INFS 4020 – Big Data Concepts

Assignment 2: Big Data Strategy Proposal (SP5 2022)

DUE: By 11PM, Nov 4th 2022

General instructions:

- This assignment is worth 50% of your final grade. It is due **no later than 11 pm Nov 4th 2022**

- You will need to submit your assignment via learnonline. The file you submit needs to be in a **pdf** format and prepared **using the template provided**.

- The word limit for this assignment is 2500~3,000 words +/- 10%. Marks will be deducted if the assignment is too short (min 2,250 words) or too long (max 3,300 words).

- Similarity index for individual sources < 5%

- **Any late submission will attract a penalty of 10% per day, or part thereof, the assignment is late. The cut-off time is 11pm each day. No submission allowed after 3 days.**

Assessment task overview:

Photos by Balázs Kétyi and Carl Heyerdahl on Unsplash

You are required to develop a proposal for senior management of a selected organisation to implement a Big Data strategy designed to meet a specific business priority. Your proposal should include recommendations of Big Data technologies and suggest analytics plus and end-user application. You should justify your recommendations and include a high-level architecture diagram to show how the proposed technologies would fit together.

Assume that the audience know little about Big Data, but they want to make better use of their data which is why you have been invited to submit a proposal. However, your proposal is not just a sales pitch – you must demonstrate that you know what you are talking about, back up your arguments with evidence and communicate your ideas effectively.

This assignment is your opportunity to bring together the knowledge you have acquired in this course, apply it to a business scenario and further develop your communication skills.

Assessment task details:

From **Table 1** below, choose an organisation and the associated priority, whichever interests you the most. Then develop a proposal for a Big Data strategy to address the chosen organisation's key priority. In developing your proposal consider the following:

- Organisations have been named to provide some context for your proposal, but these are hypothetical scenarios. ***Please <u>do not</u> contact the organisations!***

- You do not have to research what your chosen organisation does with their data. Assume that they have no Big Data capability currently and develop your proposal accordingly.

- You are recommending what would need to be done, not actually doing it – i.e. you don't have to build any Big Data capacity or do Big Data analysis.

- You will need to make use of research. However, simply finding and presenting information of 'experts' will not be enough to earn a good grade. You will need to work things out for yourself and communicate your understanding of Big Data concepts and technologies.

Organisation	Key Priority
OTR (On The Run)	Deciding what is the optimal distribution of petrol stations in metropolitan Adelaide.
University of South Australia	Need to attract more international students to South Australia and compete with Eastern States.
Domino's Pizza	Needing to recruit staff who will stay and perform well.
JB Hi-FI	Improve online shopping experience

Table 1. Selected organisations and their key priorities

Understanding the priority:

Consider what questions would need to be answered to understand the key priority. See the Microsoft resources around the questions 'Is Big Data the right solution?' and 'Determining analytical goals'. A good starting point is here: https://msdn.microsoft.com/enus/library/dn749858.aspx

Note: These are Microsoft resources so naturally they suggest Microsoft technologies. You do not have to use those technologies in your proposal.

Data sources:

From the questions to be answered you need to identify the information needed and the data sources that can provide that information. It may be that some of the data would not exist (that you know of) so you would be recommending that it be collected.

Big Data technologies:

Once you know the data needed and where it would come from, decide which Big Data technologies would be appropriate to capture, store, process, clean, share, visualise and use that data. This should include suggested analytics and an end user application – would it be a predictive app or something else?

Provide brief descriptions of the technologies required to deliver the Big Data capability and give an example of one technology for each component of your strategy (e.g.: processing of streaming data – Apache Spark). The technology choices will depend on the data types in data sources you are recommending.

Don't focus on a specific tool or vendor. You can use the Big Data/AI Landscape diagram from Week 2 to investigate and recommend technologies.

High-level architecture:

Your proposal should include a diagram of a high-level architecture showing the different technologies and how they fit together. This, once again, is intended for a non-technical audience. Refer to resources from Weeks 9 to 11 for examples.

Big Data visualisation examples:

Provide two examples (screen shots with appropriate referencing) of Big Data visualisations to give the audience an indication of what you would be providing them (or if you had built a prototype). Explain the visualisations. The more relevant to the business priority and organisation the better. Importantly, these visualisations should be based on Big Data.

