Dallo studio dell'insieme dei dati all'analisi dello spazio degli osservatori: come la geometria degli operatori equivarianti non espansivi può aiutarci nell'interpretazione delle informazioni

Patrizio Frosini

Department of Mathematics, ALMA-AI, AM^2 and ARCES, University of Bologna
patrizio.frosini@unibo.it

"Matematica per l'Intelligenza Artificiale e il Machine Learning"
Congresso UMI 2023, Pisa, 7 settembre 2023

Outline

The key role of observers in data analysis

Topological and metric basics for the theory of GENEOs

Building linear and nonlinear GENEOs

How can we use GENEOs in applications?

The key role of observers in data analysis

Topological and metric basics for the theory of GENEOs

Building linear and nonlinear GENEOs

How can we use GENEOs in applications?

Some key principles in our model

LET US ILLUSTRATE

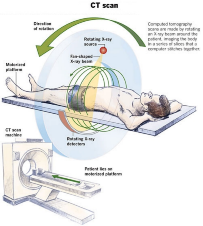
THE MAIN PRINCIPLES

IN OUR MODEL

Data can be often regarded as functions

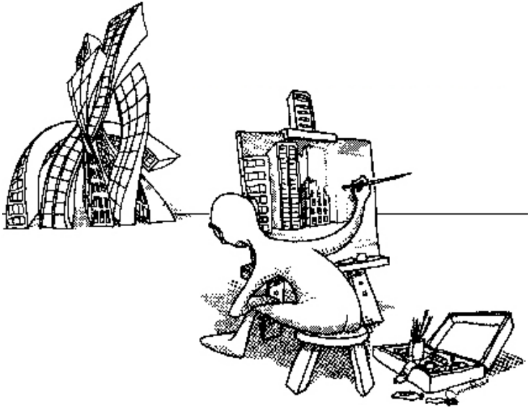
Some examples of data that can be seen as functions:

- An electrocardiogram (a function from $\mathbb{R}$ to $\mathbb{R}$);
- A gray-level image (a function from $\mathbb{R}^2$ to $\mathbb{R}$);
- A computerized tomography scan (a function from a helix to $\mathbb{R}$).



Data are processed by observers

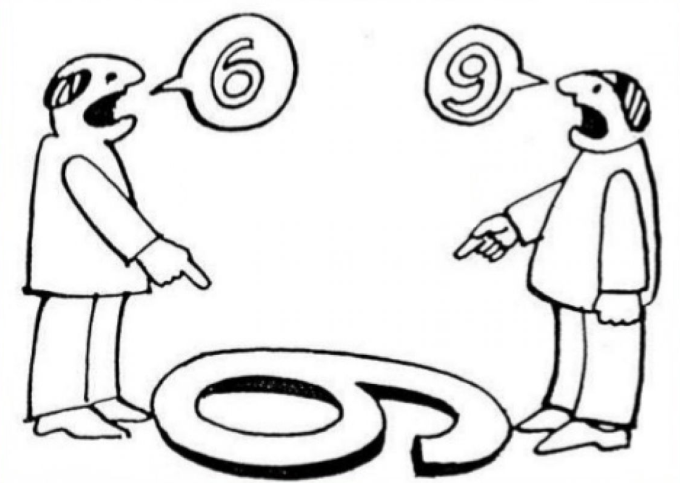
Data have no meaning if no observer elaborates them.



An observer is an agent that transforms data.

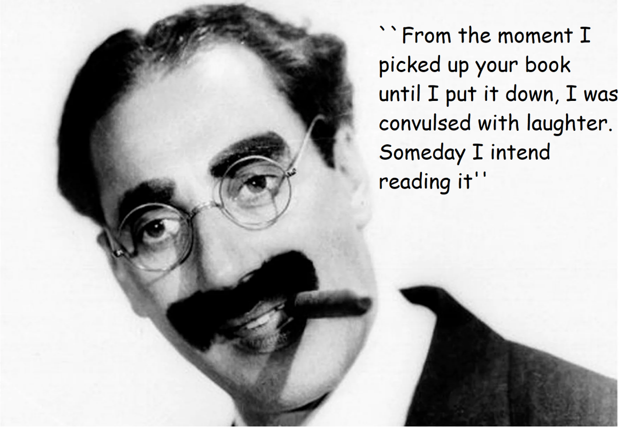
Observers are variables in data analysis

Data interpretation strongly depends on the chosen observer:



