



Rola państwowych instytucji w erze Big Data oraz AI

Jacek Leśkow
NASK

Nowa umowa społeczna cywilizacji cyfrowej

2

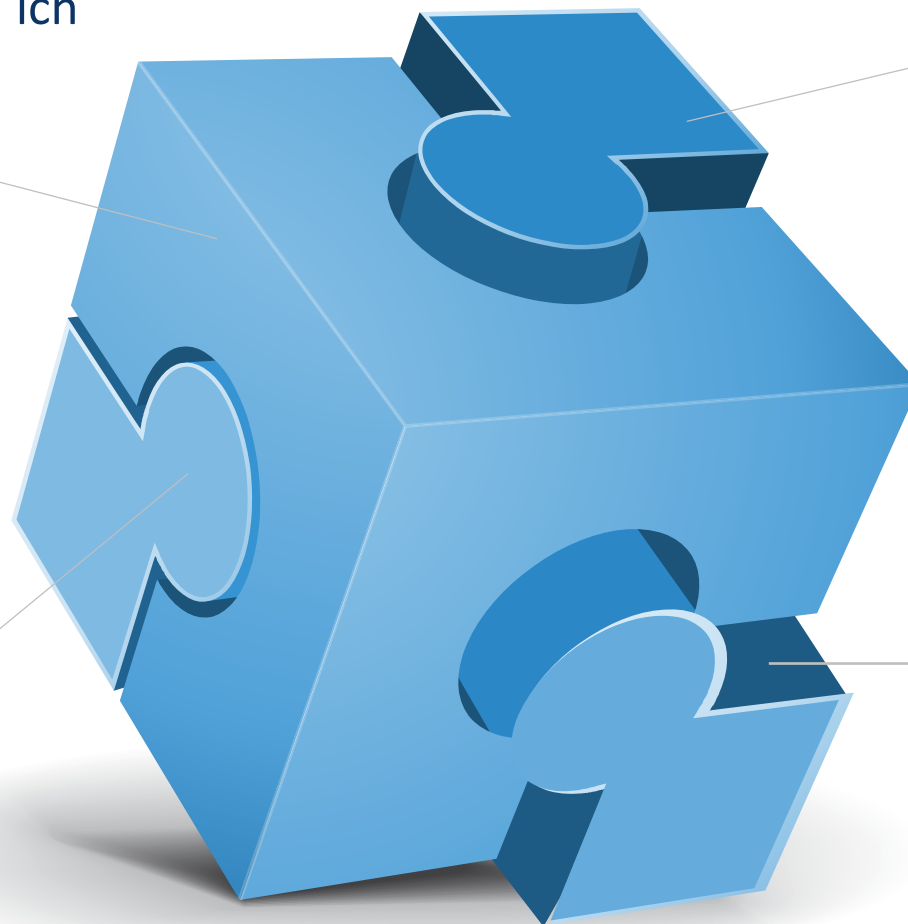
OCHRONA DANYCH

Organizacje dysponujące danymi i ich mandat społeczny

- + Prywatne korporacje
- + EU, Open Data, chmury
- + Powszechna dostępność a prywatność

KOMPETENCJE CYFROWE

- + Dostęp do oprogramowania
- + Sieci użytkowników
- + Otwarte ekosystemy dostępne



EKONOMIA 4.0

Harmonijny rozwój
Schematy mobilności
Decentralizacja
Sztuczna inteligencja

RÓŻNORODNOŚĆ POZNAWCZA

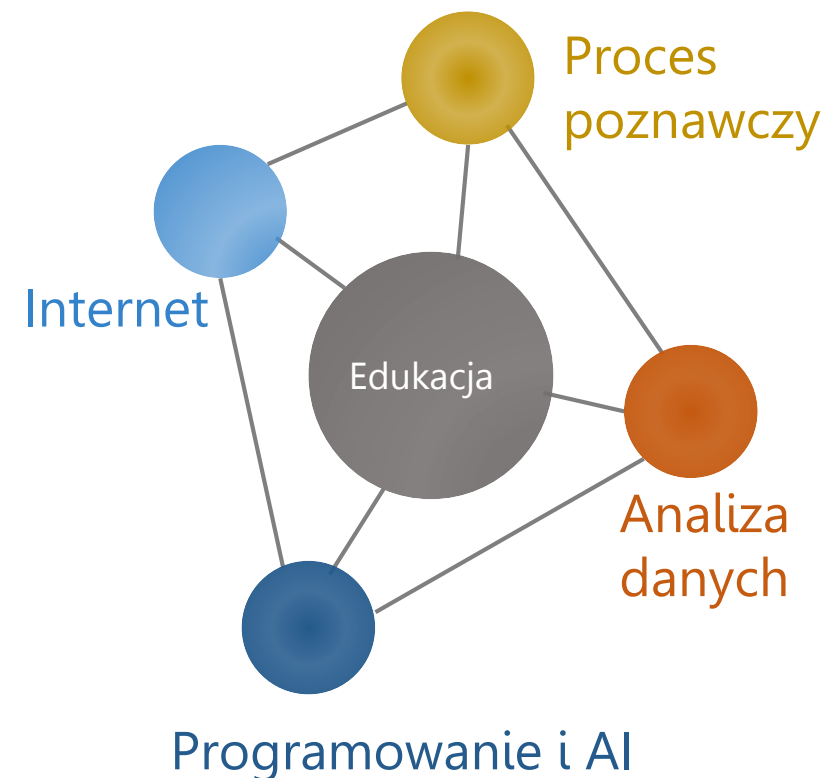
Wielość źródeł
Szybkość dostępu
Społeczności sieciowe
Media

