

No part of the candidate evidence in this exemplar material may be presented in an external assessment for the purpose of gaining credits towards an NCEA qualification.

# 3



NEW ZEALAND QUALIFICATIONS AUTHORITY  
MANA TOHU MĀTAURANGA O AOTEAROA

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

COMMON ASSESSMENT TASK

## Level 3 Digital Technologies and Hangarau Matihiko 2021

### 91908 Analyse an area of computer science

Credits: Three

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 space below. (If your NSN has 10 digits, omit the leading zero.) It should look like “123-123456789-91886”.

There are three questions in this document. **Choose ONE question to answer.**

Make sure you have the PDF of the Resource Booklet 91908R. This contains resources for Questions Two and Three.

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

Your answers should be presented in 12pt Times New Roman font, within the expanding text boxes, and may include only information you produce during this assessment session. Internet access is not permitted.

**Save your finished work as a PDF file** as instructed by your teacher.

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 this is the case.

**Merit**

**TOTAL**

**05**

ASSESSOR'S USE ONLY

## INSTRUCTIONS

There are three questions in this assessment, on the topics of:

- Formal languages ([page 3](#))
- Network communication protocols ([page 9](#))
- Big data ([14](#)).

**Choose only ONE question to answer.**

Questions Two and Three require you to refer to the separate resource booklet.

Read all parts of your chosen question before you begin.

**OR: QUESTION THREE: Big data**

This question includes references to **Resources E to G** on pages 4 and 5 of the resource booklet.

- (a) Briefly describe the “3 Vs” of big data (**volume, variety, and velocity**).

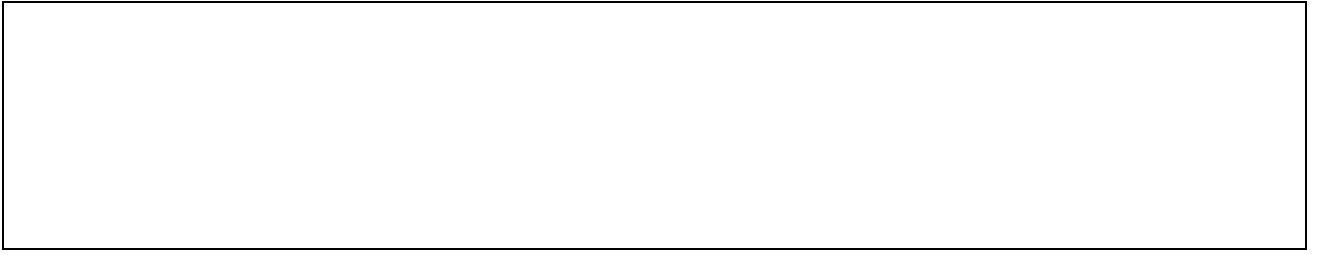
The concept of big data is gathering, storing and processing massive amounts of data, this data is then used to by corporations in order to predict outcomes and figure out a user’s, or the masses personal preferences. Volume is the amount of data that is being gathered, the variety is the types of data that is being gathered weather is be structured unstructured, personal or public. And velocity is the speed at which that data is able to be transferred from its source to a data warehouse/server.

- (b) What key problems or issues relating to the 3 Vs can you see with the collection of the dataset used to generate **Resources E and F**? How might these problems or issues be addressed?

Concerning the velocity of the data is resource e and f, it is sub optimal, but then again it is always such. Real time big data is near imposable. But with our realistic data transfer speeds there will be a lag in the time the data is collected and when it is processed and used This means that while the data is the latest data we have, it is not the most recent data collected data is still data that is being processes, and or transferred from the user’s device to a data warehouse. Addressing this issue is as simple as creating faster better machine learning algorithms in order to process the data faster and gets closer to the ideal o Realtime big data. Another way to address the issue of the velocity of the data is to increase transfer speeds of the data. This is one of the main reasons for the push for faster and faster mobile data. The reality is 4G is well faster enough for the average internet user, but the big push for 5G and next will be 6G is the increase in data transfer speed, specifically to benefit the collection big data. So, a faster connection such as 5G over 4G is another way to address the issue of velocity in big data.