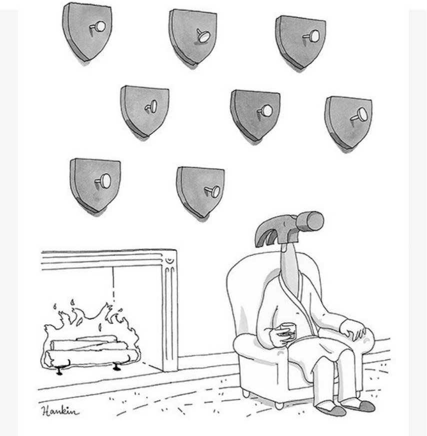
Our interest in data is greatly overrated

We are hardly ever interested directly in data but in the reaction of the observer to the presence of data.



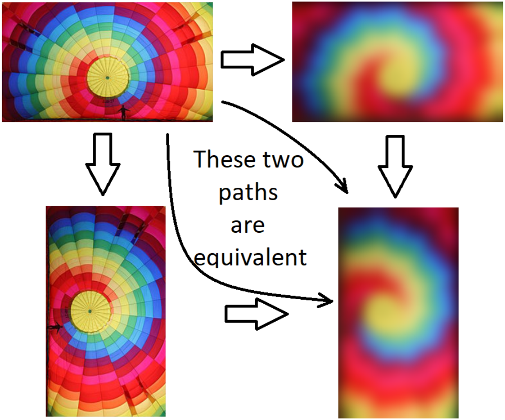
No data structure

Generally speaking, there is no structure in data. The structure of data is a projection of the structure of the observer.



Representing observers as equivariant operators

Observers are structures able to change data into other data, and usually do that by respecting some data symmetries, i.e., by commuting with some transformations (**equivariance**).



Representing observers as equivariant operators

As a first approximation, observers can be represented as **Group Equivariant Operators** (**GEOs**).
In this talk we will illustrate some results on the theory of **Group Equivariant Non-Expansive Operators** (**GENEOs**).

Why "non-expansive"?
Because

1. observers are often assumed to simplify the metric structure of data in order to produce meaningful interpretations;
2. non-expansiveness guarantees good topological properties.

Some preliminary results

LET US NOW ILLUSTRATE

SOME PRELIMINARY

TOPOLOGICAL RESULTS

How could we represent observers?

nature machine intelligence

ARTICLES

https://doi.org/10.1038/s42256-019-0087-3

Towards a topological–geometrical theory of group equivariant non-expansive operators for data analysis and machine learning

Mattia G. Bergomi[1], Patrizio Frosini[2,3*], Daniela Giorgi[4] and Nicola Quercioli[2,3]

We provide a general mathematical framework for group and set equivariance in machine learning. We define group equivariant non-expansive operators (GENEOs) as maps between function spaces associated with groups of transformations. We study the topological and metric properties of the space of GENEOs to evaluate their approximating power and set the basis for general strategies to initialize and compose operators. We define suitable pseudo-metrics for the function spaces, the equivariance groups and the set of non-expansive operators. We prove that, under suitable assumptions, the space of GENEOs is compact and convex. These results provide fundamental guarantees in a machine learning perspective. By considering isometry-equivariant non-expansive operators, we describe a simple strategy to select and sample operators. Thereafter, we show how selected and sampled operators can be used both to perform classical metric learning and to inject knowledge in artificial neural networks.

https://rdcu.be/bP6HV

All begins with the space of admissible functions

Let X be a nonempty set. Let Φ be a topological subspace of the set $\mathbb{R}_b^X$ of all bounded functions φ from X to $\mathbb{R}$, endowed with the topology induced by the metric

$$D_\Phi(\varphi_1, \varphi_2) := \|\varphi_1 - \varphi_2\|_\infty .$$

We can see X as the space where we can make our measurements, and Φ as the space of all possible measurements. We will say that Φ is the *set of admissible functions*. In other words, Φ **is the set of all functions from X to $\mathbb{R}$ that can be produced by our measuring instruments (or by other observers)**. For example, a gray-level image can be represented as a function from the real plane to the interval $[0,1]$ (in this case $X = \mathbb{R}^2$).

Perception pairs

Let us consider a group G of bijections $g : X \to X$ such that $\varphi \in \Phi \implies \varphi \circ g \in \Phi$ for every $\varphi \in \Phi$. We say that (Φ, G) is a perception pair.