DYDAKTYKA A RÓŻNORODNOŚĆ POZNAWCZA

- * Dostęp do internetu w klasie
- * Nauczyciel i uczeń współuczestnikami procesu poznawczego
- * Zdolność koncentracji a różnorodność

ROLA ANALIZY DANYCH, PROGRAMOWANIA I AI

- Analiza danych i wiedza: od empirii do teorii
- Korzystanie z edytorów tekstu i arkuszy kalkulacyjnych a programowanie
- Różnorodność źródeł, powszechność darmowych platform programowania, świat „open source”



MYŚL GLOBALNIE, DZIAŁAJ LOKALNIE

4

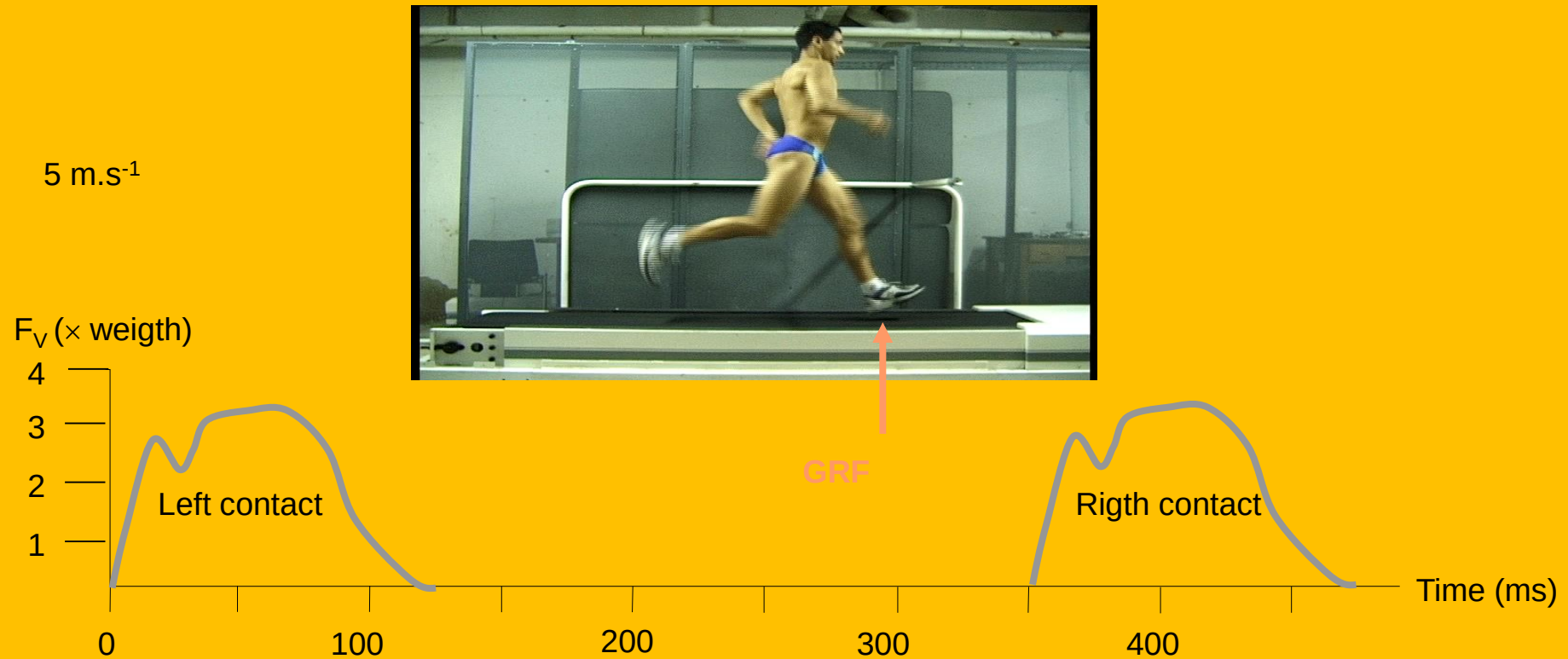
Sieć OSE – 26 tys szkół w jednym kampusie
OSE – wielki projekt cywilizacyjny

