- Industrial and Waste Permits
- Permit Breaches
- Pollution Inventory
- Carbon Reduction Commitment

4. David Driver: Project Manager, WTI Trade Association

David is the project manager recently employed to lead an advice and guidance project for Waste Trade International. This is his first role after completing a part time MSc in Waste Management whilst working for a consultancy. David has recently moved to a new city for this job, and is relishing his new found free time, by joining a local cricket team and film lovers club.

Background:

Waste Traders International* are a trade association for the waste industry.

A big reputational issue for the waste industry can be amenity issues such as noise, dust and odour. These can result in complaints from the communities surrounding waste sites, and can also be damaging for wildlife and the environment.

WTI has secured government funding to provide an advice service for waste sites with amenity issues. David has been recruited to identify waste sites that would have most room for improvement, and benefit from the free support service.

WTI want to reach out to the whole sector. This would help improve reputation of industry and quantify impact of SME members.

(*WTI is a fictitious Trade Associations)

A usage scenario:

David doesn't know which of its members or non-members is causing the problems with noise, dust and odour. He wants to be able to identify how many businesses are in their sector, which ones are causing the issues and what size the companies are.

Suggested Development Options:

- Identify which permit breaches are for amenity issues, and which ones relate to waste companies
- Identify how many amenity breaches take place at companies registered with companies house, and how many take place at SMEs or sole traders.
- Create criteria for selecting companies to approach with the advice service.
- Create a methodology for monitoring the impact of the advice service on the occurrence of amenity issues at waste sites.
- Identify if other reasons for breaches are more prevalent and could be the focus for future advice service projects.
- Develop a tool that could be used by other trade associations in different sectors to carry out similar analysis.

Data sets:

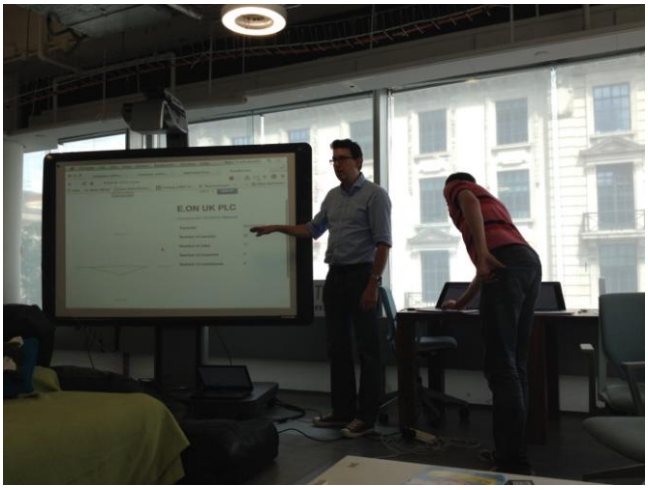
- Industrial & Waste Permits - This includes Installation sites, Waste Operation sites, permit number , and business sector information (this will be EA sector definitions), postcodes, etc
- CLEAR Info data set – data collated to parent company level, for companies matched to companies house numbers.
- Company Hierarchy data – to link site permit data to the parent company. Also Useful at company level for SIC codes (Sector Industry Classification codes – an external referencing system also used by Companies House.)
- OPRA Installations and OPRA Waste Ops (OPerational Risk Appraisal) – This gives a risk score for sites with permits,
- Permit Breaches–this information which can be potentially filtered for Amenity as type of permit breach.
- Environmental Pollution Incidents.

How would this help CLEAR Info:

Generate a solution to linking SMEs to sector.

Generate innovative ideas for analysing environmental risks by business sector and identifying targets for interventions.

1.7.3 Envhack2 photos



1.7.4 Evaluation Questions

Team One: River stuff - Simon

Is there any other data you would have liked to have access to, to use in your solution?

Actual discharge quality data – was confidential. Wanted to get type of treatment.

There was no water discharge / quality breaches included in the CCS dataset.

NIRs can't be linked to permits but do have eastings and northings that could potentially be attributed to WFD water bodies.

Pollution Inventory doesn't count volume of effluent discharged from sewers.

WIMS doesn't give you volumes emitted, just total capacity permitted. Also these vary on special circumstances e.g. storm flood conditions.

Has the quality of the data met your needs? If not, how would it need to change?

Discharge consent data was in MSAccess format – slowed things down as needs to be converted before it can be used. Would be better in open formats like CSV or XML.

The geographic data was in NGR and needed to be converted to WGS84 international globe format.

What other data sets have you integrated with the CLEAR Info data?