The choice of a perception pair states which data can be considered as legitimate measurements (the functions in Φ) and which group represents the symmetries between data (the group G).

To proceed, we need to introduce suitable topologies on X and G. Before doing that, we recall that the initial topology τ_{in} on X with respect to Φ is the coarsest topology on X such that every function φ in Φ is continuous.

A pseudo-metric on X

Let us define on X the pseudo-metric

$$D_X(x_1, x_2) = \sup_{\varphi \in \Phi} |\varphi(x_1) - \varphi(x_2)|.$$

D_X induces a topology τ_{D_X} on X.
The use of D_X implies that we can distinguish two points only if a measurement exists, taking those points to different values.

Proposition

The topology τ_{D_X} is finer than the initial topology τ_{in} on X with respect to Φ. If Φ is totally bounded, then τ_{D_X} coincides with τ_{in}.

A pseudo-metric on X

The following properties are of use in our model.

Proposition

Every function in Φ is non-expansive, and hence continuous.

Proposition

If Φ is compact and X is complete, then X is compact.

In the following, we will usually assume that Φ is compact and X is complete (and hence compact).

Some magic happens: each bijection is an isometry

- $\mathrm{Bij}_\Phi(X) = \{$**bijections** $g : X \to X$ s.t. $\Phi \circ g, \Phi \circ g^{-1} \subseteq \Phi\}$;
- $\mathrm{Homeo}_\Phi(X) = \{$**homeomorphisms** $g : X \to X$ s.t. $\Phi \circ g, \Phi \circ g^{-1} \subseteq \Phi\}$;
- $\mathrm{Iso}_\Phi(X) = \{$**isometries** $g : X \to X$ s.t. $\Phi \circ g, \Phi \circ g^{-1} \subseteq \Phi\}$.

Proposition

$\mathrm{Bij}_\Phi(X) = \mathrm{Homeo}_\Phi(X) = \mathrm{Iso}_\Phi(X)$.

A pseudo-metric on G

Let us now focus our attention on a subgroup G of $\mathrm{Homeo}_\Phi(X)$.
We can define a pseudo-metric D_G on G by setting

$$D_G(g_1, g_2) := \sup_{\varphi \in \Phi} D_\Phi(\varphi \circ g_1, \varphi \circ g_2).$$

Theorem

G is a topological group with respect to D_G and the action of G on Φ by right composition is continuous.

Theorem

If Φ is compact and G is complete, then G is compact.

The concept of GENEO

WE ARE NOW READY

TO INTRODUCE

THE CONCEPT OF GENEO

GEOs and GENEOs

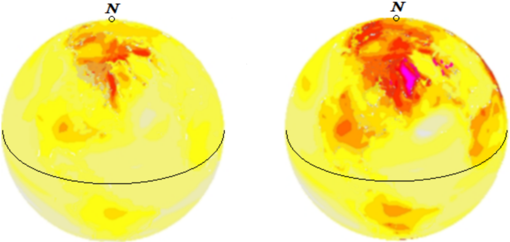
Let us assume that two perception pairs (Φ, G), (Ψ, H) are given, and fix a group homomorphism $T : G \to H$.

Each function $F : \Phi \to \Psi$ such that $F(\varphi \circ g) = F(\varphi) \circ T(g)$ for every $\varphi \in \Phi, g \in G$ is called a *Group Equivariant Operator (GEO)* associated with the homomorphism T.

If F is also non-expansive (i.e., $D_\Psi(F(\varphi_1), F(\varphi_2)) \leq D_\Phi(\varphi_1, \varphi_2)$ for every $\varphi_1, \varphi_2 \in \Phi$), then F is called a *Group Equivariant Non-Expansive Operator (GENEO)* associated with the homomorphism T.

An example of GENEO

Let us assume to be interested in the comparison of the distributions of temperatures on a sphere, taken at two different times:



Let us also assume that only two opposite points N, S can be localized on the sphere.

