Evaluating research output @ INFN

Luciano Canton
INFN – sezione di Padova

INFN: a very structured, large and complex research organization.
Research evaluation occurs at many levels.

BUT, there are two specific moments where INFN is evaluated globally...

The international evaluation committee (CVI)
5 very high-profile scientists (P. Drell chair) and 2 experts from industry and economics

- Every year, it assesses INFN performance and provides an annual report. This report reviews the overall quality of INFN activities and is a reference for future directions, it is for MIUR and for the executive board of INFN.

- Procedure and committee estabilshed since '90

National multi-year evaluation of research structures (UNI and RO)

- **Only one previous experience (VTR:2001-2003)**
  INFN selected & provided relevant infos about **660** products

- **The current evaluation process** (**VQR:2004-2010**) is the most systematic and capillary evaluation of research structures ever attempted in Italy!

- VQR is the first evaluation process started and organized by the recently established National Agency of Research University Evaluation (ANVUR)....   …........ and   it is   **BIG !!!**

  200 000 **products across ALL DISCIPLINES**

  20 000 **products selected in PHYSICS**

  about   **6100** **products selected by INFN.**

The INFN Evaluation Working Group (gruppo di lavoro sulla valutazione-GLV)

(Chair: Giorgio Chiarelli PISA)

The Group was established in 2000 under the previous research evaluation procedure VTR. The GLV consists of five groups (one for each Scientific Committee) responsible for in-depth analysis of the scientific performance of each activity.

THE DATABASE:
http://www.infn.it/pubblicazioni

The GLV, thanks to the valuable support of the **INFN dataweb service,** collects and maintain in this database all the publications by affiliates, indexed by the 'Institute for Scientific Information' (ISI, of Thomson-Reuters). This is a fundamental tool to prepare all the bibliometric data for research evaluation.

THE DB is more than just publications.

THESIS completed in the years

Outreach Initiatives (third mission events)

Formation and continuous education of personnel

Talks and Proceedings
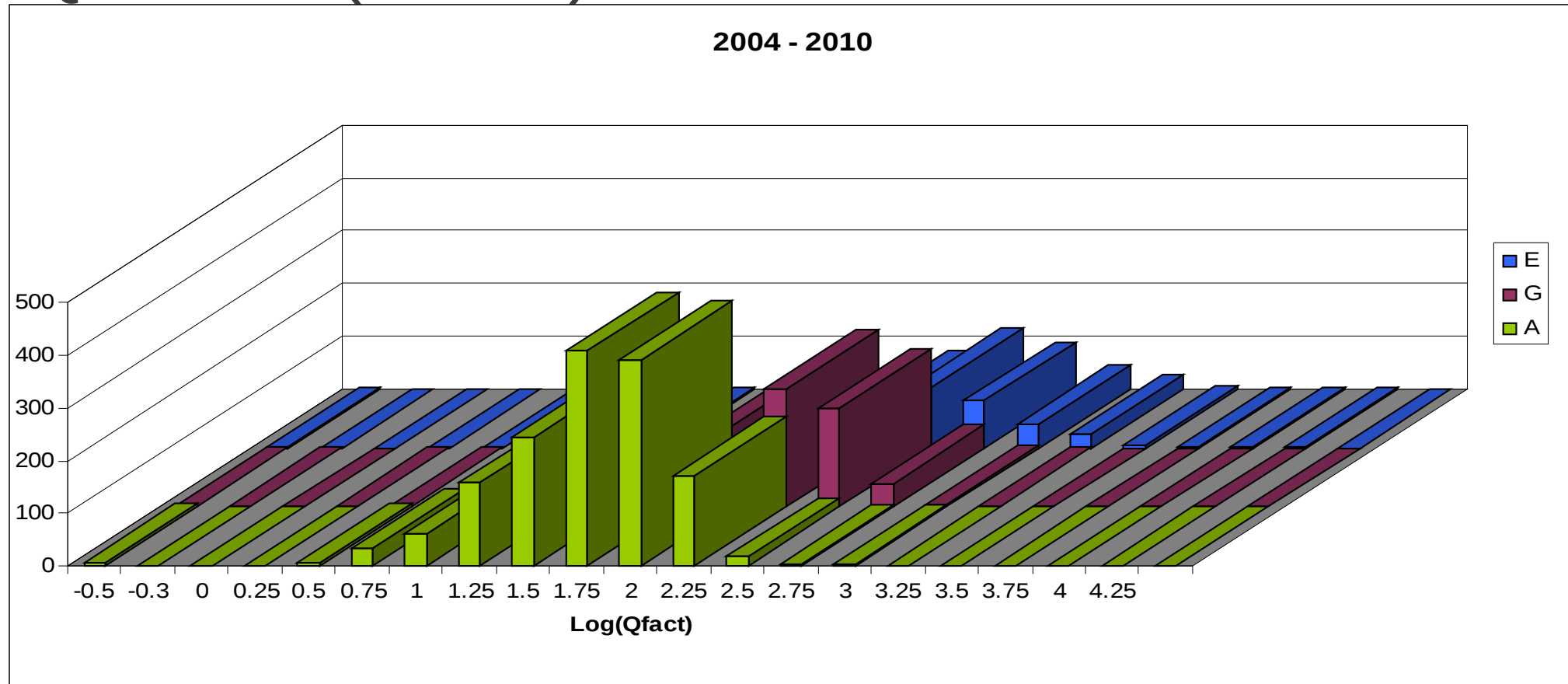
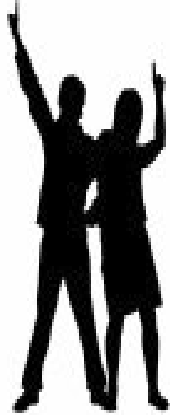Financial allocation of resourses

