NZQA

NEW ZEALAND QUALIFICATIONS AUTHORITY
MANA TOHU MĀTAURANGA O AOTEAROA

QUALIFY FOR THE FUTURE WORLD
KIA NOHO TAKATŪ KI TŌ ĀMUA AO!

COMMON ASSESSMENT TASK

Level 3 Digital Technologies and Hangarau Matihiko, 2019

91908 Analyse an area of computer science

Credits: Three

Achievement Criteria		
Achievement	**Achievement with Merit**	**Achievement with Excellence**
Analyse an area of computer science.	Analyse, in depth, an area of computer science.	Critically analyse an area of computer science.

Type your School Code and 9-digit National Student Number (NSN) into the header at the top of this page. (If your NSN has 10 digits, omit the leading zero.)

Answer all parts of the assessment task in this document.

Your answer should be presented in 12pt Arial font, within the expanding text boxes, and may only include information you produce during this examination session.

You should aim to write between **800–1500 words** in total.

Save your finished work as a PDF file with the file name used in the header at the top of this page ("SchoolCode-YourNSN-91908.pdf").

By saving your work at the end of the examination, you are declaring that this work is your own. NZQA may sample your work to ensure that this is the case.

YOU MUST HAND THIS BOOKLET TO THE SUPERVISOR AT THE END OF THE EXAMINATION.

Achievement
03

INSTRUCTIONS

Read all parts of the assessment task before you begin.

Select ONE of the following computer science areas:

- complexity and tractability

- computer vision

- big data

- computer graphics

- formal languages

- network-communication protocols.

Type your chosen computer science area in the space below:

Big Data

Begin your answers on page 3.

ASSESSMENT TASK

(a) Explain the key aspects of your chosen computer science area.

Big data is data that cannot be processed by conventional processing systems. It requires the use of special processing systems to be analyzed, which could consist of multiple computers and multiple systems.

Key aspects of big data are the 5V's: Velocity, Volume, Variety, Veracity and Value.

Velocity refers to the speed of which the data is created. Data is created all the time and at different speeds, sometimes a small amount of data may be created in a large period of time and sometimes huge amounts of data may be created in a short period of time. The data system must be able to cope with the changes in the amount of data being generated.

Volume refers to the amount of data that is created. Vast amounts of data are created every second, which need to be processed in real time or stored so that they can be processed later. Therefore storage is an essential part of big data, the system must be able to store the data.

Variety refers to the different types of data that are being processed. Big data can come in various forms, it can be structured, unstructured, in different formats and from different sources, the data must be put into an appropriate structure so that it can be processed. Therefore the data system must be able to handle all types of data.

Veracity refers to the reliability of the data, if it can be used and how reliable it is. Vast amounts of data are created and not all of it is reliable, therefore we must decide which data is reliable and which is not.

Value refers to the usability and the application that the data holds. So much data is created that we must decide which data is of value and which is not and only collect data which has a value or use, otherwise it is a waste of storage and money to store and process data that has no value.

(b)   Explain the relevant algorithms or mechanisms that support your chosen computer science area.

The mechanisms behind big data are its collection of data from it source whether it be from social media data, banking data, health data or communication data. Once the data is collected it must be processed so that value can be extracted from it, in its unprocessed form it has not value to anyone. Data arrives in various forms and must be formatted so that it can be analyzed so that something of meaning can be extracted. Once it is analyzed trends and patterns will start to emerge that can be of use to the organization that collected the data. These patterns or trends can then be used to provide personalized services to individuals, to create new products and to provide insights to companies.

(c) Explain how your chosen computer science area is used, implemented or occurs. Use examples to support your answer.

Big data is used in various applications, it can be used for analytical purposes, new product development, to provide a personalized service, etc. When data analyzed it reveals trends which can be used for a purpose.

An example of big data being used to provide a personalized service. Netflix has many images that can be used to convince users to view a program or movie and displays the image that best matches their previous viewing history. E.g. Netflix has data that an individual has viewed multiple romantic movies, Netflix might display an image that that has a couple in it to convince this user to view this movie, this provides a personalized service which would not have been possible without this data being available.

Another example is that data that is collected from google searches may reveal that people are looking for a certain product that is not yet available. E.g. Many people are searching for smart home. Google may be able to use this data to develop new products or services. This provides valuable insight to companies that develop new products, which would otherwise not exist without this data being available.

(d) Explain the key problems or issues related to your selected computer science area, AND how these have been, or may be, addressed.

A key problem with big data is that as more people create data and more things are created that collect data, new systems must be developed to handle this data all the time to ensure that this data can continue to be processed, analyzed and used. If data systems do not keep pace with the current generation rate of data than data systems may not be able to handle the future volumes of data. This may be addressed by ensuring that data systems keep pace with the current generation rate of data.

Another problem is the privacy implications behind the data, the data is collected on individuals is vast and may include their shopping habits, spending habits, internet usage habits, what they like, etc., and may be private to the individual. This has been addressed by new privacy laws that limit what can be collected and that users must be informed that their data is being collected and for what purpose.

(e)   Provide a detailed explanation of how the technical capabilities and limitations of your chosen computer science area relate to humans. Use examples to support your answer.

Big data is collected on humans and used to provide insights that are of purpose to humans. This big data is only valuable because it can be targeted to a certain human to provide a personalized service or product, if it was not targeted to a certain human than it would have no purpose.

(f) Compare and contrast different perspectives on your chosen computer science area.

> Big data perspectives consist of people who believe that it is fine that the data is being collected because it improves the services and goods that they use and therefore it is a good thing because it make their life easier.
>
> Another perspective is that our habits should not be tracked, stored and used because it breaches our privacy and we do not know what the data is being used for, therefore these organizations should not be able to collect this data. Because it enables them to know what we like, don't like, where we go, who we are friends with, etc. It is too much for an organization to know about us and is a breach of privacy.
>
> Another perspective is only data this is necessary should be collected and used, if it does not provide a purpose for providing the product or service it should not be collected or used, no matter how valuable to data is to the organization.

(g) What conclusions can you draw about your chosen computer science area?
In your answer, you could:

- explore less-obvious implications

- justify predictions that you make

- consider potential improvements

- suggest innovative and imaginative wider uses.

Big data is a complex topic and has very useful purposes, whether it be for providing a personalized service, personalized product, new product, new service, providing insights, etc. Big data is a very valuable tool to organizations and enables them to know what their customers and users want. Innovative uses include being able to judge what movie a Netflix user would like to watch next based on their previous viewing history, another innovative use would be to use location data history to locate someone that has been lost. An improvement would be to ensure that the data that is being collected is not private, so that users' privacy is not being breeched.

Achievement Exemplar 2019

Subject	Digital Technologies		Standard	91908	Overall grade	03
Q	**Grade**	**Annotation**				
		The candidate demonstrated a limited understanding in analysing their chosen area of computer science. The key aspects were described using the so-called 5Vs (i.e. velocity, volume, variety, veracity and value). Discussions of Algorithms and / or Mechanisms were partial. The candidate showed some understanding of the concepts, and some of the problems, within their chosen area of computer science. However, their analysis was weak or partial throughout. A grade of Achievement was appropriate.				