An example of GENEO

Let us introduce two perception pairs $(\Phi, G), (\Psi, H)$ by setting

- $X = S^2$
- $\Phi =$ set of 1-Lipschitz functions from S^2 to a fixed interval $[a, b]$
- $G =$ group of rotations of S^2 around the axis $N - S$

and

- $Y =$ the equator S^1 of S^2
- $\Psi =$ set of 1-Lipschitz functions from S^1 to $[a, b]$
- $H =$ group of rotations of S^1

An example of GENEO

This is a simple example of GENEO from (Φ, G) to (Ψ, H):

- $T(g)$ is the rotation $h \in H$ of the equator S^1 that is induced by the rotation g of S^2, for every $g \in G$.
- $F(\varphi)$ is the function ψ that takes each point y belonging to the equator S^1 to the average of the temperatures along the meridian containing y, for every $\varphi \in \Phi$;

We can easily check that F verifies the properties defining the concept of group equivariant non-expansive operator with respect to the isomorphism $T : G \to H$.

In plain words, our GENEO simplifies the data by transforming "temperature distributions on the earth" into "temperature distributions on the equator".

Two key results (and two good news for applications)

Let us assume that a homomorphism $T : G \to H$ has been fixed.
Let us define a metric D_{GENEO} on $\mathrm{GENEO}((\Phi, G), (\Psi, H))$ by setting

$$D_{\mathrm{GENEO}}(F_1, F_2) := \sup_{\varphi \in \Phi} D_{\Psi}(F_1(\varphi), F_2(\varphi)).$$

Theorem

If Φ and Ψ are **compact***, then $\mathrm{GENEO}((\Phi, G), (\Psi, H))$ is* **compact** *with respect to D_{GENEO}.*

Theorem

If Ψ is **convex***, then $\mathrm{GENEO}((\Phi, G), (\Psi, H))$ is* **convex***.*

Two key observations (1)

- While the space of data is often non-convex (and hence averaging data does not make sense), the assumption of convexity of Ψ implies the convexity of the space of observers and allows us to consider the "average of observers".

Two key observations (2)

- Our main goal is to develop a good geometric and compositional theory to approximate an ideal observer. In our model, "to approximate an observer" means to look for a GENEO F that minimizes a suitable "cost function" $c(F)$. The cost function quantifies the error that is committed by taking the GENEO F instead of the ideal observer. Since the space of GENEOs is compact and convex (under the assumption that the data spaces are compact and convex), if the cost function $c(F)$ is strictly convex we have that there is one and only one GENEO that best approximates the ideal observer.

Methods to build GENEOs

NOW THAT WE KNOW

THE ROLE OF GENEOS

WE NEED TO EXPLAIN

HOW WE CAN BUILD THEM

How can we build linear and nonlinear GENEOs?

frontiers
in Artificial Intelligence

ORIGINAL RESEARCH
published: 15 February 2022
doi: 10.3389/frai.2022.786091

Check for updates

On the Construction of Group Equivariant Non-Expansive Operators *via* Permutants and Symmetric Functions

Francesco Conti[1,2], *Patrizio Frosini*[3,4,5,6]* and *Nicola Quercioli*[3,7]

[1] Department of Mathematics, University of Pisa, Pisa, Italy, [2] Institute of Information Science and Technologies "A. Faedo", National Research Council of Italy (CNR), Pisa, Italy, [3] Department of Mathematics, University of Bologna, Bologna, Italy, [4] Alma Mater Research Center on Applied Mathematics, University of Bologna, Bologna, Italy, [5] Alma Mater Research Institute for Human-Centered Artificial Intelligence, University of Bologna, Bologna, Italy, [6] Research Centre on Electronic Systems for the Information and Communication Technology, University of Bologna, Bologna, Italy, [7] ENEA Centro Ricerche Bologna, Bologna, Italy

https://www.frontiersin.org/articles/10.3389/frai.2022.786091/full

Elementary methods to build GENEOs

Proposition (Composition)

If $F_1 \in \mathrm{GENEO}((\Phi, G), (\Psi, H))$ *w.r.t.* $T_1 : G \to H$ *and* $F_2 \in \mathrm{GENEO}((\Psi, H), (\chi, K))$ *w.r.t.* $T_2 : H \to K$ *then* $F_2 \circ F_1 \in \mathrm{GENEO}((\Phi, G), (\chi, K))$ *w.r.t.* $T_2 \circ T_1 : G \to K$.

Proposition (Image by a 1-Lipschitz function)

If $F_1, \ldots, F_n \in \mathrm{GENEO}((\Phi, G), (\Psi, H))$ *w.r.t.* $T : G \to H$, L *is a 1-Lipschitz map from* $\mathbb{R}^n$ *to* $\mathbb{R}$, *and* $L^*(F_1, \ldots, F_n)(\Phi) \subseteq \Psi$ *(where* L^* *is the map induced by* L*), then* $L^*(F_1, \ldots, F_n) \in \mathrm{GENEO}((\Phi, G), (\Psi, H))$ *w.r.t.* T.

