Big Data
in Arctic Studies

Alexei D. Gvishiani

**Chief Scientist, Geophysical Center of the Russian Academy of Sciences (RAS),
Chair of the RAS Scientific Council for Arctic and Antarctic Studies**

**IIASA Regional Conference "Systems Analysis in Eurasia"
April 13–15, 2021**

Birth of Big Data

- The author of the term **"Big Data"** is Clifford Lynch, editor of "Nature" journal. On September 3, 2008 he issued a special volume of the journal on the topic "How can technologies that open up high possibilities to work with big data influence the future of science?".

Explosive growth of data volume and variety → **Leap from the amount of initial data to the quality of recognizable knowledge**

A 21st Century Revolution in Data Science

- The transition of the quantity of the accumulated information into the quality of the problem solution is called Big Data (BD) phenomenon. Today it is the most discussed concept in the IT science and industry.

- Data are no longer seen as a static value. They have become the material for creating new economic values.

- In the 2000s, software technologies MapReduce (Google) and Hadoop (Yahoo) have appeared. For them, the original information no longer needs to be lined up in neat database tables. In result the Big Data processing technologies emerged without the previous rigid hierarchy and homogeneity.

Big Data Research Paradigm

- In the Big Data world, the search for causality will have to be abandoned in favor of the search for correlations. The data processing in order to answer the question WHY? is replaced by the search for the answer to question WHAT IS IT?

- This change in the research paradigm shortens the transition to prediction.

Big Data. Definition

- **Horizontal scaling** is dividing the system into smaller structural components and spacing them into separate physical machines and / or increasing the number of servers, nodes and processors that simultaneously perform the same function.

- **Big Data** is structured and unstructured data of **huge volumes** and **significant diversity**, efficiently processed by horizontally scalable software, alternative to traditional DBMS.

Systems Analysis Methods to Extract Knowledge from Big Data

- Artificial neural networks, network analysis

- Optimization, genetic algorithms

- Image recognition, dynamic pattern recognition

- Predictive analytics

- Data mining, cluster analysis

- Simulation

- Spatial analysis

- Statistical analysis, time series analysis

- Visualization

THE INTERNET. Basic Big Data example.

2020 This Is What Happens In An Internet Minute

- **Facebook – 1,3 million logins**

- **Instagram – 694,444 people scrolling**

- **Apple Messages – 19 million texts sent**

- **YouTube – 4,7 million videos viewed**

- **Twitter – 194,444 people tweeting**

- **Email – 190 million emails sent**

- **WhatsApp – 59 million messages sent**

- **NETFLIX – 764,000 hours watched**

Sources of Big Data in Earth Sciences

- Meteorological Data

- Earth remote sensing Data

- Data ecological superstations SMEAR II information

- Oil and gas seismic surveys

- Seismograms monitoring

- The Arctic comprehensive information is a comings source of Big Data

# Polar Circle Big Data, 66 33' Northern Latitude

- The area of the Arctic Circle – **21 million sq. km.**

- Circumference of the Arctic Circle – **15,948 km**

- Population – **4,6 million people**

- **2,5 million** live in the Russian Arctic

– the warming rate is **5 times higher** than the average on Earth

Arctic Exploration RF Today

Characterized by the developments in the areas in which the Arctic undeniably demonstrates its usefulness:

1. Mining of the metal, gold and diamond deposits (14% of global nickel and 41% of palladium production)
2. Production of about 85% of Russia's gas
3. Development and production of mineral fertilizers
4. Northern Sea Route as a unified infrastructure of the Northern Sea Transport Corridor
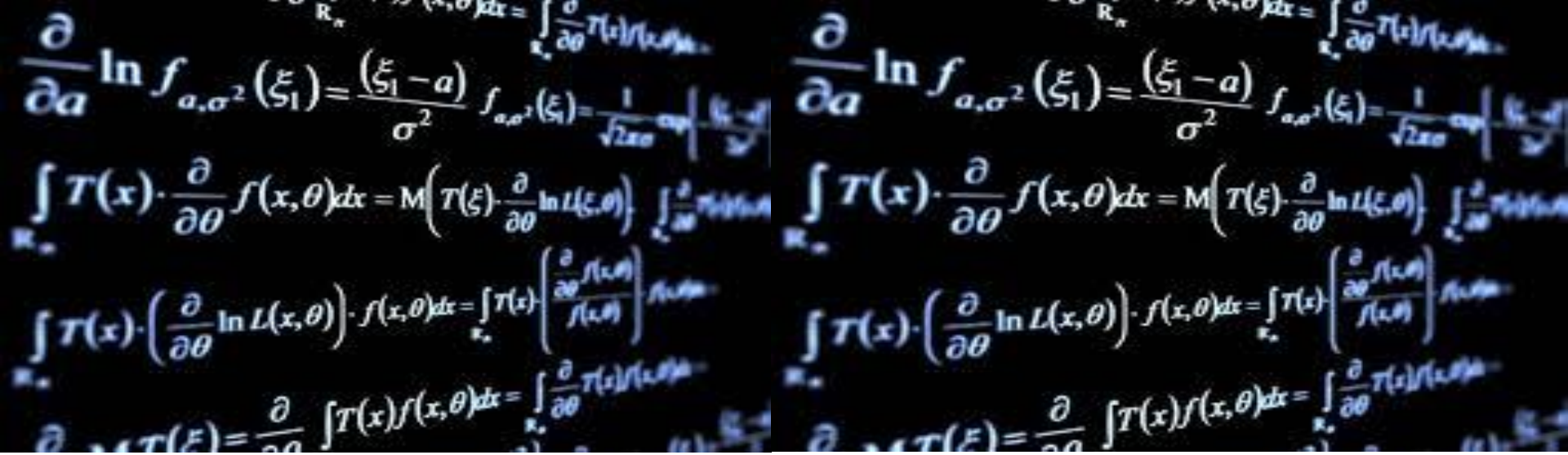
«The highest purpose of mathematics is to find order in the chaos which surrounds us»

Norbert Wiener (1894 – 1964)