Społeczność uczniów i nauczycieli
Różnorodność poznawcza
Oprogramowanie, dane i AI



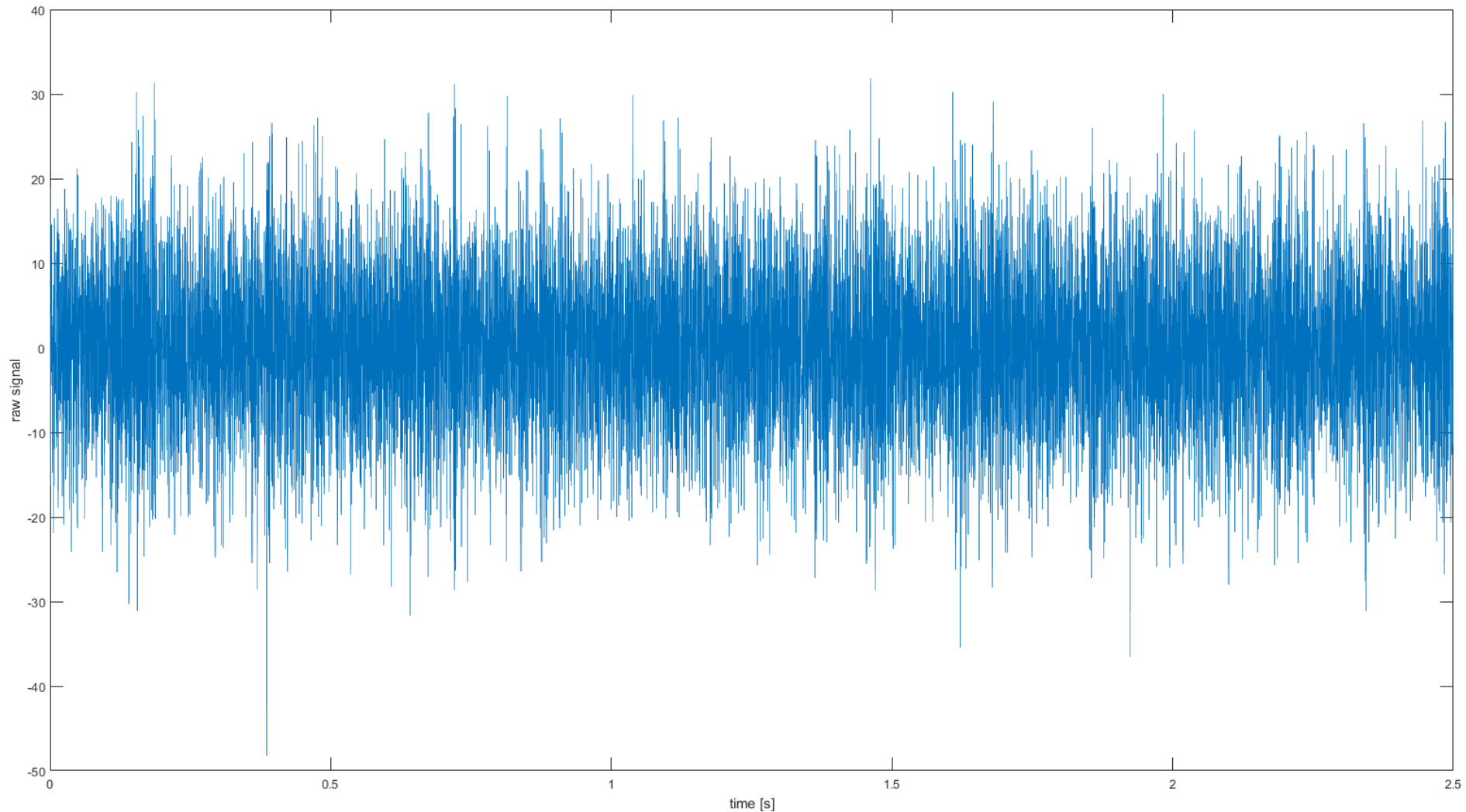
- **Natrętna cywilizacja cyfrowa**
- * **Trzy modele: EU {GDPR}, USA {korpo}, Chiny {totalitaryzm 2.0}**
- **„Dziecięca choroba” mierzenia {Steffen Mau, The Metric Society}**
- * **Big Data, sztuczna inteligencja, diagnostyka**

