1959
Data set of **1,000 cases**

1965
Data sets of **10,000 cases** (e.g. observations)

1978
technology to analyse **500 to 5,000 variables** at the same time

1981
Statistics for **20,000 samples,** as histograms with 800 cells

1981
Software to handle **88,000 variables**

1982
Moderate data sets have less than **500 cases** and large more than **2,000**

1986
Very large data sets referred as **11,000 cases**

1987
50,000 points, but the representation of 1 million or more data is feasible

1990
5,000 cases with 6 variables for immediate evaluation (c. 3 seconds.)

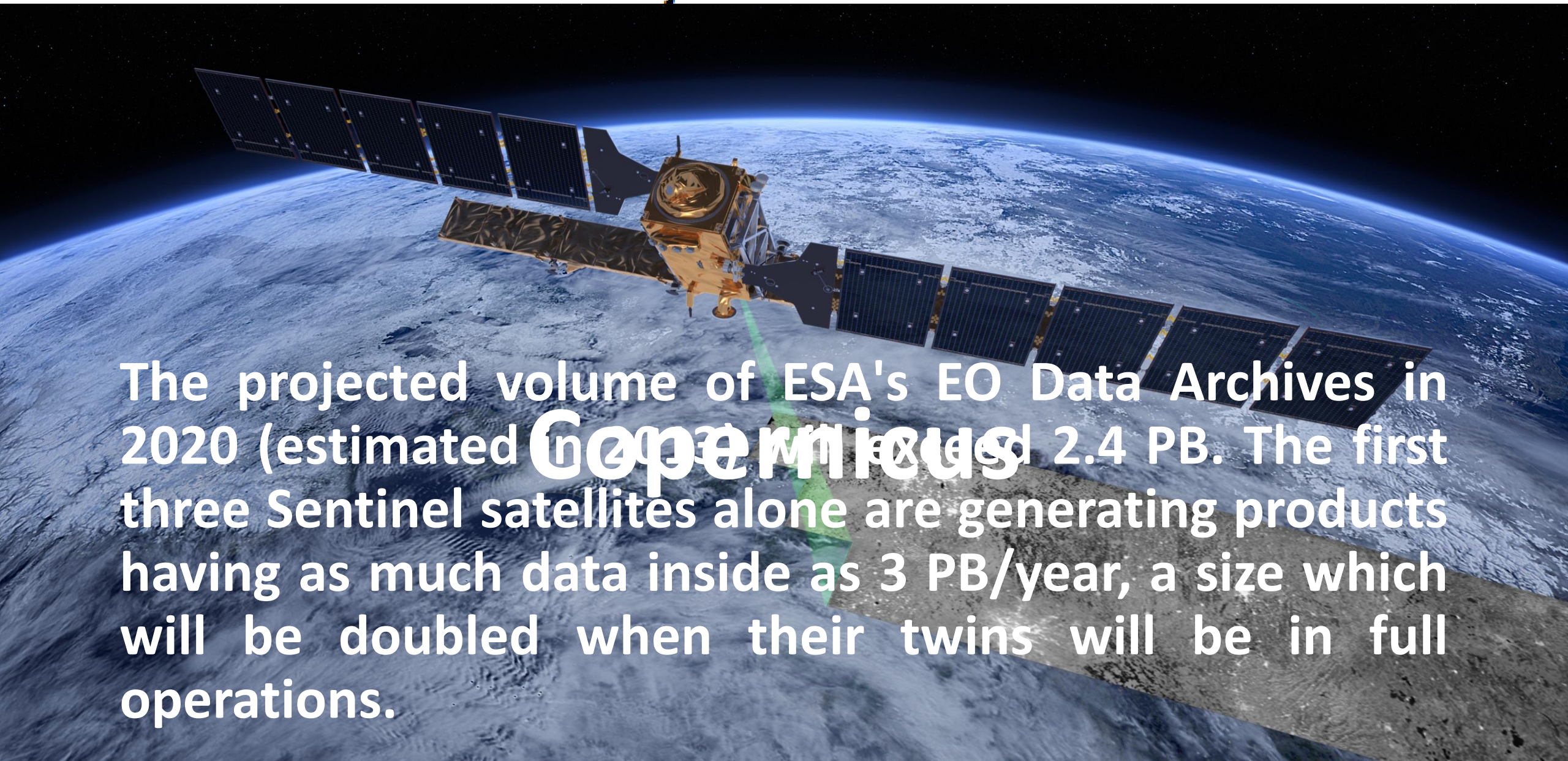
1991
1 million cases a very large data file