and more..

people

From VTR experience, every year we selected "eccellent, good & acceptable" publications
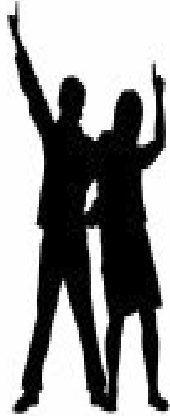
Qfact =IF(CIT+1)

QUESTION n.1!

HOW TO ALLOCATE FUNDS AMONGST GROUPS/ STRUCTURES **(WITHOUT MERITS)**?

ANSWER:

JUST COUNT HEADS!!!

QUESTION n.2!

HOW TO ALLOCATE FUNDS STILL WITHOUT MERITS, BUT **BETTER**?

ANSWER:

JUST COUNT ACTIVE HEADS!!

IN CSN4,

it's called BECCHI rule

QUESTION n.3   !

HOW TO ALLOCATE FUNDS AMONGST GROUPS/ STRUCTURES **(WITH MERIT OR SCIENTIFIC VALUE)**?

ANSWER:

JUST COUNT HEADS, EACH ONE WEIGHTED THROUGH EVALUATION OF THEIR RESEARCH PRODUCTS!!!

THIS IS THE VQR WAY …   it costs

The FINAL VQR CALL on nov, 7th, 2011

THE FINAL VERSION WAS RELEASED AFTER A THOROUGH DISCUSSION WITHIN THE ITALIAN RESEARCH COMMUNITY. THE GLV ACTIVELY PARTECIPATED SUGGESTING MODIFICATIONS TO IMPROVE THE METHOD AND CORRECTING SOME EVIDENT DISTORTIONS.

