Notes on Completion: Please refer to the appropriate NIA Governance Document to assist in the completion of this form. The full completed submission should not exceed 6 pages in total.

NIA Project Registration and PEA Document

Date of Submission

Sep 2013

Project Reference Number

NIA_NGN_030

Project Registration

Project Title

Predictive Analytics

Project Reference Number

NIA_NGN_030

Project Licensee(s)

Northern Gas Networks

Project Start

December 2012

Project Duration

1 year and 5 months

Nominated Project Contact(s)

Tony Pearson, Alec Breen and Tom Bell

Project Budget

£220,000.00

Summary

IBM/Presidion follows a structured process known as CRISP-DM. As part of this process, there will be clear milestones that will allow NGN and IBM/Presidion to monitor project progress and success to date. This process has 6 high level stages:

- Business Understanding: Typically, all interested and relevant parties are engaged to fully scope and define the business problem and resources available. A project plan will result from these meetings that will outline specific tasks and milestones and what resources will be needed for each.

- Data Understanding: At this stage, the IT team are initially involved so that Presidion can best understand the processes behind data collection. An assessment will be made as to the quality of the data held, and what is initially considered as likely factors in affecting failure to meet the 7/28 target. If the relevant data cannot be readily accessed, the project will be immediately halted at this point, and only resumed when the data access issue is fully resolved.

- Data Preparation: As the data for the project comes from a variety of different sources, and typically these are recorded at different levels, the information will be required to be manipulated and collated in a manner efficient for predictive modelling.

- Modelling: Using IBM-SPSS Modeller, Presidion will build predictive models to associated risk scores with cases. These risk scores will identify the major factors associated with a case failing to meet both the 7 and 28 day targets. The score itself, can act as a scientific tool for resource allocation and deployment strategies.

- Evaluation: Before deployment of the predictive models, Presidion will evaluate their accuracy, understanding and usability. This may result in alternative models being produced. Expected levels of prediction can then be used to inform NGN of likely successful deployment strategies.

- Deployment: As part of the deployment, Presidion will interface with existing NGN processes and systems to best get the score information to the Area Managers. Additionally, documentation will be provided to describe the process behind building the models

and what main factors were identified.

Preceding Projects

NIA_NGN_120 - Predictive Analytics Part Two

Third Party Collaborators

IBM

Nominated Contact Email Address(es)

innovation@northerngas.co.uk

Problem Being Solved

While NGN have an understanding of the existing reported gas escapes, they cannot forecast the number, type or locations of the next probable leaks which makes fleet management/resource planning difficult. Specifically this means putting the right resources/ equipment in the most likely locations to respond as quickly and effectively as possible to the incoming customer calls.

NGN has the following performance targets within the current safety case:

1. 85% of all escapes from the Network repaired within 7 days
2. 98.5% of all escapes from the Network repaired within 28 days
3. Any Escapes older than 28 days, with exception reporting on the actions being taken and reasons for any escapes over 40 days old.

At present, there are approximately 81% of cases closed within 7 days and 94% within 28 days. In order to hit the 7/28 target, NGN needs to consistently perform above 85% and 98.5% respectively over the majority of the year to prevent underperformance annually.

Method(s)

For the proof of concept stage, Predictive Maintenance Solutions from IBM will access multiple data sources in real time to predict equipment failure so the organization can avoid costly downtime and reduce maintenance costs. Driven by predictive analytics, these solutions can detect even minor anomalies and failure patterns to determine the areas that are at the greatest risk of failure. This early identification of issues will help NGN deploy limited maintenance resources more cost-effectively, maximize equipment uptime, and improve service levels for customers.

Once proof of concept has been established, other areas of investigation and analysis will be identified so that a range of methods (and potentially partner organisations) can be used to identify the best solutions for a range of problems.

Once suitable sample areas across a wide range of business activities have been identified, intensive investigation will be carried out on the available data to confirm its availability, quality, sufficiency, suitability for a range of analytical techniques / solutions and any data currently not currently available that would have a significant positive impact if it could be obtained.

Scope

For the proof of concept, IBM/Presidion follows a structured process known as CRISP-DM. As part of this process, there will be clear milestones that will allow NGN and IBM/Presidion to monitor project progress and success to date. This process has 6 high level stages:

1. Business Understanding: Typically, all interested and relevant parties are engaged to fully scope and define the business problem and resources available. A project plan will result from these meetings that will outline specific tasks and milestones and what resources will be needed for each.
2. Data Understanding: At this stage, the IT team are initially involved so that Presidion can best understand the processes behind data collection. An assessment will be made as to the quality of the data held, and what is initially considered as likely factors in affecting failure to meet the 7/28 target.If the relevant data cannot be readily accessed, the project will be immediately halted at this point, and only resumed when the data access issue is fully resolved.
3. Data Preparation: As the data for the project comes from a variety of different sources, and typically these are recorded at different levels, the information will be required to be manipulated and collated in a manner efficient for predictive modelling.