None. Tried FAME, list of companies and their activities. Wanted to explore companies in a specific area. Defaulted data that was provided.

What has worked well?

Reformatting has prevented notable progress. Good to have the data available.

Team Two: Alan, Gregory's suppliers

What have you done with the CLEAR Info data:

Two things: Created an API that allows people to easily access the data around which we've built some scoring metrics. Second, created a front end to visualise the metrics

What uses have you created for the CLEAR Info data?

A key for end users to access the 'risk' of dealing with a particular company, both in terms of supply and reputation.

What level of CLEAR Info data have you used?

Parent Company and site level data- users search by company but we generate our metrics based on the sites linked to that company and its child companies.

What benefits have you seen to having the company level data?

We can aggregate the metrics at all levels of the hierarchy, users can get an assessment of the company they are interested in without having to drill down. Data users or companies investigating their supply chain are often interested in a midlevel company rather then at site or parent company level. Having the full hierarchy to interrogate had investigations at individual levels / companies possible. Often Parent Companies might have very different assets in several different sectors and more diverse impacts / breaches then the company of concern.

Which CLEAR Info data set has been the most useful?

Company hierarchy data.

What limitations have you found using the CLEAR Info data?

Only a lack of knowledge. You do need an 'expert' on hand to help understand what the is and how is can be used.

Difficulty with normalisation to compare companies impacts or performance. Limited data availability with turnover or number of permits but not very accurate – production values would be much more useful but not available.

Understanding where to start to pull information together. If looking at a company performance the logical place to start would be with Breach data. However, if looking for a new supplier in a given sector this would not pick up good performing companies. Start with Company / sector information and look for permits and breaches from those companies.

How have you found the format of the data and what could be improved?

Flat files are difficult on a hack because time has to be spent processing them before we can actually make use of them.

Is there any other data you would have liked to have access to, to use in your solutions?

Flood risk, whether an area had low, medium or high levels of abstraction.

It would also have been good to have CAMs data on water availability to evaluate if a water intensive supplier could be at resource risk.

Actual amounts of water abstracted rather then total permitted would have been useful

Has the quality of the data met your needs?

On the basis that what was there was accurate, it was just where we felt data appeared to be missing was an issue.

What other data sets have you integrated with the CLEAR Info data?

We couldn't on the day, but we wanted to integrate flood risk and abstraction stress with the CLEAR Info data.

What has worked well?

The company hierarchy was key to our project, so it was awesome to have access to that data.

What problems have you faced with integrating the data?

We often found data either seemed to be missing or could be difficult to link across certain datasets. Also, just lack of knowledge we really relied on the data experts.

Lack of compatible data for WIMS catchment codes and area catchment names on data share catchment area files – cannot relate datasets.

How could we make the CLEAR Info data easier to integrate with other data sets?

Move it out of the flat files into some form of database or even better out an API over it to save us the work of having to do it!

How have you visualised the data?

We created scoring metrics on 4 points; pollution, water abstraction, breeched and flood risk and plotted them on a radar chart for a given company.