Volume of data is also another concern the amount of data collected thought the GPS systems on people’s phone is massive, and to store and process all of that data is a mammoth task. To address the issues concerning the volume of the data we can, as before, try and build better machine learning algorithms in order to process the massive amounts of data that are being collected. This make the lag from the time of collection to having real useful statistics less and helps to deal with the massive amounts of incoming data more quickly and efficiently. Another concern with the volume of data is the opposite, not having enough data. While having too much data can pose problems with how we can manage that data, not having enough data make that data less accurate and as google says less ‘anonymous’ and less of an aggregate view of trends. So when there is not enough data to create a well rounded idea of the current trends that data that has been collected essential becomes useless as is the case for the data collected for the transit stations.

In reference to the variety of the data, I can’t imagine too many problems with it, but a couple that come to mind are the fact that this data set only tracks the data of those with a google account and or those who have not disabled the GPS data collection feature. So those who fit these categories are not included in the resulting graphs, and this leaves out a portion of the population. As well as people who went to a smaller supermarket that may have not been one of the ones the algorithm driving the data had registered, in this case, someone who went to one of these places would not have been included in the dataset despite going out and to a supermarket.



- (c) With reference to **Resources E to G**, how do the technical capabilities and limitations of big data affect people's lives? Give examples to support your answer.

The big data that has been collected over the lockdown, of the locations of people is able to have an impact everyone's lives, a negative of this is well, we are being tracked, while this tracking may not be used on a personal level is just something to think about, it may be a scary fact to many that google is able to use the data from your GPS, which you properly didn't even know was actively sending data to google in, and use it in conjunction with everybody else's data to create these trend graphs. Keeping in mind that they are not only using the data they collect for good. Google has a business in selling your data.

But in this case the data being collected and used for a good reason, the technical capabilities of big data can be very useful and, in this case, may be able to save lives. Google says that the big data that they have collected can be 'used by health officials and others to understand people's response to social distancing guidelines' The data that google have collected can show officials the trends in the amount of people going to places like cafes and supermarkets at a given point, and this may be able to give them an insight as to better ways to organise lockdowns, and or how effective the current ones are.

Another concern with the technical capabilities of this big data analysis is that people may have unwillingly signed up to be tracked. The collection of big data is so valuable, why would google make you have to opt in to be one of the sources for this data, when they can just and legally make you have to opt out. They just need to make it clear that by signing up to google you essentially forfeit your right to any privacy. And to know this take reading the huge privacy policy and terms and conditions that I don't think anyone has ever read. So the lack of privacy in an age where big data is so valuable, should be a great concern of everyone. And it is not just google that has access to your data, as I stated before Google is a company with a business in *selling data*.

And the limitations of big data namely velocity. With the velocity, it is in the best interest of corporations to increase the speed for data transfer between your phone and their server, this has the subsequent effect of making your mobile data connection on your phone much faster. The real push for faster mobile data is not for the consumer but to the corporation are trying to increase the velocity of their big data. But this has a great effect for the consumer, a quicker advance in mobile data technology. While there may be no need for the blazingly fast 5G connection on the consumers end it is a little nice to have.

- (d) With reference to **Resources E to G**, what are some of the benefits and disadvantages of big data? Discuss whether you think it is possible to find a balance between these benefits and disadvantages.

A couple of benefits of big data are that it is increasing the rate at which the speed of connectivity increases. It is also helping to advance our machine learning algorithms. The future is in machine learning for pretty much everything, and the fact that machine learning goes hand in hand means that if we want to advance our big data sector, we must do the same for machine learning. Right now, big data is really driving advancements in machine learning. Because big data is useless if it is stored on a server with no way to derive real, useful insights from the massive amounts of data. Big data is paving the way for the future of everything really, it is a great tool for companies, and if done right can have huge benefits for both the producer and the consumer.

A disadvantage of big data is that it gives companies a great reason to want to know as much as possible about you, your habits and your life. This creates a real conflict of interest between the

consumer and the company. Google wants to collect as much data on you as possible, but I can assume that the average consumer properly does not want all of this data collected on them. But there is a balance to be found here, and it is already occurring, people just turn a blind eye to all of the data collection, and uncaringly or unknowingly commit all of their data to be collect and used. Google says that the data collect is anonymous. And that no data that is identifiable to any individual person is displayed, while individuals may not be identifiable from the data that google publishes publicly, there is no doubt that google themselves have the raw data available to them for purposes mainly concerning advertising.