4. Modelling: Using IBM-SPSS Modeller, Presidion will build predictive models to associated risk scores with cases. These risk scores will identify the major factors associated with a case failing to meet both the 7 and 28 day targets. The score itself, can act as a scientific tool for resource allocation and deployment strategies.

5. Evaluation: Before deployment of the predictive models, Presidion will evaluate their accuracy, understanding and usability. This may result in alternative models being produced. Expected levels of prediction can then be used to inform NGN of likely successful deployment strategies.

6. Deployment: As part of the deployment, Presidion will interface with existing NGN processes and systems to best get the score information to the Area Managers. Additionally, documentation will be provided to describe the process behind building the models and what main factors were identified.

Following successful completion of a Proof of Concept, NGN will develop a strategic approach to the adoption of structured analytics. NGN will examine the full range of its business activities for areas which could potentially benefit from analytical modelling. From these, a small number of opportunities covering a wide range of functions (including asset management, operational delivery, investment efficiency, customer impact and environmental performance) will be identified. These will be intensively investigated to determine data availability and its suitability for a range of analytical solutions.

Test the use of Predictive Analytics in other areas of the business, using different approaches, structures and solution paths to determine the best techniques for a variety of situations

- Identify a range of areas where statistically-driven modelling could be beneficially used to inform intervention decisions and / or forecast future outcomes
- Identify the most appropriate Predictive Analytic approach to be taken in each of these areas
- Implement solutions
- Assess and compare the output benefits based on problem type and approach taken

Discovery phase to “deep dive” the data for the identified opportunities in order to :-

- Confirm its availability, quality, sufficiency, suitability
- Identify any additional data not currently available that would significantly improve the analysis
- Identify the range of analytical techniques / solutions that could be applied for each of the areas of investigation.

Objective(s)

Stage 1 is a proof of concept. Focusing on external escape reports, using existing easily available data from across the business.

Success criteria for the Predictive Maintenance project have been identified:

1. Accurate/repeatable identification of when and where a gas leak is most likely to be reported.

2. Understanding of the main factors affecting a case.

3. Demonstration of the benefit of deploying predictive modelling to business areas to enable more effective fleet/resource planning.

4. Capability to take ownership of the data models created, as part of the NGN roll out of Predictive Analytics

Stage 2 Undertake Research & Development of full live data trial Select full area model to extend trial to use a more sophisticated workable model using:

- Use structured and unstructured data
- Internal and external data
- Apply those predictions to real-world constraints to optimise decisions

- Develop “What If" simulations to measure risks and benefits
- Refresh live models with new data to “learn” from past decisions.

Financial Forecast Stage 1

Based on our experience, the propensity models will take 35 days to complete. Presidion will deliver a report and repeatable outputs that can be immediately used as part of the NGN Predictive Analytics implementation.

Consumer Vulnerability Impact Assessment (RIIO-2 Projects Only)

n/a

Success Criteria

The challenge is to test if Predictive Analytics can outline, using available NGN data, the factors associated with reported gas leaks and predict where the next geographic location for a leak is most likely to be. Would such a model enable more effective deployment of the resources by the Area managers, and give valuable additional insights towards planning and decision making. Success criteria for the Predictive Maintenance project have been identified:

1. Accurate/repeatable identification of when and where a gas leak is most likely to be reported
2. Understanding of the main factors affecting a case
3. Demonstration of the benefit of deploying predictive modelling to business areas to enable more effective fleet/resource planning
4. Capability to take ownership of the data models created, as part of the NGN roll out of Predictive Analytics.

Project Partners and External Funding

n/a

Potential for New Learning

n/a

Scale of Project

The proof of concept will be undertaken on historical "dead" data to provide evidence that this methodology could provide an improvement. After a review and re-planning exercise the trial will continue onto limited "live" data over a larger geographical area with the relevant changes in operational processes and human behaviour.

If this project is successful NGN will move into roll out within a specific area.

Technology Readiness at Start

TRL7 Inactive Commissioning

Technology Readiness at End

TRL8 Active Commissioning

Geographical Area

Initial trial will cover the West Yorkshire Area, Bradford, Leeds, Huddersfield, Dewsbury

Revenue Allowed for the RIIO Settlement

None Within the RIIO Business Plan:

- Innovation Section
- Project 11
- DLO Resource Management & Scheduling

Setting up a policy, process and systems to manage DLO resources across Emergency, Connections, Repair and Replacement so

that they can respond immediately to customer needs.

