Final report

Project code: V.RHM.0060

Prepared by: Kimon Kotronis and Roberto Mosca
Clarivate Analytics

Date published: 30 December 2017

PUBLISHED BY
Meat and Livestock Australia Limited
Locked Bag 1961
NORTH SYDNEY NSW 2059

Smart Packaging Interactive Dashboard

Meat & Livestock Australia acknowledges the matching funds provided by the Australian Government to support the research and development detailed in this publication.

Abstract

Clarivate Analytics was engaged in a project with Meat & Livestock Australia Ltd (MLA) to provide information and useful insight from patent and published literature data into the structure and nature of innovation and research within the 'smart packaging' technology field, specifically to identify technologies that can assist Australian food producers to deliver food products to exports markets that deliver on food safety, provenance, shelf life, traceability and integrity that can contribute to capturing a price premium.

Initial deliverables were provided in Microsoft Excel and PDF format, which were, subsequently, further analysed by Clarivate Analytics Professional Services Enabling Technology team in order to define and design an interactive dashboard. Clarivate Analytics defined the functional and technical specification of the dashboard under a "Define & Design" project. This project builds on the preliminary findings reported in MLA project V.RDP.1003 which collected 12,930 patent invention families and 4,441 published literature documents identified for analysis and construct of a dashboard

MLA contracted Clarivate Analytics to develop and deploy the interactive dashboard. The project was run using Agile Scrum methodology and was successfully completed in September 2017.

The aim of this final report is to present the work performed as part of this engagement, the methodology used for the development of the dashboard as well the lessons learned from the project for future reference.

Executive summary

Clarivate Analytics understands that MLA's objective is to enable food producers to have easy access to information related to smart packaging technologies from Patent and Non-Patent Literature, in order to assist them with identifying technologies that will allow them to export their products in packaging that ensures food safety, provenance, maximises shelf life, enables traceability and integrity and enables them to position themselves at premium segment of their markets.

The objective of the project is to develop an interactive dashboard that allows end users to easily query and filter datasets previously provided in MS Excel in order to discover suitable technologies, trends and packaging suppliers for their needs.

Under the Define and Design project, Clarivate Analytics' design team analysed the data and worked on the following:

1. Graphic design language: matching both Clarivate Analytics and MLA visual style. That includes but is not limited to: page layout, logos, font and common elements such as text blocks and buttons.
2. Graphic design of all proposed pages /sections.
3. Options and graphic design for chart types, layouts, visual and interactive features.
4. Options and graphic design for tabular representation of data with filters and other convenience features.
5. Technical specification of the solution, including hardware and software requirements, browser compatibility.
6. Pricing for multiple options and timelines.

Additionally and in order to enable the performance of the tasks mentioned above, the team also worked on the following:

1. User and task analysis: understanding user needs and defining workflows and use cases
2. Information architecture: organizing the information to be displayed in the dashboard

Following MLA's acceptance of the Define & Design deliverable and associated proposal for the development and deployment of the interactive dashboard, Clarivate Analytics developed and deployed the interactive dashboard on own servers in September 2017. At the time of writing of this report, final hosting arrangements are still in progress and the system remains hosted on Clarivate Analytics servers, although hosting services are not in scope of the current Maintenance and Support agreement.

Table of contents

1 Background 5

2 Project objectives 5

2.1 Business Objectives 5

2.2 Scope 6

3 Methodology 9

3.1 Define and Design approach 9

Option 1: Complete solution 9

Option 2: Minimum Viable Product 9

4 Design specifications 10

4.1 Option 2: Functional and Technical Requirements 10

4.2 Option 2: Technical Requirements 11

4.3 Technical Architecture 12

4.3.1 System Architecture 12

4.3.2 Technology Stack 12

4.3.3 Data Extract Transform Load (ETL) 12

5 Delivery 13

5.1 Deliverables & Methodology 13

5.2 Work Packages 13

5.3 Implementation methodology 14

5.4 Maintenance and Support 15

6 Discussion 16

6.1 Lessons learned 16

6.2 Challenges and Risks 17

1 Background

Clarivate Analytics was engaged in a project with Meat & Livestock Australia Ltd (MLA) to provide information and useful insight from patent and published literature data into the structure and nature of innovation and research within the 'smart packaging' technology field, specifically to identify technologies that can assist Australian food producers to deliver food products to exports markets that deliver on food safety, provenance, shelf life, traceability and integrity that can contribute to capturing a price premium.

