BIG DATA

WHAT ARE YOU DOING?

SHH... I'M TRYING TO *COUNT* THE *STARS*.

YOU'RE *CRAZY!* LOOK AT ALL THE STARS! THERE MUST BE *MILLIONS* OF THEM UP THERE, AND THESE ARE ONLY THE ONES WE CAN SEE -- THERE ARE GALAXIES OUT THERE THAT WE CAN'T SEE BECAUSE THEY ARE TOO FAR AWAY!

COUNTING THEM ALL WOULD TAKE FOREVER! ...AND OUR POOR *BRAINS* CAN'T EVEN STORE THAT MUCH INFORMATION.

OUR BRAINS MIGHT NOT BE ABLE TO, BUT *SUPER-COMPUTERS* CAN. THE *SKA* USES SUPER-COMPUTERS TO PROCESS ALL THE *DATA* THEY RECEIVE FROM THE STARS.

CHALLENGE: A UNIVERSE OF DATA

THE *SQUARE KILOMETRE ARRAY (SKA)* IS THE MOST AMBITIOUS SCIENCE DEVELOPMENT IN HUMAN HISTORY. HUNDREDS OF RADIO TELESCOPES AROUND THE WORLD ARE GOING TO BE POINTING TO THE SKY TO OBSERVE OUR UNIVERSE. BUT THE UNIVERSE IS A BIG PLACE, AND THAT'S A LOT OF DATA TO TAKE IN, WHICH THEN HAS TO BE PROCESSED AND STORED. THIS IS CALLED *BIG DATA*, AND IT REQUIRES *SUPERCOMPUTERS* TO PROCESS IT.

ACTIVITY: COLOUR BY NUMBERS

FOR EACH ROW ON THE GRID, COLOUR IN THE SQUARES INDICATED UNDER 'DATA'. WE'VE DONE THE FIRST 3 ROWS FOR YOU!

DATA

1: –	**13:** I – Q
2: F – H	**14:** I – Q
3: E – I, L – M	**15:** I – P
4: D – N	**16:** I – P
5: C – N	**17:** I – P
6: B – O	**18:** J – P, R
7: B – O	**19:** J – P, R
8: B – P	**20:** J – O, R
9: B – Q	**21:** J – O
10: C – S	**22:** J – N
11: C – S	**23:** K – M
12: D – E, H – R	**24:** K – L
	25: –

WHAT'S HAPPENING HERE?

Did you know that you have just been *processing data*? The strings of numbers are data, and you processed them to produce an image!

Data is quantities and figures collected about something, which can then be *processed into information*. When you send a message on Whatsapp, you are *transmitting data via radio waves*. The phone you are sending it to will then receive the data and process it to perform the function it is being told to do: these functions include things like displaying a text message, or displaying an image, or playing a song, and so on. The more *complex* the function, the more data it uses, and the longer it takes to send to other devices.

Cell phones and computers also *store* this data, so you can look at it again later.

BYTE (B): 1 BYTE
SMS: ABOUT 190 B

KILOBYTE (KB): 1000 BYTES
PICTURES: 100 – 900 KB

MEGABYTE (MB): 1000,000 BYTES
SONGS: ABOUT 3 – 4 MB

GIGABYTE (GB): 1000,000,000 BYTES
MOVIES: UP TO 5 GB

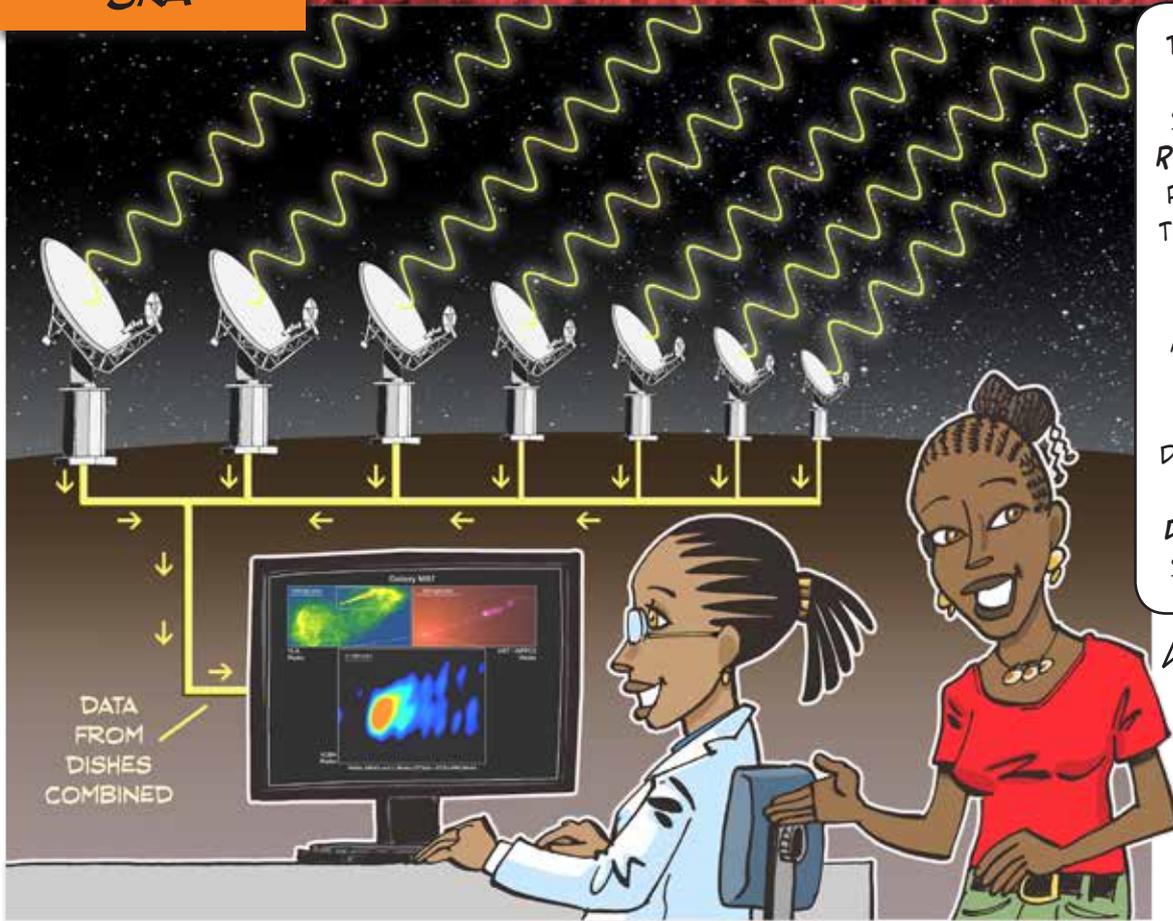TERABYTE (TB): 1000,000,000,000 BYTES
STORAGE SPACE ON A COMPUTER: UP TO 3 TB