1.7.5 Post event communications

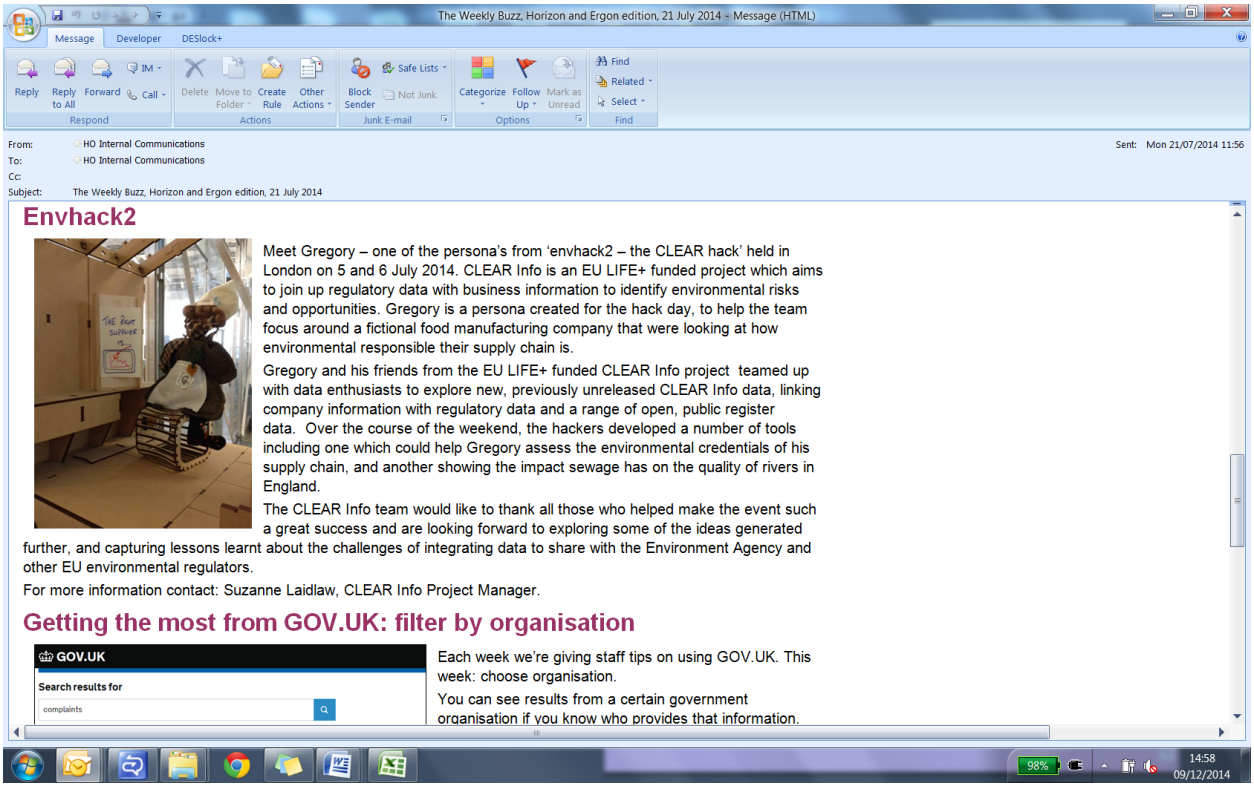


Image 1: Screenshot of Weekly buzz article

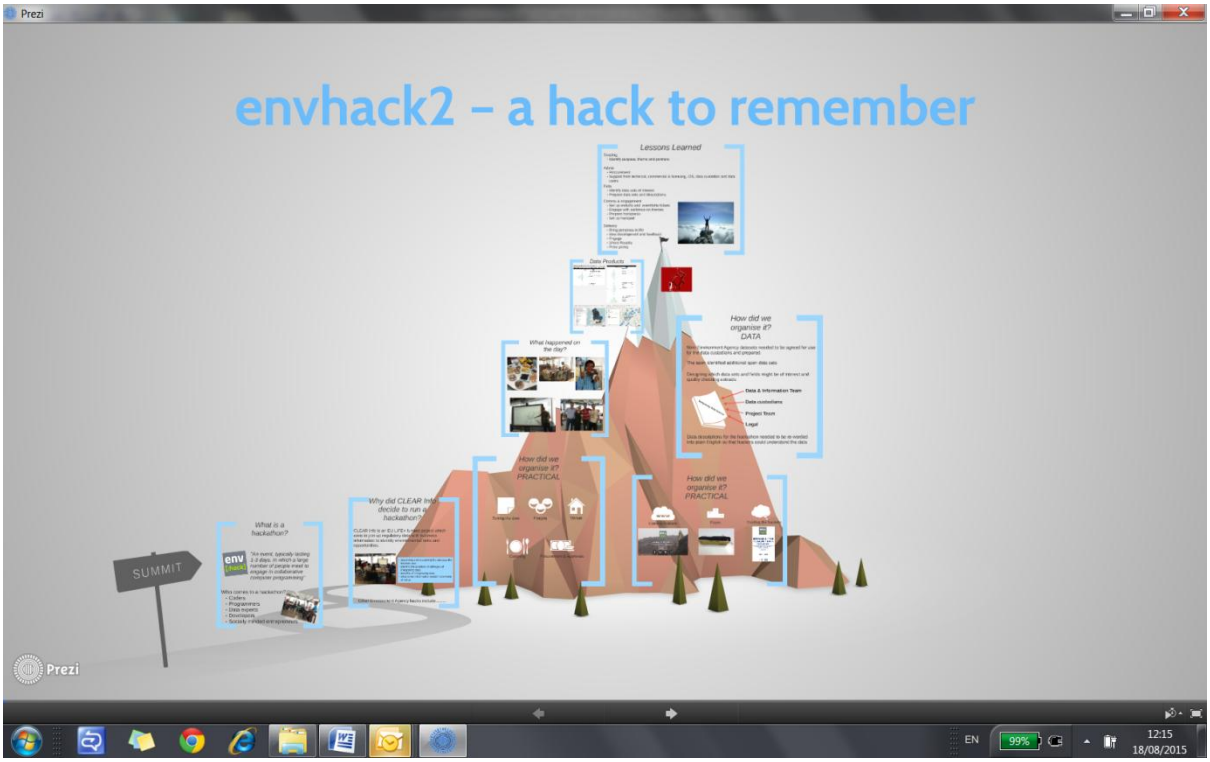