Subsequently, Clarivate Analytics was asked to define and design the functional and technical specifications for an interactive, web-based dashboard that allows the intuitive querying and user-friendly visualization of the underlying data, to enable wider sharing of knowledge with food producers.

Clarivate Analytics developed and deployed the interactive dashboard on own servers in September 2017. At the time of writing of this report, final hosting arrangements are still in progress and the system remains hosted on Clarivate Analytics servers, although hosting services are not in scope of the current Maintenance and Support agreement.

The aim of this final report is to present the work performed as part of this engagement; the methodology used for the development of the dashboard as well the lessons learned from the project for future reference.

2 Project objectives

2.1 Business Objectives

Clarivate Analytics understands that MLA's objective is to provide access to this information to food producers, in order to assist them with identifying technologies that will allow them to export their products in smart packaging that ensures food safety, provenance, maximises shelf life, enables traceability and integrity and enables them to position themselves at premium segment of their markets.

We believe that a number of other user categories may have an interest in the information contained in this dashboard and may benefit from its use. Additionally, it is expected that MLA will require access to information on site usage so that the ROI can be calculated and the investment can be justified.

Our analysis has identified the following four types of users (personas):

- Food producers, accessing the system directly or via MLA agents.
- Packaging producers and their agents
- Researchers in packaging technologies, either in commercial or academic institutions
- System owners or MLA executives, interested in system utilization.

The dashboard was designed in such way so as to allow the flexible interrogation of the underlying datasets by any and all of the above types of users, in order to quickly and easily produce summary

visualizations that answer their smart packaging technology queries and allow them to drill down to the underlying matching data.

2.2 Scope

The scope of the project was restricted to the development of Option 2: Minimum Viable Product as described in the Define and Design document delivered in April 2017. Further, the build was based on key data sources derived from preceding MLA project V.RDP.1003 (original contracted with Thomson Reuters) which generated key deliverables such as - ThemeScape mapping – patent landscape and literature landscape mapping - and Entities filing bubble charts and Taxonomy by Technology Categories in the Smart packaging IP and patent search completed for MLA as shown below:

Figure 1 - Patent Landscape:

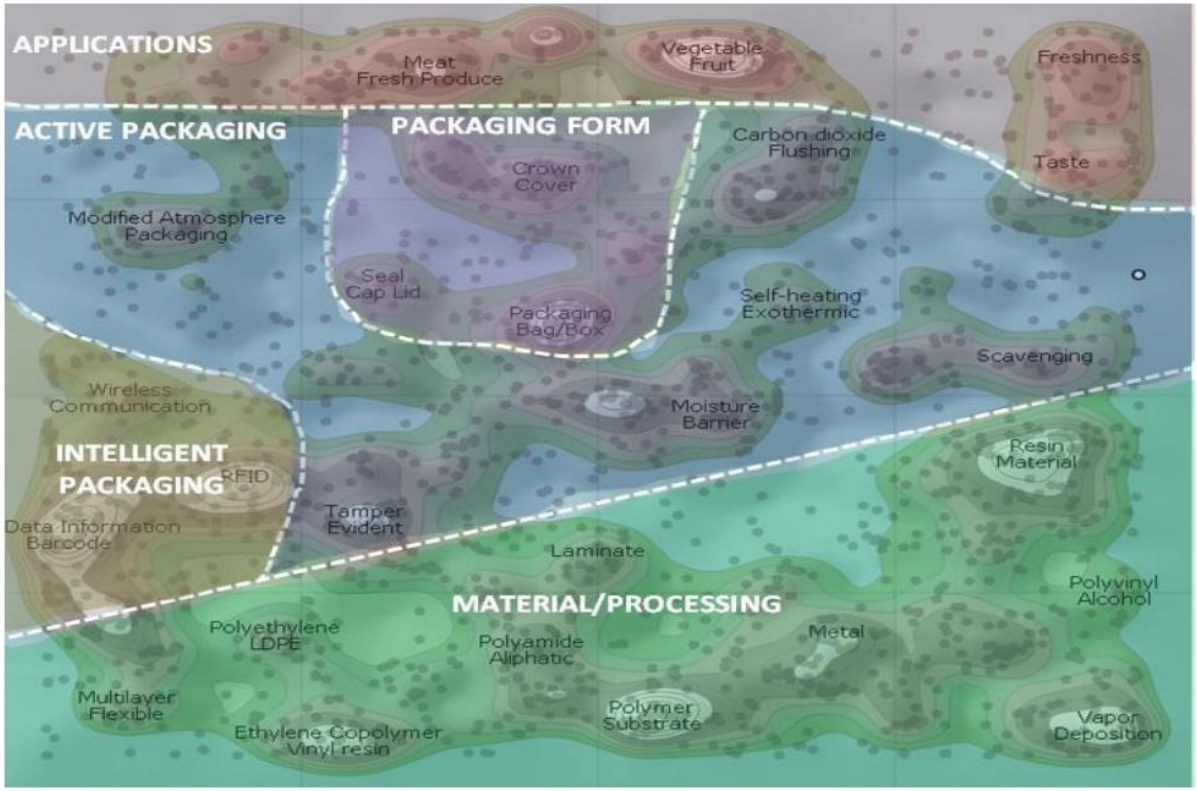

Figure 2 - Literature Landscape:

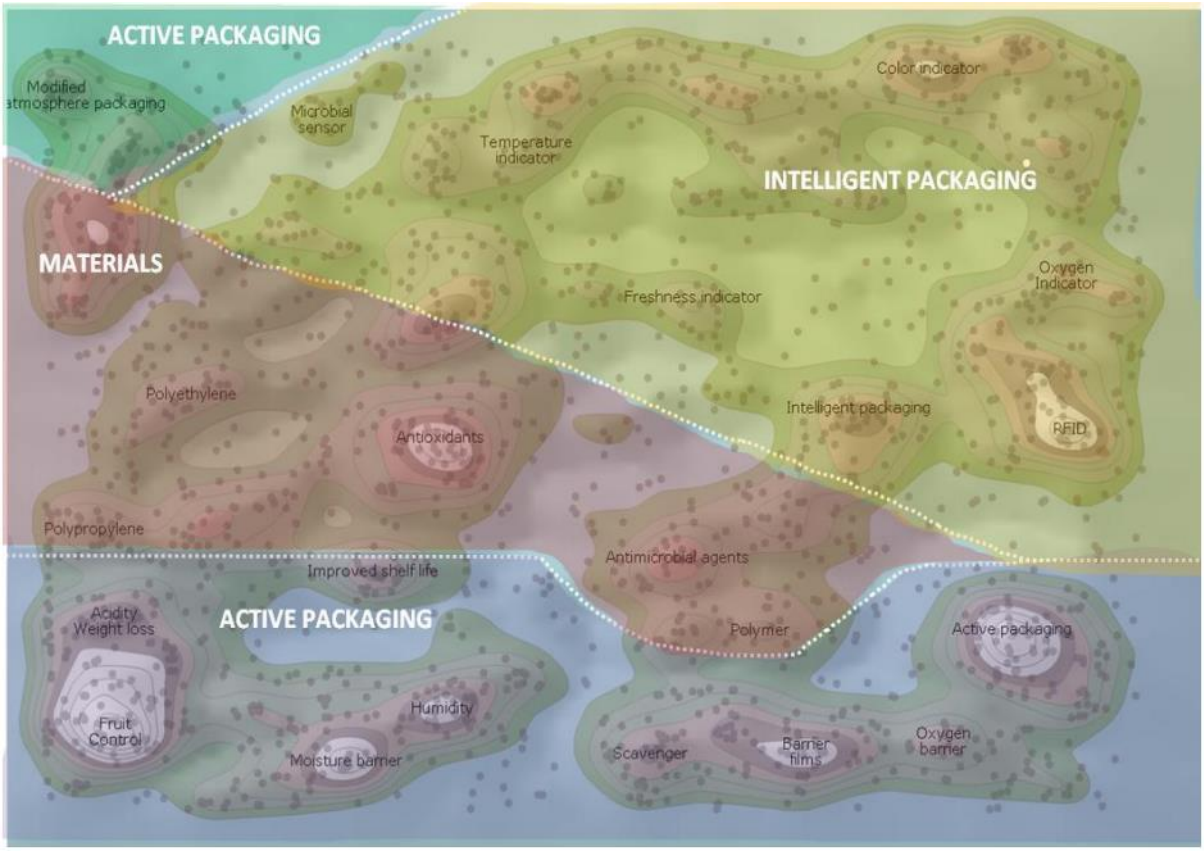

Figure 3 – Tier 1 Entities – Technical Breadth Bubble Chart

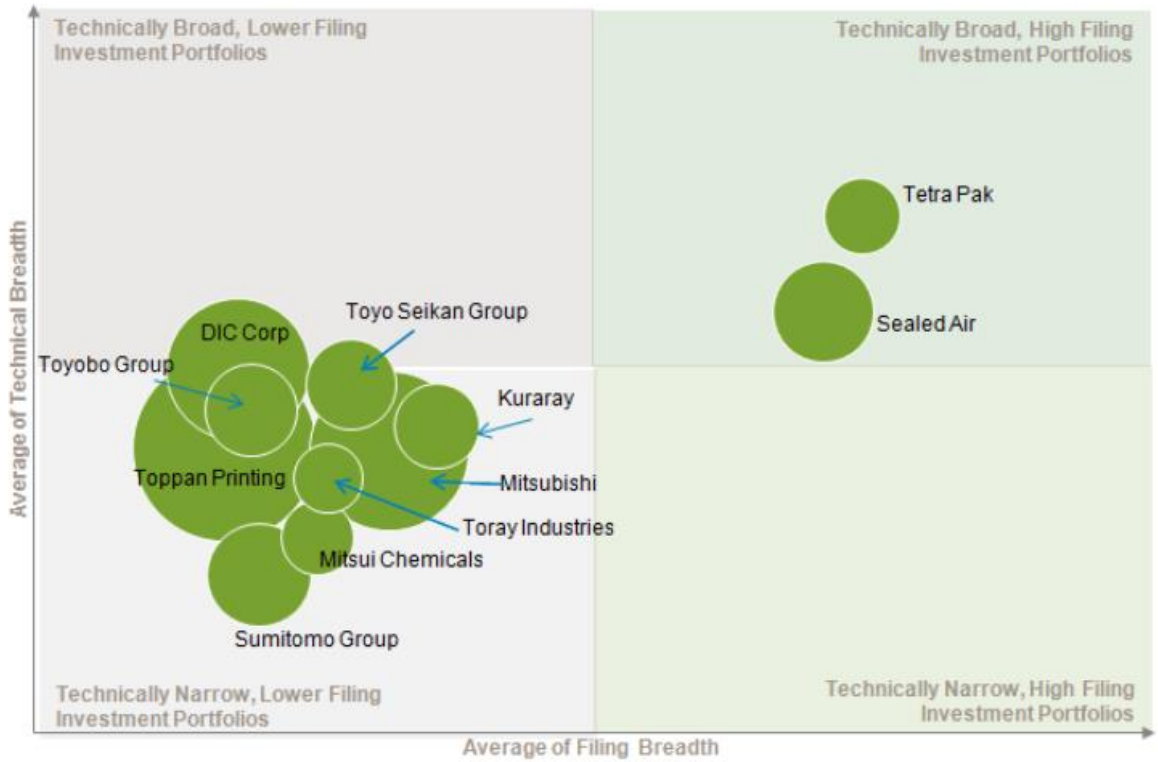

Figure 4 – Taxonomy – Technology Categories

	Technical Categories	Inventions
1. Function / Technical Features	1.01.01.01 Microbial growth sensor	219
	1.01.01.02.01 Anti-microbial agent	108
	1.01.01.02.02 Enzyme based microbial inactivation	23
	1.01.01.02.03 Other Anti-microbial systems	256
	1.01.02.01.01 Oxygen sensor	809
	1.01.02.01.02 Oxygen scavenging	4497
	1.01.02.01.03 Anti-oxidation systems	552
	1.01.02.02.01 Ethylene sensor	54
	1.01.02.02.02 Ethylene scavenging	877
	1.01.02.03 Other gases of interest	4612
	1.01.02.04.01 Low oxygen (or anoxic) packaging	983
	1.01.02.04.02 High oxygen packaging	260
	1.01.02.04.03 Vacuum packaging	907
	1.01.02.04.04 Other Modified/Controlled atmosphere packaging	1129
	1.01.03.01 Temperature sensor	1019
	1.01.03.02 Thermostatic packaging	32
	1.01.03.03 Self-heating/Self-cooling	539
	1.01.04.01 UV blockage	177
	1.01.04.02 Light barrier	387
	1.01.05.01 Moisture sensor	949
	1.01.05.02.01 Moisture barrier	3044
	1.01.05.02.02 Moisture content or humidity maintenance	1532
	1.01.06.01 Odor absorbing	76
	1.01.06.02 Flavor releasing	25
	1.01.07 Corrosion inhibition	113
	1.01.08 Integrity sensor	72
	1.01.09 Freshness sensor	430
	1.01.10 Eating profile enhancement	141
	1.01.11 Migration prevention	97
	1.01.12 Pathogens/microbial spoilage	66
	1.02.01 Temperature sensor	1021
	1.02.02 Physical shock/motion sensor	127
	1.02.03 Other environmental factors	690
	1.03.01 Tamper-proof	99
	1.03.02.01 Near field communication enabled	63
	1.03.02.02 RFID tags	818
	1.03.02.03 Printable electronics	90
	1.03.02.04 QR code	746
	1.03.03 Electronic article surveillance	29
	1.03.04 Wireless communication network	355
	1.03.05 Augmented reality	39
	1.03.06 Environment sensing and responding	2
	1.04.01 Easy opening of food packs	155
	1.04.02 Easy dispensing	17
	1.04.03 Portioning/sizing	6
	1.04.04 Resealable	87
	1.04.05 Cook within package	574
	1.04.06 On-the-go enabling	345

More specifically, the system consists of a single page offering a faceted search and analysis functionality as follows:

1. **Search filters by technical category (5 subcategories)**