1996
Large data sets have **60,000 or more observations**

2006
1 million (bytes, cases, variables, combinations or tests) as a guideline to refer to a dataset as 'large'

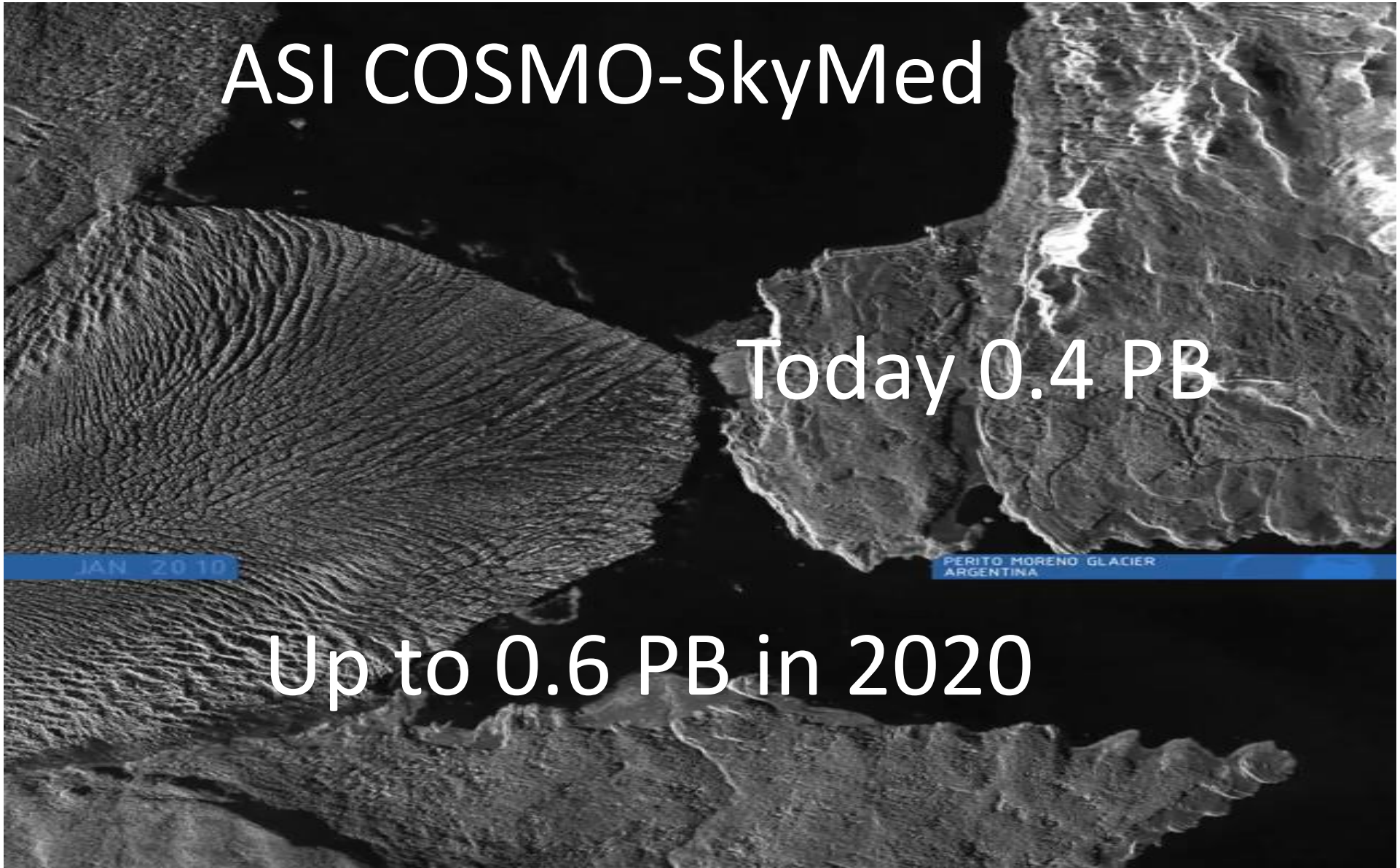
Extensive datasets are not today's phenomenon