Another benefit of big data is the information that we are able to get out of it. The speed at which we can process big data has never been so fast and it is only getting faster. The information and insights that we are able to derive from big data sets in invaluable to company's who can use this data to target adverting, to suggests products, to suggest videos as is the case on YouTube, and to suggest movies and tv series as is the case with Netflix and Google to auto-complete your searches and find trends in the mobility of people during a Covid lockdown. The trends found and the insights derived are so important to these company's, they help to drive revenue and in the case of the resource benefit public health.

To be honest I think that a balance between the positives and the negatives has already been archived, large corporations are able collect and analyse data in ways which benefit their company, they are able to use the information the collect in order to drive revenue and profit, and they are able to do it without too much concern for the privacy of the masses, because the reality of it is the general population is largely unknowing and or uncaring that their 'personal data' is being collected. But this uncaring for personal privacy is not a given, if company's go too far people will begin to realize and care, but at the moment companies are at a happy equilibrium with most of the population.

The number of internet-connected devices is increasing globally. In 2018 there were approximately 8.4 billion networked devices. A forecast at the beginning of 2020 predicted that this number will increase to 29.3 billion by 2023, and that 71 per cent of the global population will be mobile subscribers by then.

Two-and-a-half quintillion ( $2.5 \times 10^{18}$ ) bytes of data are created daily. This could potentially rise even more thanks to better internet access. Today, we can collect big data from every imaginable field we're able to monitor digitally, including science, commerce, medicine, government, and sport, just to name a few.

Sources (adapted):

Hill, K. (2020). Connected devices will be 3x the global population by 2023, Cisco says. RCR Wireless News, 18 February 2020.  
<https://www.rcrwireless.com/20200218/internet-of-things/connected-devices-will-be-3x-the-global-population-by-2023-cisco-says>

Mills, T. (2018). Big data is changing the way people live their lives. Forbes, 16 May 2018.  
<https://www.forbes.com/sites/forbestechcouncil/2018/05/16/big-data-is-changing-the-way-people-live-their-lives>

- (e) With reference to the statement above, discuss how a particular organisation or technology sector you have studied this year uses the concept of big data. What are some future uses or improvements you think are possible in this area, and what are the implications of these?

Self-driving cars are a huge one. Big data is an integral part of having self-driving cars work safely and efficiently. Self-driving cars use big data in order to teach their machine learning algorithms to drive. Data that is collected from the various sensors on the cars is sent back to the company who made the car in order to be analysed to make the car safer and more reliable. The sensors on the cars, cameras, depth sensors, motion sensor, accelerometers all send data back to a server where it is used to help the machine learning algorithms learn and perform better in a given situation. The more self-driving cars we have on the road the more data can be collected, analysed and used to make the software on the cars, safer and more effective. Tesla for example gathers all of the data from the cars they have on the road, analyse it to figure out the problems with the software on the car, rectify the issues, and change a few things about, and then that gets packaged up and sent as a software update to all of the tesla's. The data collected is what drives the machine learning algorithms that drive the cars so the more data they are able to collect and process the safer and more advanced the cars will be.

In the field of big data the two main improvements to be made are the volume and the velocity of data. As mentioned before the more data that is collected the more information we have for programs to learn and advance. But this is in contradiction with what is ideal for the consumer. As alongside more data collection comes more effective machine learning and better insights into a company comes less privacy for the average person.

The other improvement to be made is in the processing and transfer of the data, in other words the velocity. The faster we can collect and process the data the more up to date and useful that data is. Striving to increase the velocity of the data drives the advancement of the machine learning algorithms that analyse that data, as well as the transfer speeds of that data. Striving for this has the implications of an increased dependency on machine learning as well as better more effective machine learning, which is the future of commuting, and the use of big data is driving advancements in machine learning, as well as the speed of data transfer.

## Merit Exemplar 2021

Subject	Level 3 Digital Technologies		Standard	91908	Total score	05
Q	Grade score	Annotation				
3	M5	<p>The candidate has completed all required components of the assessment. The response to part (b) covers the evidence expected in the otherwise too-brief response to part (a).</p> <p>The candidate has interpreted the resources in parts (c) and (d) and attempted to link the information back to the computer science concept. While the responses include some generalised or incorrect statements, these do not detract from the overall understanding demonstrated.</p> <p>The response in part (e) is too generalised, meaning the answers as a whole do not meet the criteria for Achievement with Excellence. The example of the self-driving car is a good one, and while the candidate has identified ways in which concepts relating to Big Data are relevant, they have not adequately explained or discussed how these could lead to future improvements, or the possible implications of these.</p>				