   - Users are able to select any number of technical categories from each subcategory. Criteria inside each subcategory are joined by “OR” and between subcategories by “AND”.

- User can also select any number of Countries of publications as well as a date range

2. **For a set of criteria the following charts will be produced:**
   - A line chart with two lines, showing the total number of publications for Patents and NPL respectively, by year
   - A clustered stacked bar chart, showing the contribution of each technical category to the total number of publications (Patents and NPL)
   - A global map with Pie Charts for Patent and NPL publications by country for the selected criteria
   - A tree map showing the number of publications by assignee or author for the selected criteria
3. **All charts are interactive:**
   - Clicking on a point on the line chart will display a table with all Patents or NPL corresponding to the criteria and the selected year
   - Clicking on a box on the stacked bar will display a table with all Patents or NPL corresponding to the Technical Category for the selected criteria
   - Clicking on a piece of the pie chart on the global map with display a table of relevant publications for the chosen country
   - Clicking on a box in the tree map will display a table of publications corresponding to that author or patent assignee for the selected criteria.

3 Methodology

3.1 Define and Design approach

Clarivate Analytics performed a requirements analysis and solution design project, producing the functional and technical specifications for the following two options:

Option 1: Complete solution

This solution was designed to make use of all available datasets, additional reports, as well as to include a number of value-adding features for user management, usage tracking, presentation of pre-compiled reports and more advanced search and query features.

Option 2: Minimum Viable Product

This option is a restricted version of the complete solution, keeping the core visualization features but without any of the value-add features.

Option 2 was selected on the basis of affordability and cost effectiveness, as it provides all the necessary functionality at the minimum cost.