Indicative Total NIA Project Expenditure

£220,000

Project Eligibility Assessment Part 1

There are slightly differing requirements for RIIO-1 and RIIO-2 NIA projects. This is noted in each case, with the requirement numbers listed for both where they differ (shown as RIIO-2 / RIIO-1).

Requirement 1

Facilitate the energy system transition and/or benefit consumers in vulnerable situations (Please complete sections 3.1.1 and 3.1.2 for RIIO-2 projects only)

Please answer **at least one** of the following:

How the Project has the potential to facilitate the energy system transition:

n/a

How the Project has potential to benefit consumer in vulnerable situations:

n/a

Requirement 2 / 2b

Has the potential to deliver net benefits to consumers

Project must have the potential to deliver a Solution that delivers a net benefit to consumers of the Gas Transporter and/or Electricity Transmission or Electricity Distribution licensee, as the context requires. This could include delivering a Solution at a lower cost than the most efficient Method currently in use on the GB Gas Transportation System, the Gas Transporter's and/or Electricity Transmission or Electricity Distribution licensee's network, or wider benefits, such as social or environmental.

Please provide an estimate of the saving if the Problem is solved (RIIO-1 projects only)

- IRR 129%
- PV of Costs £144,836
- PV of Benefits (including PoS) £1,035,856
- Project NPV(inc Prob of Success) £ 935,570

This NPV is based on an operational saving of around £80,000 per annum removing the potential of failing standards of service and completing escapes sooner

Please provide a calculation of the expected benefits the Solution

The assumption made in the calculation is that the return on investment once fully rolled out will be 2:1 over a ten year period. Total investment £150,000. This is based on the reduction in emissions, ability to resolve leakage much earlier and having resources closer to location at time of reports occurring.

Please provide an estimate of how replicable the Method is across GB

NGN believe this is fully replicable across all GDN's and will share the results of our development with other GDN's.

Please provide an outline of the costs of rolling out the Method across GB.

The cost to roll out across other GDN's would be in line with those incurred by NGN.

Requirement 3 / 1

Involve Research, Development or Demonstration

A RIIO-1 NIA Project must have the potential to have a Direct Impact on a Network Licensee's network or the operations of the System Operator and involve the Research, Development, or Demonstration of at least one of the following (please tick which applies):

☐ A specific piece of new (i.e. unproven in GB, or where a method has been trialled outside GB the Network Licensee must justify repeating it as part of a project) equipment (including control and communications system software).

☐ A specific novel arrangement or application of existing licensee equipment (including control and/or communications systems and/or software)

☑ A specific novel operational practice directly related to the operation of the Network Licensees system

☐ A specific novel commercial arrangement

RIIO-2 Projects

☐ A specific piece of new equipment (including monitoring, control and communications systems and software)

☐ A specific piece of new technology (including analysis and modelling systems or software), in relation to which the Method is unproven

☐ A new methodology (including the identification of specific new procedures or techniques used to identify, select, process, and analyse information)

☐ A specific novel arrangement or application of existing gas transportation, electricity transmission or electricity distribution equipment, technology or methodology

☐ A specific novel operational practice directly related to the operation of the GB Gas Transportation System, electricity transmission or electricity distribution

☐ A specific novel commercial arrangement

Specific Requirements 4 / 2a

Please explain how the learning that will be generated could be used by the relevant Network Licensees

Once predictive analytics has been proven it could be implemented across all UK GDN's.

Or, please describe what specific challenge identified in the Network Licensee's innovation strategy that is being addressed by the project (RIIO-1 only)

Within the RIIO Business Plan:

- Innovation Section
- Project 11
- DLO Resource Management & Scheduling

Setting up a policy, process and systems to manage DLO resources across Emergency, Connections, Repair and Replacement so that they can respond immediately to customer needs.

☑ Has the Potential to Develop Learning That Can be Applied by all Relevant Network Licensees

Is the default IPR position being applied?

☑ Yes

Project Eligibility Assessment Part 2

Not lead to unnecessary duplication

A Project must not lead to unnecessary duplication of any other Project, including but not limited to IFI, LCNF, NIA, NIC or SIF projects already registered, being carried out or completed.

Please demonstrate below that no unnecessary duplication will occur as a result of the Project.

n/a

If applicable, justify why you are undertaking a Project similar to those being carried out by any other Network Licensees.

n/a

Additional Governance And Document Upload

Please identify why the project is innovative and has not been tried before

n/a

Relevant Foreground IPR

n/a

Data Access Details

n/a

Please identify why the Network Licensees will not fund the project as apart of it's business and usual activities

n/a

Please identify why the project can only be undertaken with the support of the NIA, including reference to the specific risks(e.g. commercial, technical, operational or regulatory) associated with the project

n/a

This project has been approved by a senior member of staff

☑ Yes