PETABYTE (PB): 1000,000,000,000,000 BYTES
DATA THAT THE SKA WILL GENERATE EVERY DAY: 960 PB!

SCIENCE SPAZA

THE **SKA TELESCOPES** RECEIVE DATA FROM SPACE IN THE FORM OF **RADIO WAVES**, EMITTED FROM PULSATING STARS. THERE WILL BE HUNDREDS OF SKA TELESCOPE DISHES, AND THEY WILL ALL BE RECEIVING HUGE AMOUNTS OF DATA CONSTANTLY. WHEN DEALING WITH THIS MUCH DATA IT IS CALLED **BIG DATA**, BECAUSE OF THE SHEER QUANTITY OF IT.

DATA FROM DISHES COMBINED

Various functions then need to be performed, such as removing the interference, calibrating all of the data and finally turning it into an image that we can see (like the MeerKAT First Light image). Normal computers aren't powerful enough to do this; in fact, it would take 100 million normal computers to do all of this! For this reason, the SKA will need to use *supercomputers*.

Supercomputers are huge machines with a lot of processing power, and are designed specifically to handle huge amounts of data.

The supercomputer needed for the SKA will be three times more powerful than any existing computer - that will make it the most powerful supercomputer in human history!

CAREERS:

Computer and Information Research Scientists

are experts in computer technology. They develop new ways to use computer technology and improve existing computing technology.

Computer Programmers

write the code for software programs, which tell computers how to process information and perform specific tasks.

Shagita Gounden from Pretoria is a signal analyst for the Council for Scientific and Industrial Research (CSIR) and is a Computer Engineer at the Square Killometre Array South Africa. Gounden obtained an Honours degree in computer engineering from the University of Pretoria.

CURRICULUM LINKS

- **GRADE 11 & 12:** Computer Applications Technology — Solution Development (Databases)

- **GRADE 11 & 12:** Computer Applications Technology — Systems Technologies (Hardware)

- **GRADE 10, 11 & 12:** Computer Applications Technology — Information Management (Process Data and Information)

Knowledge is Ncah!

SCIENCE SPAZA

PUZZLE YOUR MIND!!!

CALCULATIONS

IF EACH MINUTE OF MUSIC IS ROUGHLY 1 MB, FILL IN THE BLANK SPACES TO WORK OUT HOW LONG IT WOULD TAKE TO PLAY BACK ALL THE DATA RECEIVED BY THE SKA IN A DAY. YOU MIGHT NEED A CALCULATOR!

960 PB = _____ MB

_____ MB of music would be roughly _____ minutes.

_____ minutes ÷ _____ (minutes per hour) = _____ hours.

_____ hours ÷ _____ (hours per day) = _____ days.

_____ days ÷ _____ (days per year) = _____ years.

Answer = _____ years!

Answer = 182648 years

WE WANT YOUR FEEDBACK!

We want to hear from you! Send us a picture of you doing the activity, or a picture of your completed puzzle, or just any news from your science club.

You can send us your feedback in any of the following ways: Whatsapp or SMS number 📞 **076 173 7130**; email us at **info@sciencespaza.org**; Facebook us at **ScienceSpaza** or contact us through our website **www.sciencespaza.org**

Remember to include your name, age and gender, as well as the name of your school and name of your science club. Also say which Activity Resource you are replying to.