Image 2: Bite Size Learning Presentation Screenshot

Image 3: Bite Size Learning Presentation Slide 1

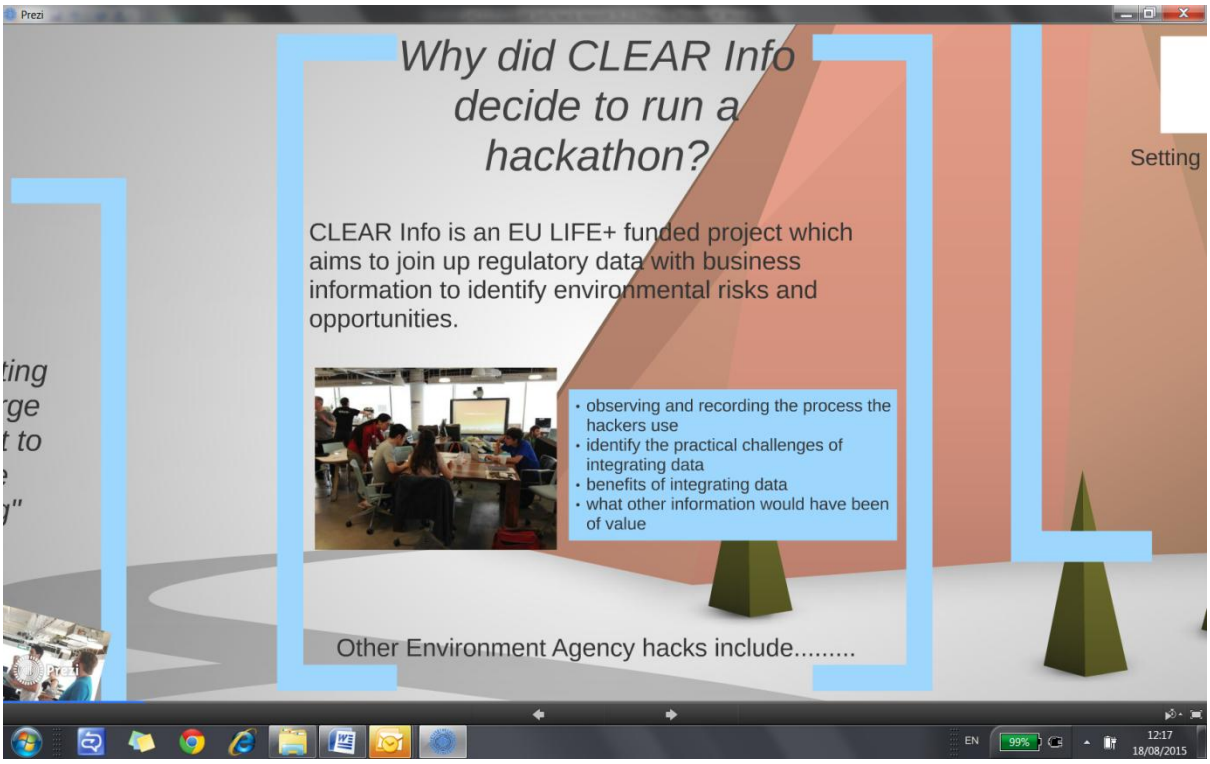