ANVUR demonstrated a balanced and open attitude to the voices of the research communities.
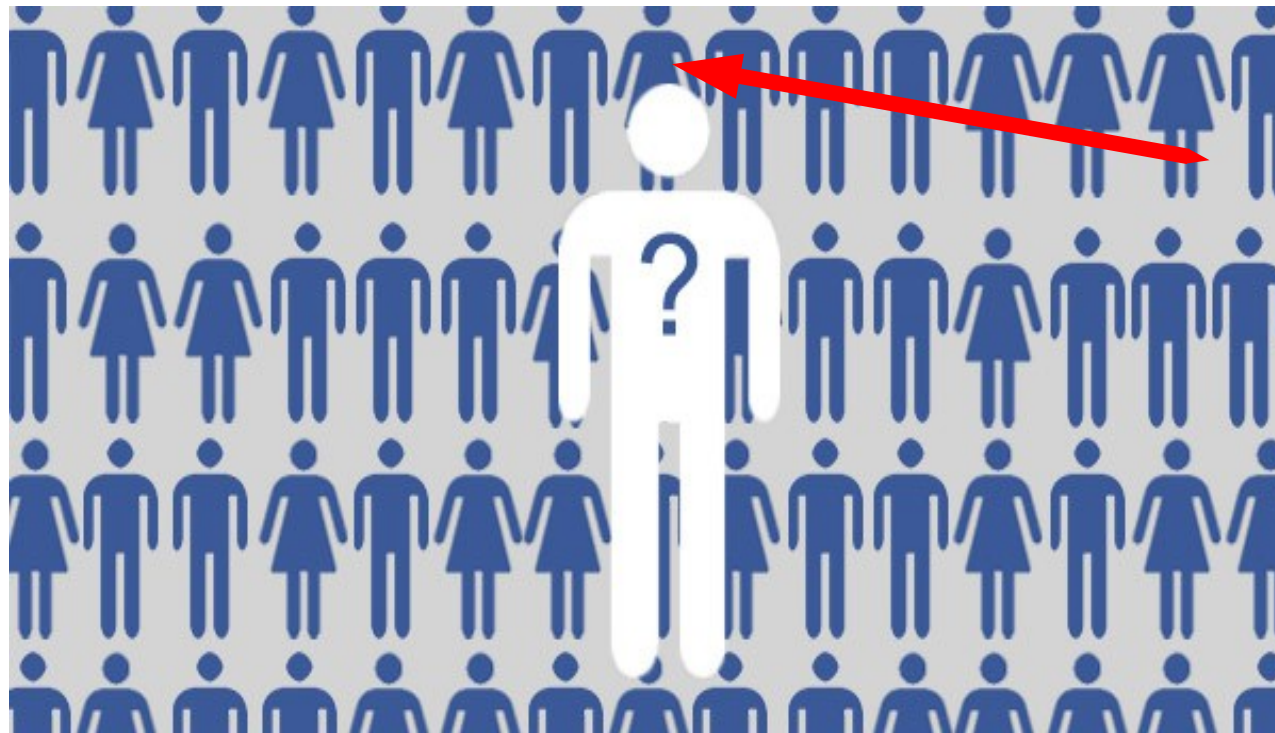
OUTLINE OF VQR 2004-2010

W-max = 3 TECHNOLOGIST FROM RO (Like INFN)
W-max = 6 RESEARCHER   FROM RO (Like INFN)
W-max = 3 RESEARCHER FROM UNIVERSITY
W-max = 6 (3 University + 3 INFN) RESEARCHER FROM UNIVERSITY
    with established association (**incarico**) with INFN

HOW TO GAIN
W-max = "N" ?

NEED "N "
EXCELLENT
PUBBLICATIONS
IN THE 2004-2010
BY THAT AUTHOR

A PUBLICATION CAN BE ...