Przykład



Od biegu do sygnału. Big Data = 20 kHz / 10 min

7



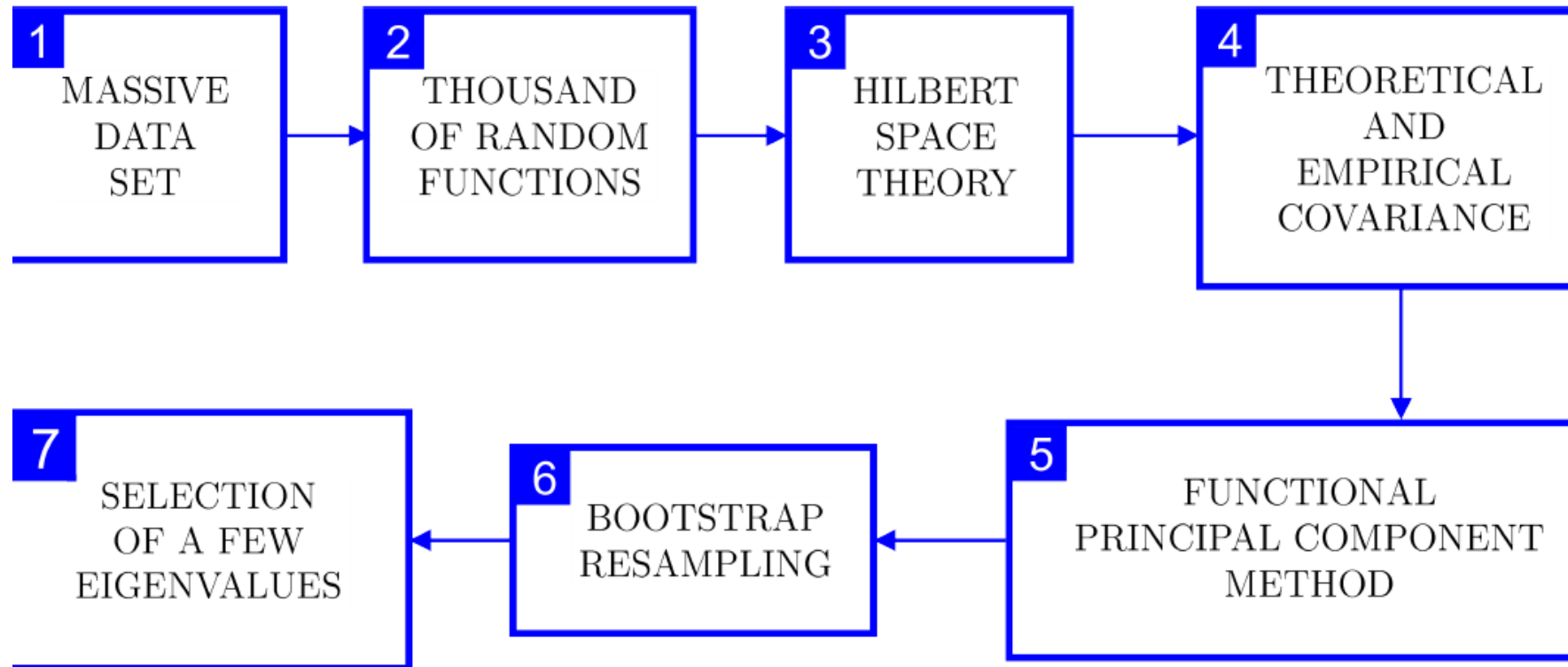
We will say that the functional time series $\{Z_i, i \geq 1\}$ fulfills the PFAR(1) model if for each $i = nT + \nu$ with $\nu = 1, \dots, T$ we have

$$Z_{nT+\nu} = \Phi_\nu(Z_{nT+(\nu-1)}) + \varepsilon_{nT+\nu}. \quad (1)$$

In the model (1) the operators $\Phi_\nu, \nu = 1, \dots, T$ are Hilbert-Schmidt integral operators in L^2 with corresponding kernels ϕ_ν fulfilling the assumption

$$\Phi_\nu(x)(t) = \int \phi_\nu(t, s)x(s)ds.$$

Moreover, the sequence ε_i is a sequence of i.i.d. mean zero elements in $\mathcal{H}$.