Image 4: Bite Size Learning Presentation Slide 2

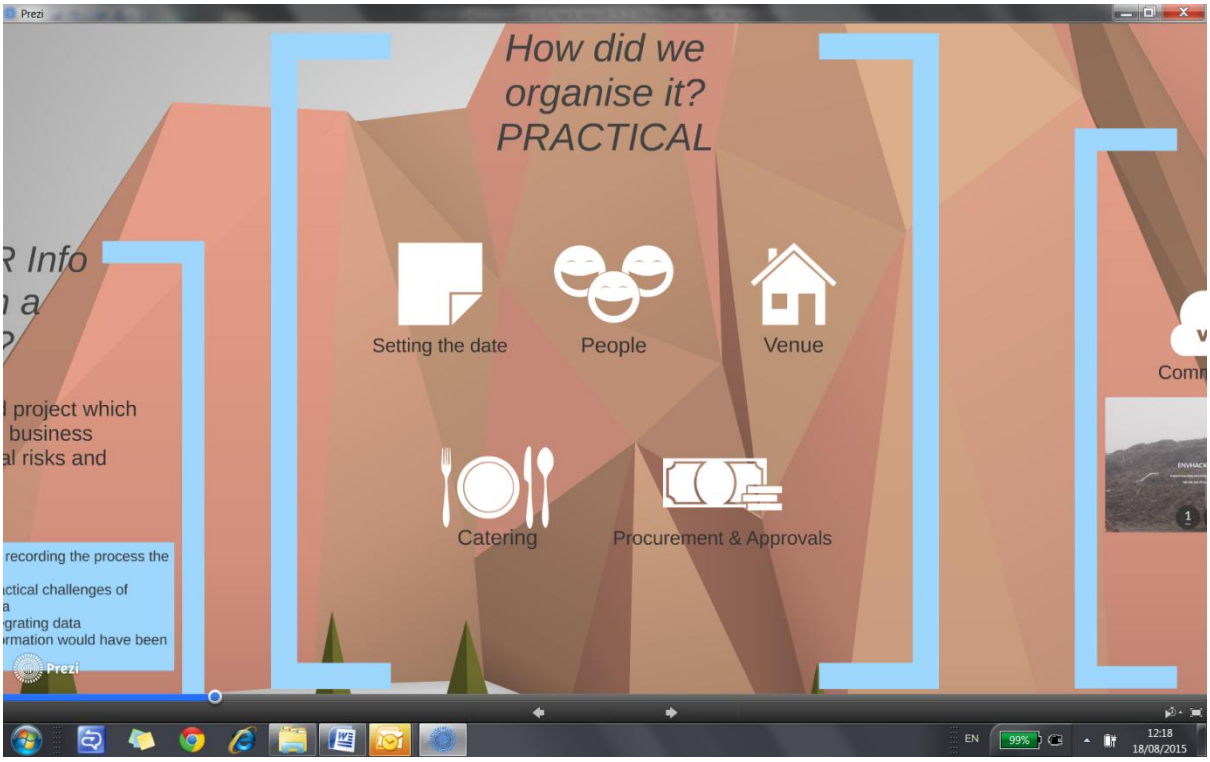


Image 5: Bite Size Learning Presentation Slide 3