ECCELLENT ! GRADE "A"   WEIGHT = 1

GOOD !   GRADE "B"   WEIGHT = 0.8

ACCEPTABLE! GRADE "C" WEIGHT = 0.5

LIMITED !   GRADE "D" WEIGHT = 0

MISSING !   WEIGHT = -0.5
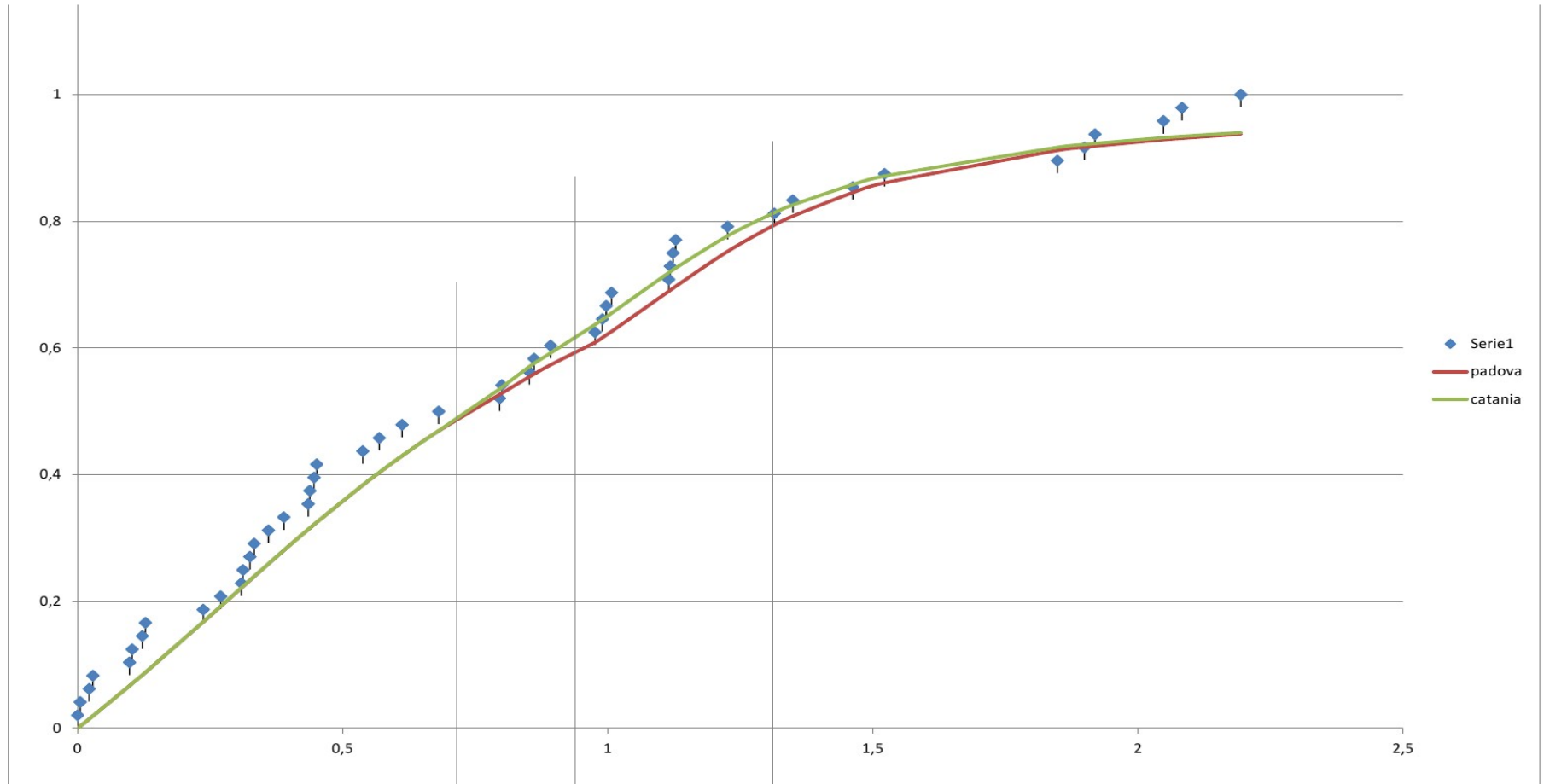
NOT ACCEPTABLE   WEIGHT = - 1

FORGERY, PLAGIARISM   WEIGHT = - 2

TO GRADE AN ISI PUBLICATION ONE NEEDS

- THE "IF" OF THE JOURNAL

- THE CITATION # UP TO END OF 2011, and YearPub

- THE SCIENCE CATEGORIES BY ISI

- 1 or 2 PACS

- TO KNOW THE OVERALL CUMULATIVE OF ALL WORLD PUBLICATIONS FOR GIVEN YEAR AND SCIENCE CATEGORIES. (FOR BOTH IF AND CIT)

- THE 50, 60, 80 QUANTILES of the two cumulatives, defining 4 intervals with prob 0.2, 0.2, 0.1, 0.5.

Cumulative functions

IF Instruments & Instrumentation 2004

THE EVALUATION MATRIX

IF

CIT

A	A	IR	IR
IR	B	B	IR
IR	IR	C	IR
IR	IR	IR	D

29, Feb 2012 published by Physics Expert Group eValuators (GEV)
18 high profile scientists led by Giorgio Parisi

THE INFN WEIGHTED HEADS

INFN SELECTION

800 INFN pers ---> 3930 products

726 University pers --> 2178 prod

UNIV SELECTION

726 University pers --> 2159 prod

SELECTION OF TOTAL 6108 prod

INTERACTION WITH 2159 UNI PRODUCTS AND 42 ITALIAN UNIVERSITY STRUCTURES AND THEIR (sometimes ANGRY) DIRECTORS

Products in our 2004-2010 database

- TOTAL PRODUCTS: 19180
  - Isi: 17709   **processed bibliometrically**
  - Infn: 898   (not ISI)
  - Jacow: 416 (these in accelerator sciences)
  - Infnman: 133   manufacts!
  - Inspec: 20   (these in computer sciences)
  - Medline: 4 (these in biomedicine)