The next three statements follow from the last proposition.

Elementary methods to build GENEOs

Proposition (LATTICE OF GENEOS)

If $F_1, \ldots, F_n \in \mathrm{GENEO}((\Phi, G), (\Psi, H))$ *w.r.t.* $T : G \to H$ *and* $\max(F_1, \ldots, F_n)(\Phi), \min(F_1, \ldots, F_n)(\Phi) \subseteq \Psi$, *then* $\max(F_1, \ldots, F_n), \min(F_1, \ldots, F_n) \in \mathrm{GENEO}((\Phi, G), (\Psi, H))$ *w.r.t.* T.

Proposition (Translation)

If $F \in \mathrm{GENEO}((\Phi, G), (\Psi, H))$ *w.r.t.* $T : G \to H$, *and* $F_b(\Phi) \subseteq \Psi$ *for* $F_b(\varphi) := F(\varphi) - b$, *then* $F_b \in \mathrm{GENEO}((\Phi, G), (\Psi, H))$ *w.r.t.* T.

Proposition (Convex combination)

If $F_1, \ldots, F_n \in \mathrm{GENEO}((\Phi, G), (\Psi, H))$ *w.r.t.* $T : G \to H$, $(a_1, \ldots, a_n) \in \mathbb{R}^n$ *con* $\sum_{i=1}^n |a_i| \leq 1$ *and* $F_\Sigma(\Phi) \subseteq \Psi$ *for* $F_\Sigma(\varphi) := \sum_{i=1}^n a_i F_i(\varphi)$, *then* $F_\Sigma \in \mathrm{GENEO}((\Phi, G), (\Psi, H))$ *w.r.t.* T.

Permutant measures

Let us consider the set $\Phi = \mathbb{R}^X \cong \mathbb{R}^n$ of all functions from a finite set $X = \{x_1, \ldots, x_n\}$ to $\mathbb{R}$, and a subgroup G of the group $\mathrm{Bij}(X)$ of all permutations of X.

Definition

A finite (signed) measure μ on $\mathrm{Bij}(X)$ is called a *permutant measure* with respect to G if every subset H of $\mathrm{Bij}(X)$ is measurable and μ is invariant under the conjugation action of G (i.e., $\mu(H) = \mu(gHg^{-1})$ for every $g \in G$).

Proposition

If μ is a permutant measure with respect to G, then the map $F_\mu : \mathbb{R}^X \to \mathbb{R}^X$ defined by setting $F_\mu(\varphi) := \sum_{h \in \mathrm{Bij}(X)} \varphi h^{-1} \, \mu(h)$ is a linear GEO. If $\sum_{h \in \mathrm{Bij}(X)} |\mu(h)| \leq 1$, then $F_\mu(\varphi)$ is a GENEO.

How can we represent linear GENEOs?

Annals of Mathematics and Artificial Intelligence (2023) 91:465–487
https://doi.org/10.1007/s10472-022-09830-1

On the finite representation of linear group equivariant operators via permutant measures

Giovanni Bocchi[1] · Stefano Botteghi[2] · Martina Brasini[2] · Patrizio Frosini[2] · Nicola Quercioli[3]

https://rdcu.be/c5Obw

Representation Theorem for linear GENEOs

The following theorem strengthens our previous result about building linear GENEOs via permutant measures.

Theorem (Representation Theorem for linear GENEOs)

Let us assume that $G \subseteq \mathrm{Bij}(X)$ *transitively acts on the finite set* X *and that* F *is a map from* $\mathbb{R}^X$ *to* $\mathbb{R}^X$. *The map* F *is a linear GENEO from* $\mathbb{R}^X$ *to* $\mathbb{R}^X$ *with respect to the identical homomorphism* $\mathrm{id}_G : g \mapsto g$ **if and only if** *a permutant measure* μ *with respect to* G *exists, such that* $F(\varphi) = \sum_{h \in \mathrm{Bij}(X)} \varphi h^{-1} \mu(h)$ *for every* $\varphi \in \mathbb{R}^X$, *and* $\sum_{h \in \mathrm{Bij}(X)} |\mu(h)| \leq 1$.

The key role of observers in data analysis

Topological and metric basics for the theory of GENEOs

Building linear and nonlinear GENEOs

How can we use GENEOs in applications?