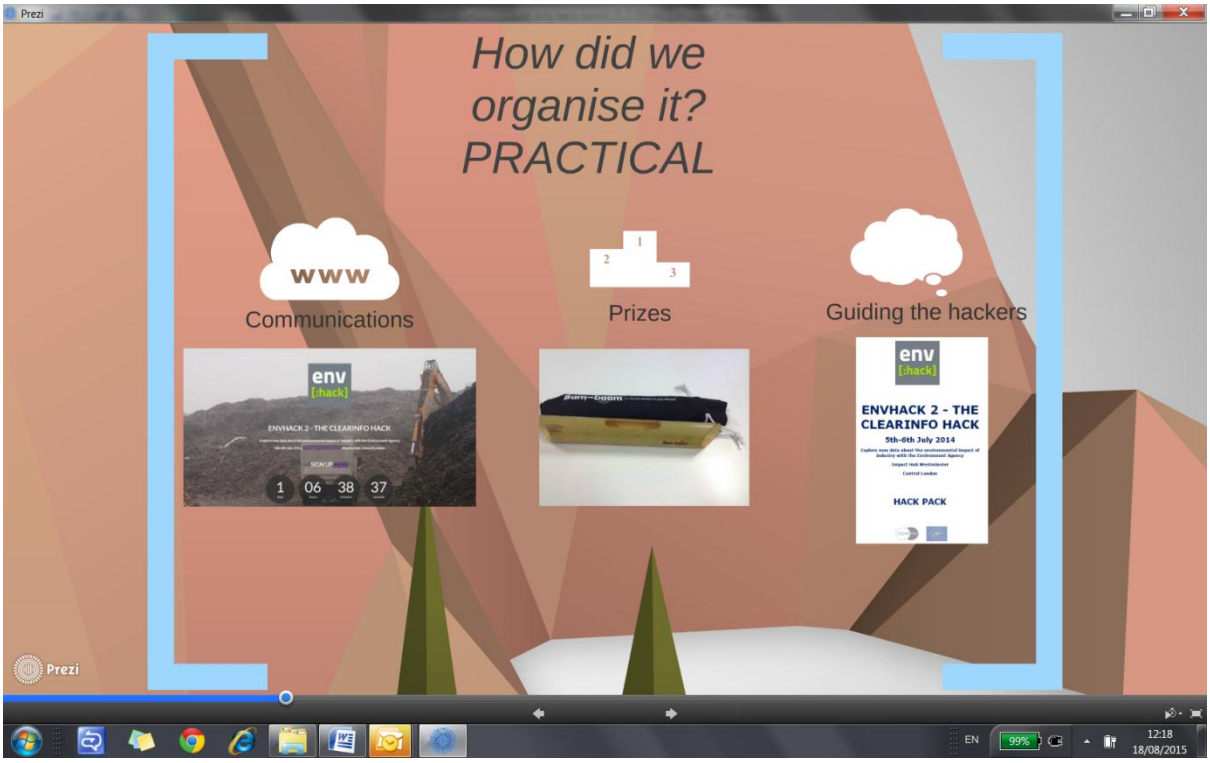


Image 6: Bite Size Learning Presentation Slide 4

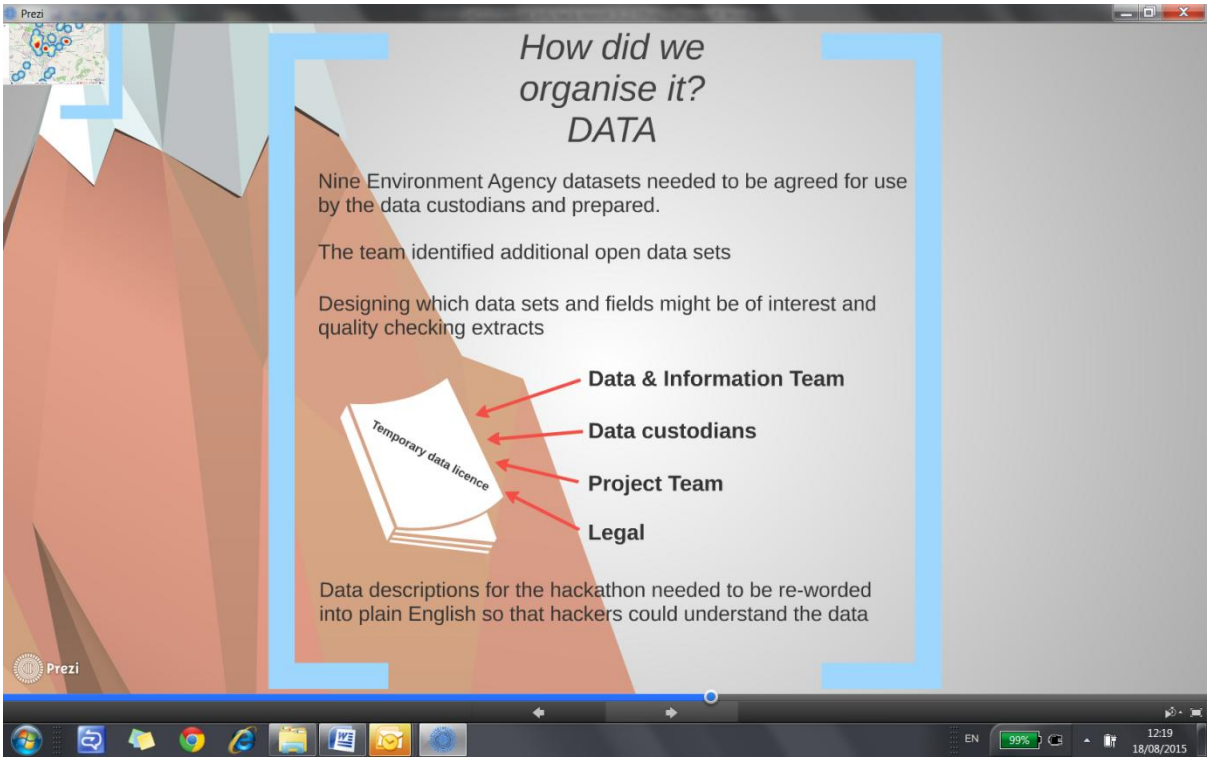