4 Design specifications

4.1 Option 2: Functional and Technical Requirements

Option 2 was deemed to be the most cost-effective solution for MLA and the one selected for implementation. More specifically, the functional for Option 2 are summarised in the table below.

ID	Category	Requirement	Priority
F1	**Home Page**	The home page must be accessible on any modern, desktop Internet browser on a public URL.	Must have
F2		The user will be able to return to the Home Page from any of the pages in the system without having to log out.	Must have
F3	**Interactive Dashboard**	Users are able to select any number of technical categories from each subcategory. Criteria inside each subcategory are joined by "OR" and between subcategories by "AND".	Must have
F4		Users can also select any number of Countries of publications as well as a date range	Must have
F5		Users can create a line chart with two lines, showing the total number of publications for Patents and NPL respectively, by year, that match the selected criteria	Must have
F6		Users can create a clustered stacked bar chart, showing the contribution of each technical category to the total number of publications (Patents and NPL) that match the selected criteria	Must have
F7		Users can create a global map with Pie Charts for Patent and NPL publications by country for the selected criteria.	Must have
F8		Users can create tree map showing the number of publications by assignee or author for the selected criteria.	Must have
F9		Clicking on a point on the line chart will display a table with all Patents or NPL corresponding to the criteria and the selected year	Must have
F10		Clicking on a box on the stacked bar will display a table with all Patents or NPL corresponding to the Technical Category for the selected criteria	Must have
F11		Clicking on a piece of the pie chart on the global map with display a table of relevant publications for the chosen country	Must have

		Clicking on a box in the tree map will display a table of publications corresponding to that author or patent assignee for the selected criteria.	

4.2 Option 2: Technical Requirements

The technical requirements for Option 2 are summarised in the table below.

ID	Category	Requirement	Priority
T1	**System requirements and accessibility**	The System must be accessible by any Internet connected device.	Must have
T2		The System must operate fully on current and previous major version of all modern Web browsers, namely: Microsoft Internet Explorer and Edge, Google Chrome, Mozilla Firefox, Apple Safari.	Must have
T3		The System will not follow responsive design therefore its use on devices with a screen resolution of less than 1280 x 1024 will not be supported	Won't have
T4		The must work without the need for additional web plugins, such as Java or Flash	Must have
T5		The System may work on other Web Browsers and older versions but they will not be formally supported	Won't have
T6		The System will not comply with any Accessibility standards.	Won't have
T7		The System must be usable on Internet connections of >0.5 MBps	Must have
T8	**General appearance and User Interface**	The System must follow MLA branding guidelines and color palettes	Must have
T9		The System must follow good UI principles	Must have
T10		The System must not require any Browser plugins	Must have
T11		The System must operate fully in all modern desktop PC web browsers	Must have
T12		The System could be usable on mobile devices, but some features may not work fully.	Could have
T13	**Security**	The System will be hosted on MLA servers.	Must have
T14		The System will not integrate with any federated authentication systems, either public (e.g. Google, Facebook etc.) or private (e.g. institutional Single Sign On systems)	Must have
T15		The System will provide a Secure Socket Layer (SSL) interface	Must have
T16		The System will encrypt data in transit.	Must have
T17		The System will not encrypt data at rest.	Won't have

T18	**ETL**	Data on the system must be able to be periodically - biannually or annually- refreshed through an ETL process run by Clarivate Analytics.	Must have

4.3 Technical Architecture

4.3.1 System Architecture

The System was built using a three-tier architecture:

1. A front-end web server serving the content that is rendered by the browser. The content is a mixture of static (e.g. decoration, navigation menus, help files) and dynamically generated content based on the underlying datasets.
2. A middle dynamic content processing and generation level application server using Spring Framework on Apache Tomcat.
3. A back-end database comprising both data sets and the database management system software that manages and provides access to the data, based on PostgreSQL.

4.3.2 Technology Stack

The system was built using the following technology stack:

Grails

Server-side web framework, powers GUI and APIs. Grails is an enterprise framework built on top of Spring, designed for rapid development and ease of deployment.

AngularJS

Front-end framework. Originally developed by Google, it is #1 framework to build modern web applications.