Evolution of ESA's EO Data Archives between 1986-2010 and future Projections

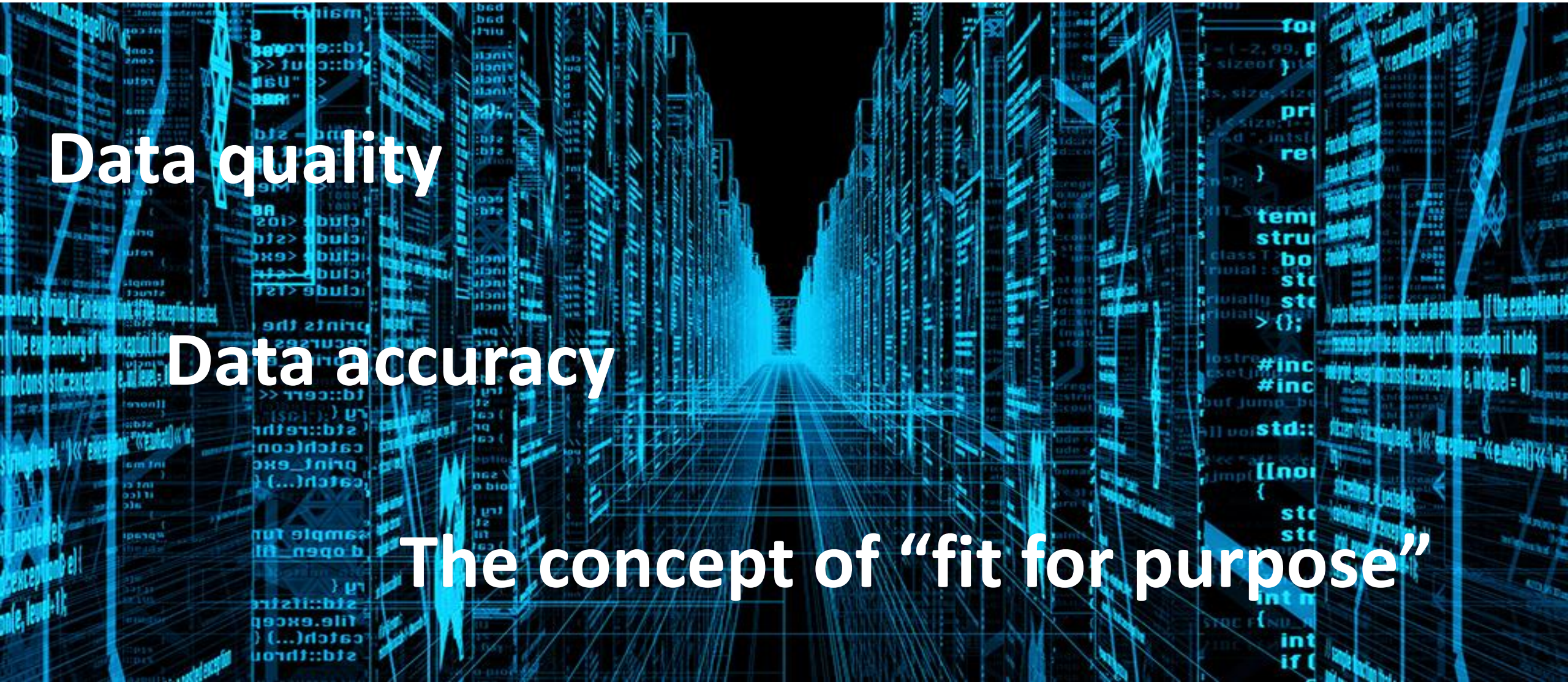


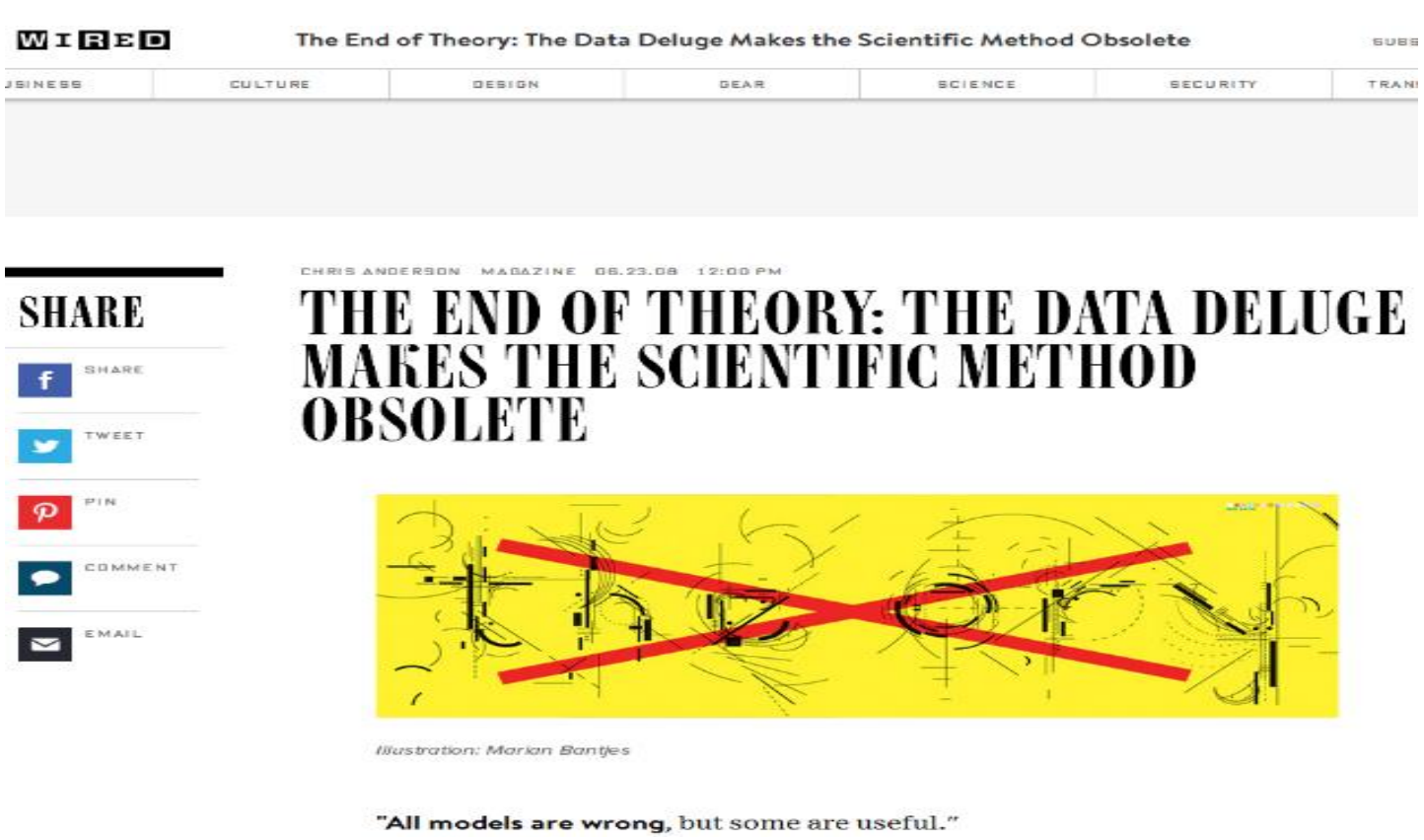
The projected volume of ESA's EO Data Archives in 2020 (estimated in 2015) will exceed 2.4 PB. The first three Sentinel satellites alone are generating products having as much data inside as 3 PB/year, a size which will be doubled when their twins will be in full operations.

Volume of Our Big Data



Veracity of Big Data





WIRED The End of Theory: The Data Deluge Makes the Scientific Method Obsolete

BUSINESS | CULTURE | DESIGN | GEAR | SCIENCE | SECURITY | TRANS

SHARE

SHARE
TWEET
PIN
COMMENT
EMAIL

CHRIS ANDERSON MAGAZINE 06.23.08 12:00 PM

THE END OF THEORY: THE DATA DELUGE MAKES THE SCIENTIFIC METHOD OBSOLETE

Illustration: Marian Bantjes

"All models are wrong, but some are useful."