Image 7: Bite Size Learning Presentation Slide 5

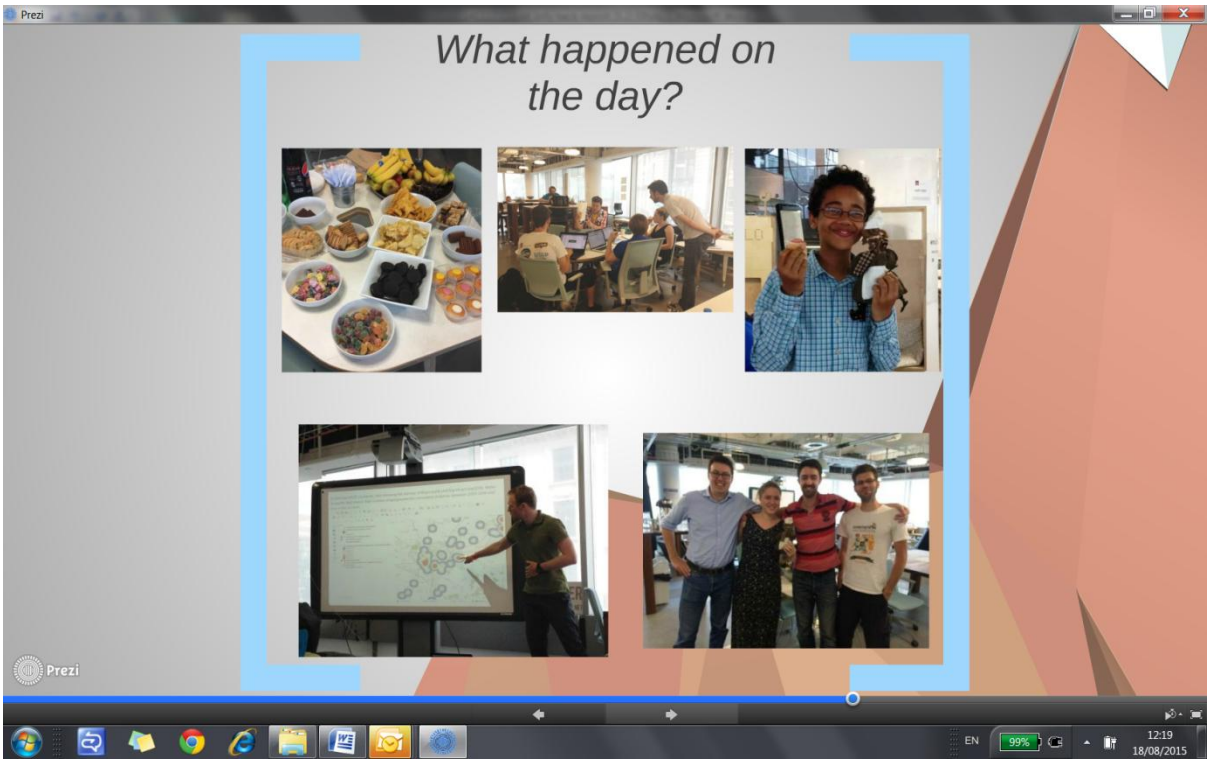


Image 8: Bite Size Learning Presentation Slide 6

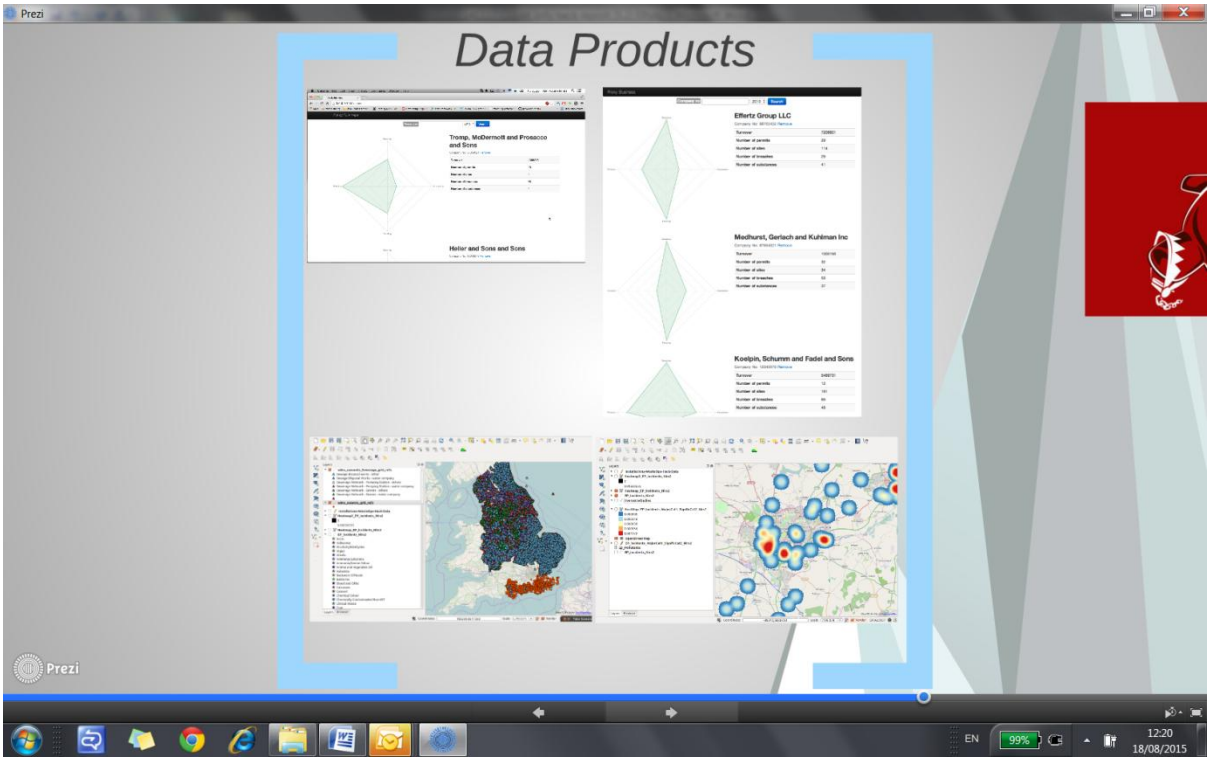