D3JS

D3 is a JavaScript visualization library and the De facto standard for building interactive graphic visualizations on the web.

Bootstrap

An Open Source HTML, CSS and JS framework that helps create responsive web applications what work on all devices.

PostgreSQL

This will be the primary database, holding all structured (numerical) and unstructured (documents) data.

4.3.3 Data Extract Transform Load (ETL)

The System is fed with the existing datasets, produced by the IP team. The Datasets can be easily imported into the database by Clarivate Analytics developers.

The files that will be used for ETL are:

1. MLA - Smart Packaging - NPL - Categorized Dataset – YYYY.xlsx
2. MLA - Smart Packaging - Patent - Categorized Dataset - YYYY.xlsx

As mentioned in the assumptions above, there is no provision for continuous data refresh. Periodic updates can be applied on an ad-hoc basis or according to a schedule that can be agreed separately, provided that data is provided in the agreed template that can be automatically imported into the database without the need for manual curation.

5 Delivery

5.1 Deliverables

Clarivate Analytics assigned a Delivery/Project Manager with the following responsibilities:

- Manage the development team following Agile Scrum
- To act as the "voice of the customer" with the development team, translating the design specifications into actionable user stories and tasks.
- To manage the project's resources and budget
- To manage the project risks and issues
- To be the primary point of contact with MLA.

The following deliverables were delivered:

1. Software solution outlined in this Define & Design document, deployed and ready to use
    - Milestone achieved 30th September 2017.
2. System documentation (help files)
    - Milestone achieved 30th September 2017.
3. System deployed on MLA/Intercept servers
    - Milestone achieved 1st December 2017

5.2 Work Packages

ID	Work Package	Task	Notes
W1	**Infrastructure**	Create project, setup development environment	
W2	**Database Management Systems, Direct**	Create domain models, set relations	

	Query Service and CRUD API and other services	Create database, indexes etc	
		Service & endpoints for executing PostgreSQL queries	
W3	**Single Analysis Page**	Design	Includes the design of reports from a Data as well as UI point of view
		Facets and Queries	
		Line Chart and table	
		Clustered and stacked chart and table	
		Geographic Map and tables	
		Tree map and table	
W4	**Other**	Knowledge transfer from IP team	
		User Experience design	Marginal additional effort for UI designer to work with developers
W5	**Testing**	Testing	
W6	**Adjustments**	Styling and polishing	Marginal improvements and bug fixes
		Bug-fixing	

5.3 Implementation methodology

Development followed the Agile Scrum methodology with work tasks being grouped in Sprints, lasting 2 weeks.

Tasks are described in User Stories, that take the form of “As a {Role}, I want to be able to {action}, so that I can {benefit}”. Each User Story is supported by one or more wireframes, demonstrating the User Interface features and design.

Work for the next sprint is planned in Sprint Planning sessions, which were held either separately or, in some cases, combined with retrospectives, depending on workload. The purpose of Sprint planning sessions was to:

- Prioritize User Stories for the coming sprint
- Go through User Stories in detail to ensure that the development team have a good understanding and that the stories are well written with clear acceptance criteria

- Prioritize bugs and improvements discovered in previous sprints
- Organize QA activities and review test cases

At the end of each Sprint, a retrospective meeting was held with the development team. The team reviewed the functionality delivered in that period, as well as the bugs that were both discovered and resolved. All team members contributed their views on the following:

- What worked well and we should continue doing
- What did not work and we should stop doing or change
- Ideas on new ways of doing things

All software development and QA (test cases and bugs) documentation was maintained in a private instance of GitLab issue tracking.

5.4 Maintenance and Support

The current maintenance and support contract does not include hosting services, as they were planned to be provided by MLA. Hosting arrangements are currently being reviewed, but it has not been decided yet if the maintenance contract will need to be amended to include hosting costs borne by Clarivate Analytics.