ONE PRODUCT USABLE ONLY ONCE (FOR EACH STRUCTURE)

THIS DOES NOT CONSIDER A "SCALE LAW" !!!!!!
EX-HEP WORKS IN BIG EXPERIMENTS AND LARGE COLLABORATIONS (300/400 ITALIANS AND 3000 TOTAL)

ALL ITALIANS IN ONE BIG STRUCTURE, **impossible to saturate all with A, B, C, D, papers.** Somebody will get **poor publications**, and somebody **very poor** products (astracts?)

FEW ITALIANS IN A SMALL STRUCTURE, easy to saturate all weights "w" with **excellent** publications from big experiments

BIG STRESS BETWEEN GLV AND INDIVIDUALS.

- "WHY ARE YOU NOT GIVING TO **ME MY** BEST pub?"

- **BIG STRESS BETWEEN GLV AND DEPARTMENTS**

- **WHY ARE YOU NOT GIVING TO MY UNI PERSONNEL THE BEST pub ?**

- **BIG STRESS AMONGST DEPARTMENTS**

- **YOU (GLV) ARE FAVORING MY COMPETITOR DEPT!!!**

  ….

- AND COMPLAINTS...

- THE INFN SELECTION INTERFERES THE PERFORMANCES OF THE DEPARTMENTS

THE GLVsoft SELECTION OF PRODUCTS

- The complexity of the selection could be handled efficiently only by a computing code, **GLVsoft** (By **F.Arneodo**, P. Pedroni, & LC). Optimization by Linear Programming Techniques using the Simplex Algorithm.

- Not clear from the VQR call if the selection was bottom-up or top-down procedure. CONFUSION!

  The complexity of "collisions" of publications at the INFN size forced to use top bottom procedure.

- With GLVsoft we provided also an optimization service for the selection of the UNI departments

DISTRIBUTION OF SELECTED PRODUCTS IN SCIENCE CATEGORIES

Category	Count
physics, multidisciplinary	1203
physics, particles & fields	1009
nuclear science & technology	926
physics, nuclear	784
astronomy & astrophysics	674
instruments & instrumentation	494
physics, mathematical	137
physics, applied	97
optics	73
radiology, nuclear medicine & medical imaging	60
physics, condensed matter	59
physics, fluids & plasmas	43
mechanics	31
computer science, interdisciplinary applications	28

VQR: ADDITIONAL ASPECTS

1. Quality of products (main index)

2. Ability to attract external resourses

3. Quality of products of researchers recently advanced in their carrier

4. Level of internationalization of research.

5. Number of PhD students and post docs (training).

6. other minor corrections (ability to allocate internal funds for research, etc)

CONCLUDING REMARKS

- THERE IS A WELL ESTABLISHED "EVALUATION CULTURE" IN GLV AND IN THE WHOLE INFN.

- SOUND EVALUATION PROCESSES ARE WELCOME HERE

- NOT EVERYTHING WORKED WELL IN THE VQR PROCESS. THE DEADLINES WERE NOT WELL PLANNED. Criteria were published on 29 Feb. Product selection (6100 products!) had to be completed by 30 May. INFN succeeded to match the deadlines but with great difficulty. Later on, ANVUR extended the deadlines, invalidating the efforts of those who did match the deadlines.

- A better planning of information releases and deadlines is mandatory for the future, for all requested metadata.

- This system of weighting products (up to 1) is clearly unsuited for several outstanding research products by INFN. It is humiliating and reductive to rate the whole CNAO project as a manufact, at the same level of a single short communication of three pages in a journal.

- THE PAST EXPERIENCE WITH VQR TELLS US THAT WE HAVE TO BE READY EVERY YEAR, AND IMPROVE CONTINUOSLY THE EVALUATION TOOLS AND SKILLS

- WE ARE WAITING WITH EXPECTATIONS THE OUTCOME OF THE VQR EVALUATION PROCESS:. THE RESULTS OF THIS BIG EVALUATION PROCESS ARE EXPECTED IN MID 2013. MAY – BE?