Image 9: Bite Size Learning Presentation Slide 7

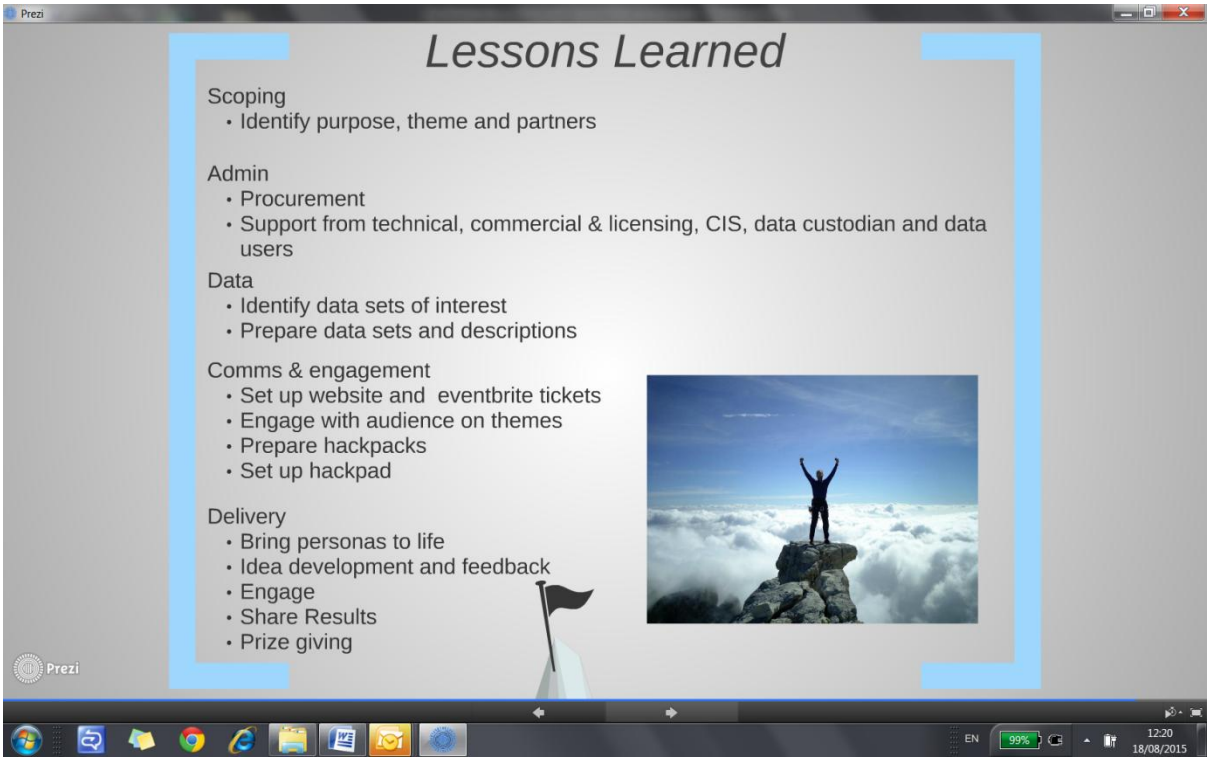


Image 10: Bite Size Learning Presentation Slide 8