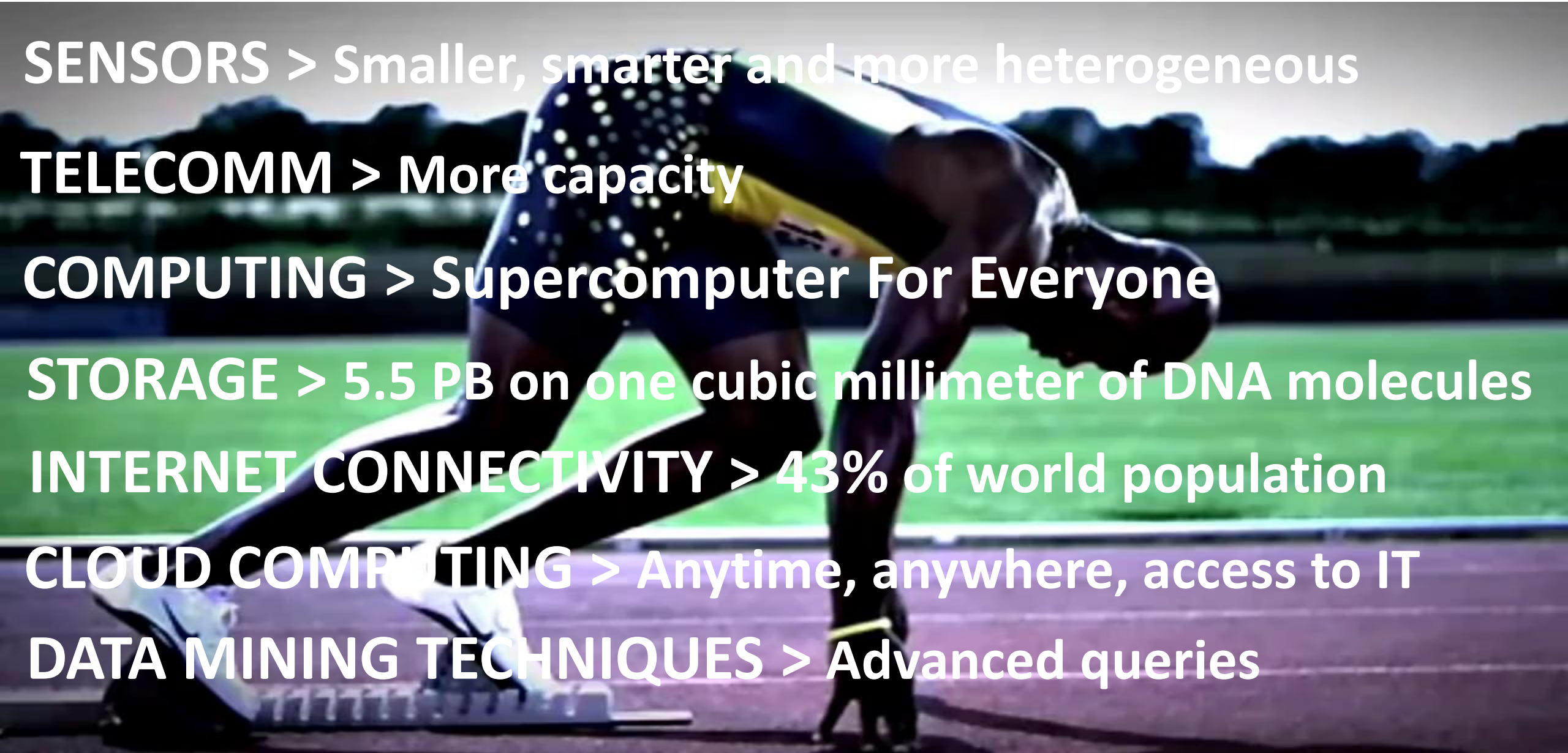
Big Data

Improve the ability to analyze and shape the real world.

Big Data enable decision making process.

Value of Big Data

Techniques of Big Data



SENSORS > Smaller, smarter and more heterogeneous

TELECOMM > More capacity

COMPUTING > Supercomputer For Everyone

STORAGE > 5.5 PB on one cubic millimeter of DNA molecules

INTERNET CONNECTIVITY > 43% of world population

CLOUD COMPUTING > Anytime, anywhere, access to IT

DATA MINING TECHNIQUES > Advanced queries