The use of GENEOs in applications

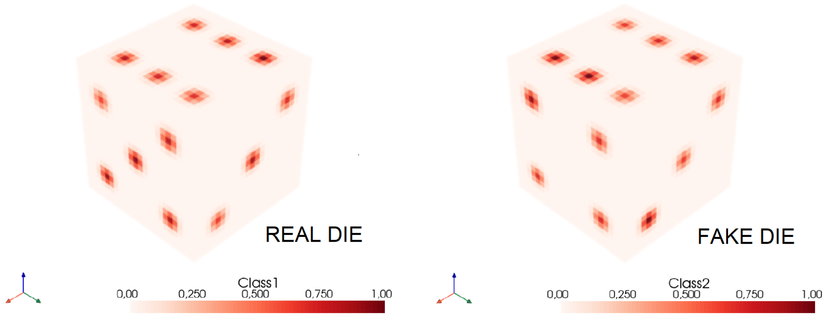
WE WILL NOW SHOW

HOW GENEOS CAN BE USED

IN APPLICATIONS

What happens when we apply GENEOs to our data?

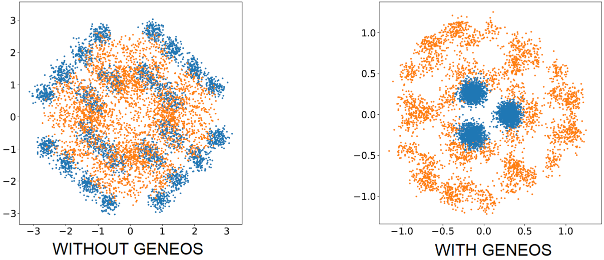
An example of use: comparison between real dice and fake dice.



(Experiment and computations by Giovanni Bocchi)

What happens to data when we apply GENEOs?

We produced 10000 dice (a training set of size 7000 and a test set of size 3000), then we applied PCA to the test set and to the test set transformed by a suitable GENEO, optimized on the training set:



For each die the first two principal components are plotted. Blue points are associated with **real dice**, while orange ones with **fake dice**. The GENEO we use was built by a convex combination of 3 GENEOs defined by permutant measures.

Using GENEOs to remove impulsive noise in TDA

entropy

an Open Access Journal by MDPI

A Probabilistic Result on Impulsive Noise Reduction in Topological Data Analysis through Group Equivariant Non-Expansive Operators

Patrizio Frosini; Ivan Gridelli; Andrea Pascucci

Entropy **2023**, Volume 25, Issue 8, 1150

https://www.mdpi.com/1099-4300/25/8/1150

GENEOs and Machine Learning

For more details about the use of GENEOs in Machine Learning:

- A. Micheletti, *A new paradigm for artificial intelligence based on group equivariant non-expansive operators*, In: *EMS Magazine*, Online First, 24 April 2023.
- https://ems.press/content/serial-article-files/27673

Some current lines of research

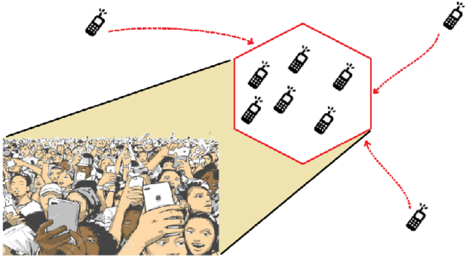
We are presently studying these problems:

- How can we extend the theory of GENEOs to graphs? (F. Ahmad, M. Ferri, P. Frosini, *Generalized Permutants and Graph GENEOs*, `https://arxiv.org/pdf/2206.14798.pdf`.)
- How can we extend the theory of GENEOs to probability spaces of signals? P. Cascarano, P. Frosini, N. Quercioli, A. Saki, *On the geometric and Riemannian structure of the spaces of group equivariant non-expansive operators*, `https://arxiv.org/pdf/2103.02543.pdf`

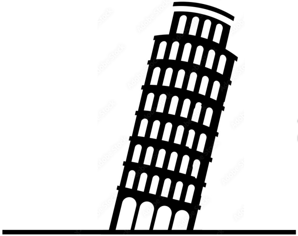
A current research project

CNIT / WiLab - Huawei Joint Innovation Center (JIC)

Project on GENEOs for 6G



WILAB HUAWEI



THANKS
FOR YOUR
ATTENTION