Clarivate Analytics will provide minimal application maintenance on a “as needed’ basis

The Scope of the Maintenance and Support Services is limited to:

1. Application Support services related to all software components comprising the System developed and deployed by Clarivate Analytics on behalf and for the Customer. Issues that are covered by this agreement include:
    - System malfunctions not related to the underlying Network, Server hardware or Operating System
    - Application software defects
    - Low performance due to poor System Design
2. Monitored email support
3. Remote assistance using Remote Desktop and a Virtual Private Network where available
4. Planned or Emergency Onsite assistance (extra costs apply), in cases of on-premises System hosting.
5. Annual data refresh.

System faults that are caused by Network, Server Hardware or Operating System malfunctions are outside the scope of this Agreement. Network, Server Hardware and Operating System support is provided by the Customer or their chosen subcontractor.

6 Discussion

6.1 Lessons learned

To ensure an intuitive and effective User experience, two objectives were agreed:

- A seamless User Interface that would require minimum training as there was an unknown population of users with varied IT skills
- Meaningful visualizations that can answer key Use Cases and end-user questions.

To achieve these objectives, we found that the following tasks were crucial to the project's success:

1. **Detailed design of the layout and functionality of the dashboard**: High-level visualizations needed to be thoroughly redesigned to allow the visualization of edge cases that had not been expected at the time of writing the SoW. An iterative process of internal reviews of wireframes was put in place before development started, to ensure that the proposed solution met the objectives outlined above, as well as being clear to the development team.

2. **Agile Scrum works well:** Regardless of the amount of initial planning, User Story description and design specifications, regular Scrum meetings and an iterative development process is paramount to ensuring that development remains on track both from a functional as well as technical perspective.

3. **Iterative QA:** Iterative development must be followed by iterative QA to minimise technical debt and to ensure that built functionality matches specifications and performs as expected. Apart from system acceptance testing by experienced QA personnel, real-life testing by users who are not that familiar with the underlying data or system design is also important to ensure that end-user personas are represented as much as possible prior to release.

4. **Regular customer meetings:** Keeping project sponsors up to date with development progress, raising issues making them aware of risks, challenges and issues is fundamental to buy-in and also helps manage expectations. It is understandable that the final solution may differ from the original design; ensuring that all stakeholders are on board avoids delays and last-minute changes.

5. **Ensure that technical requirements are addressed:** Agile scrum focuses more on functional requirements, aiming to address user needs and to build systems that match user expectations in controlled environments. To ensure that the application functions as designed in multiple environments, technical specifications need to inform the chosen technology stack. In cases where compliance with technical requirements will result to reduced functionality or a suboptimal user experience, a balance must be achieved with all stakeholders' approval.

6.2 Challenges and Risks

From our analysis of the project we have identified a range of risks and challenges that may adversely impact the development of the software solution. Some of the potential risks and challenges are:

- Use on mobile devices and tablets
    - Responsive design and support for smaller screens is out of scope of this design.
    - Applying responsive design to the system will have to be managed as a System enhancement through a Change Request.
- Protection of IP
    - Intelligence data and content may be stored on a cloud hosted server and will be vulnerable to hacking attempts.
    - Clarivate Analytics will follow industry best practice in Server management, security updates and patching if hosting is provided as a service. It is understood that MLA will provide hosting services, therefore it is advisable that similar policies and procedures are followed by MLA, if hosting is provided by internal resources.
- Legacy browser compatibility
    - The application is built using the latest HTML 5 techniques to provide enhance User experience.
    - Clarivate Analytics has implemented a number of fall backs to ensure wide compatibility with legacy browsers. Unfortunately, compatibility with legacy browsers cannot be guaranteed.
- NPL dataset: multiple regions
    - The NPL dataset contains multiple regions, corresponding to the region or country of each of the authors. This is used to indicate research collaboration between universities from different regions.
    - As this would result in double-counting of publications, we chose to take into account the first region only, assuming that it corresponds to the primary.
- User interface improvement: zooming in.
    - In some cases and on smaller screens, data labels may be hard to see or clickable areas may too small.
    - If this becomes an issue, we could investigate a solution that could implement some type of zooming either automatically or with user intervention.
- Stacked bar chart: analysis by category
    - Depending on the use case, it may be useful to transpose the Year/category arrangement, having years on the Y axis and categories as stacked groups.