LFPCs

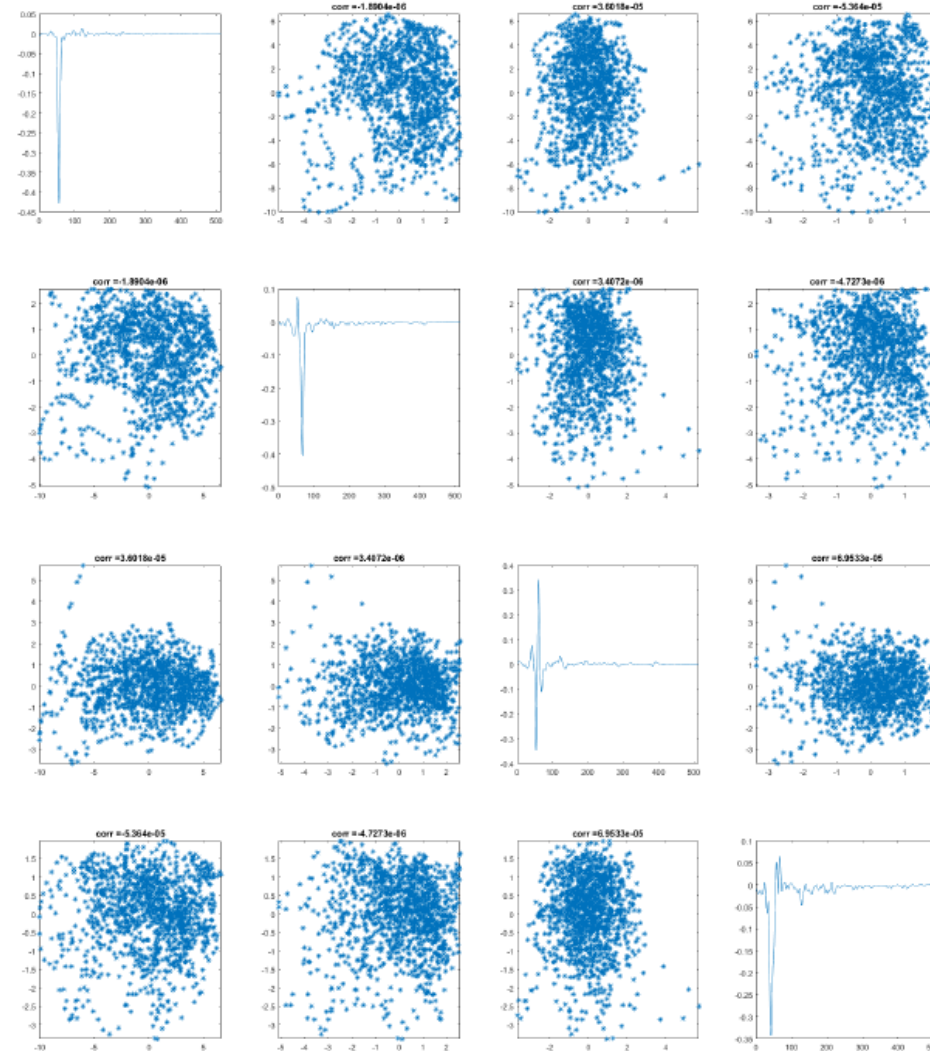
$$CPV(p) = \frac{\sum_{i=1}^p \hat{\lambda}_i}{\sum_{i=1}^N \hat{\lambda}_i}$$

We choose p for which $CPV(p)$ exceeds a desired level, 85% is the recommended value.

Tab.: Percentage of variability explained by each empirical functional component.

FPC1	FPC2	FPC3	FPC4	FPC5	FPC6	FPC7	FPC8
52.62 %	9.91%	5.48 %	4.07 %	3.36 %	2.38 %	2.10 %	1.96 %
Total explained:		81.88 %					

Weryfikacja wyboru funkcjonalnych wartości własnych



- **Zamiast milionów danych – kilka funkcjonalnych składowych głównych i odpowiadających im wartości**
- * **Diagnostyka oparta na testowaniu {bootstrap}**
- **Big Data, sztuczna inteligencja, diagnostyka – w rękach mandatariuszy społecznych**
- **Suwerenność cyfrowa i wiodąca rola statystyków {Data Scientist?}**

Projekty bieżące NASK – Państwowego Instytutu Badawczego z perspektywy Big Data, AI i inteligencji obliczeniowej

- **Cyberbezpieczeństwo ZUS**
- * **AI we Wspólnej Infrastrukturze Informacyjnej Państwa**
- **Diagnostyka medyczna**
- **POSZUKUJEMY STATYSTYKÓW !!!**

Dziękuję za uwagę