Benefits and adoption challenges:

Clearly articulate the benefits of the proposed strategy. You should also discuss any challenges of Big Data adoption and Big Data analytics, including data quality, privacy and security, in general and specific to your proposal. Include some recommendations of how your organisation could address these challenges in the context of the strategy you are proposing.

Executive summary:

An executive summary is a short document or section of a document produced for business purposes. It summarises a longer report so that readers can become acquainted with the contents of the report without having to read it all.

Write your executive summary after you have finished your Big Data strategy proposal. You should use short, concise paragraphs and write your executive summary in the same order as the full proposal.

Referencing:

Key resource is this website: www.unisa.edu.au/referencing. You should use the Harvard UniSA or APA referencing style.

Reference all diagrams appropriately, both in-text and in your list of references. This includes your visualisation examples. If you adapt a diagram for your purposes, acknowledge the original source.

Referencing is important for assignments to: (a) expand your knowledge of the assignment topic and (b) provide evidence to the claims you make and (c) demonstrate you know what you are talking about to make a convincing proposal and (d) provide other examples or case studies.

The general rule is if you are using information or data that is not of your own creation then you need to acknowledge it. This includes the screenshots and any data you use. Not only is this for academic integrity but to add weight to your recommendations – to show they are just not opinions.

The more you can back up your suggestions with research, examples, etc the higher mark you will receive.

How many references?

That depends on how many points you are making. Generally, more is better because you have used more sources to understand the topic and reinforce your points.

A minimum of 5 references is required. However, just adding as many references as possible without using them in the assignment will not earn maximum marks.

Do not plagiarise, i.e. do not copy directly from references without using quotation marks or without including a reference, and make sure that you follow the rules when paraphrasing.

Keep direct quotes to a minimum.

We want your understanding on the topic, not copied words from experts – this only demonstrates that you can research well, not apply your learning.

Reference quality:

The type (quality) of references makes a difference and this is considered in the marks as well. Feel free to use the course readings.

Avoid marketing/vendor sites and general websites - the quality is not assured because anyone can get a website up regardless of their expertise and marketing material from software companies is usually biased. The exception would be news sites when you want to report an event or where they are the sole vendor of a technology. References from global research companies like Gartner, Forrester, McKinsey incorporate insight gained from customers who have actually implemented technologies with various levels of success and failure. This breadth of experience, and in particular applicability to different industries, is well worth including.

Relying exclusively on Google to find references is a poor approach – try the library catalogue instead and include **_at least two_** academic sources, e.g. journal articles.

Since this is a fast-moving area, look for references from the last 5 years.

Presentation and structure:

The structure should be in a logical format that flows well. A sample template for the assignment is on the course website. Please use this structure – you can add to it with sub-headings if you wish.

Since this is proposal for a business audience, it should be presented in a professional format making it easy to read. An efficient layout is also important but do not spend too much time on making it look good and not enough time on the content.

Using bullet points are OK occasionally but you will need sentences for each point (i.e. just a bullet point list with no explanation is not suitable).

Word limit:

2500~3,000 words +/- 10% (minimum 2,250 words to maximum 3,300 words) overall

You need to include an executive summary **exactly one page** in length.

Marks will be deducted if the assignment is too short or too long. Keeping to a word limit requires a focus on what the audience most needs to know.

These are included in the word count:

- Executive summary
- Headings
- Direct quotes (you will gain more marks by writing using your own words than using lots of direct quotes)
- Diagram headings and captions

These are excluded from the word count:

- Title page
- Table of contents (if you add one)
- References
- Footnotes
- Text within diagrams

Other:

- Do not write in the first person ("I")
- Use formal language – this is a report intended for business.

Marking criteria:

The assignment will be marked on how well you cover the following:

Area	Weighting
Business priority questions and data sources	15%
Big Data technologies and high-level architecture	15%
End-user application and visualisation examples	15%
Benefits and challenges	15%
Executive summary	10%
Referencing • Correct referencing as per UniSA guidelines • Quality of references • How recent references are	10%
Use of formal business or academic language, including correct grammar and spelling	10%
Layout and professional presentation	10%