CTF implementation of Open Science and Data Sharing

HRA Member's Meeting – Manhattan Beach, CA

March 23, 2017

Salvatore La Rosa, PhD

VP Research and Development
slarosa@ctf.org

CHILDREN'S
TUMOR
FOUNDATION
ENDING NF
THROUGH RESEARCH

Reasons for creating a data-portal

- There is little data because of few researchers

- Data is scattered and sometimes duplicated in different reports

- Hard scientific questions are rarely addressed

- Conflicted results often reported separately because of different model systems employed

- Confusion in the clinical community on which drug to progress to clinical trials

Reasons for CTF to have a data-portal

- Being a small/rare disease community we believe it's critical to have a single go-to repository to find useful data
  - Few researchers, scattered data, limited funding

- Have an active role in the research community and contribute with providing useful resources

- Use the data to launch specific RFAs
  - Generate questions
  - Validate Hypotheses

- Attract out-of-the-field researchers and business opportunities

Reasons to have an open-Data policy:

- Have researchers collaborate more

- Disseminate results faster

- Report negative data

- Avoid re-funding of what has already been done (...and not worked)

- Increase transparency and reproducibility

What should a data repository contain?

- We envision two phases for our NF Data portal :
  - Phase I: mainly preclinical data (no clinical or patient data)
  - Phase II: connect research data to clinical phenotypes or clinical trials data
    - Requires appropriate consent, regulatory compliance for privacy

- Big emphasis on metadata and annotation/classification

- Connect to other databases - no need to duplicate datasets at different locations, but be able to obtain data in the right format for easier integration

What are the incentives to make it work?

- Tie funding opportunities to data submission
    - i.e. launch 'special edition' grant awards as registered reports

- Provide smart tools to aid data analysis
    - Provide dedicated personnel who can look into the datasets and help if needed

- Provide opportunities for collaboration

- Data quality in line with scientific journal standards

- Create incentives for researchers to use the data (workshops, challenges, specific RFA)

- Require data users to return their enriched data back to the platform

CTF Consortia Funding Conditions

- **1.1** - *Periodic reports on the projects are accompanied by submission of all data in an electronic format*
- **1.2** - *During the Confidentiality Period, access to this data in this repository will be limited to CTF and any Funded Party only*
- **1.3** - *At the end of the Confidentiality Period, CTF shall give access to this data in this repository to anyone without restriction.*

- **2.4** - *During the Confidentiality Period, CTF shall have the right to use the Results in order to make decisions as to whether or not to fund other projects or grants.*

- **3.2** - *Confidentiality period in no event shall be less than six months and exceed one calendar year from the disclosure within the Periodic Report*

Consortia
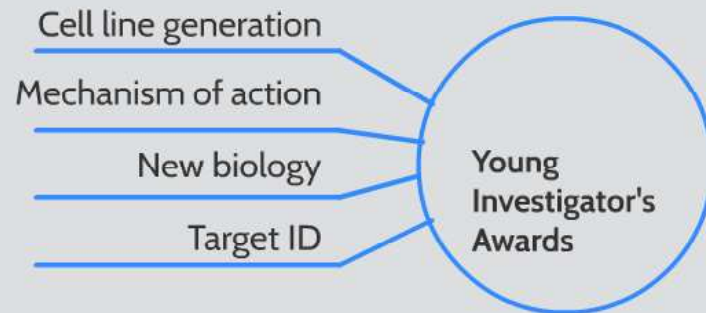

Synodos NF2

Synodos NF1

Synodos NF1

Synodos NF1 LLG
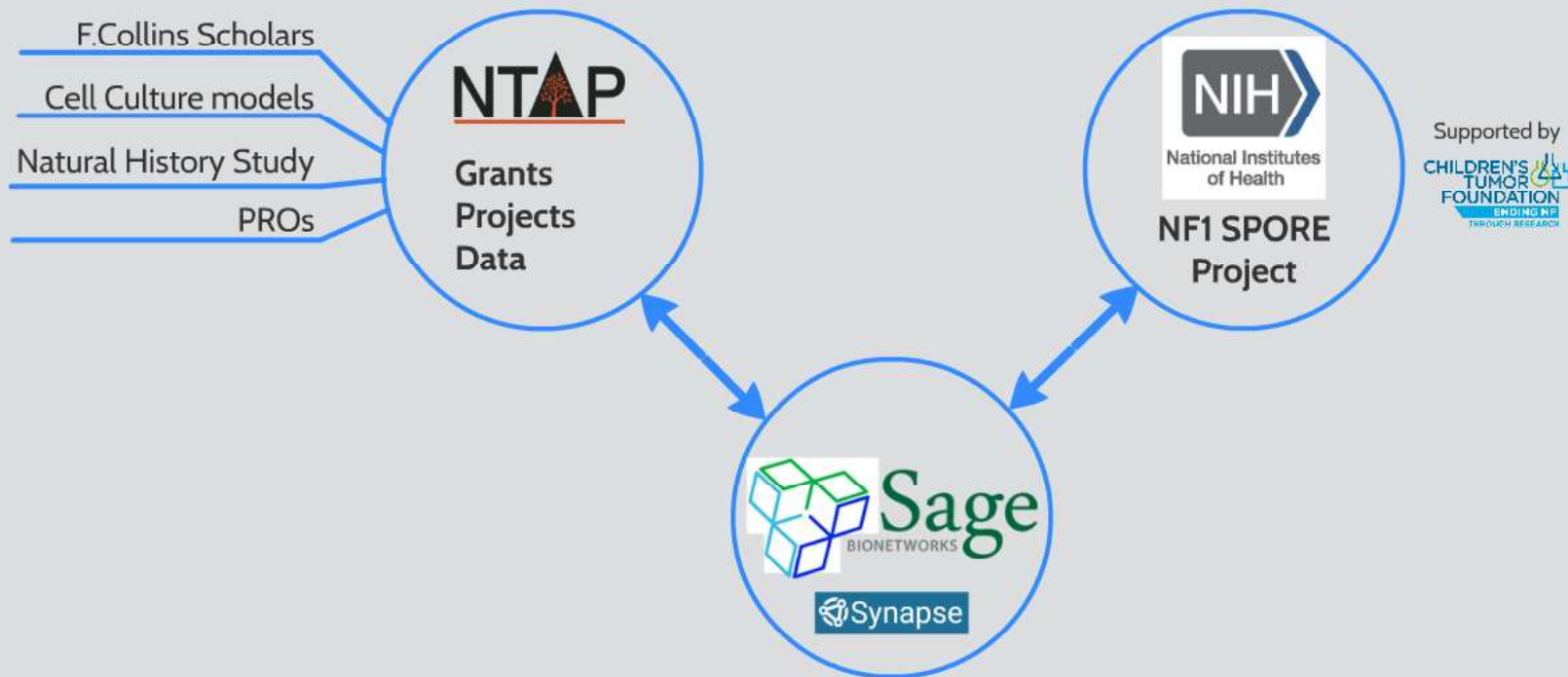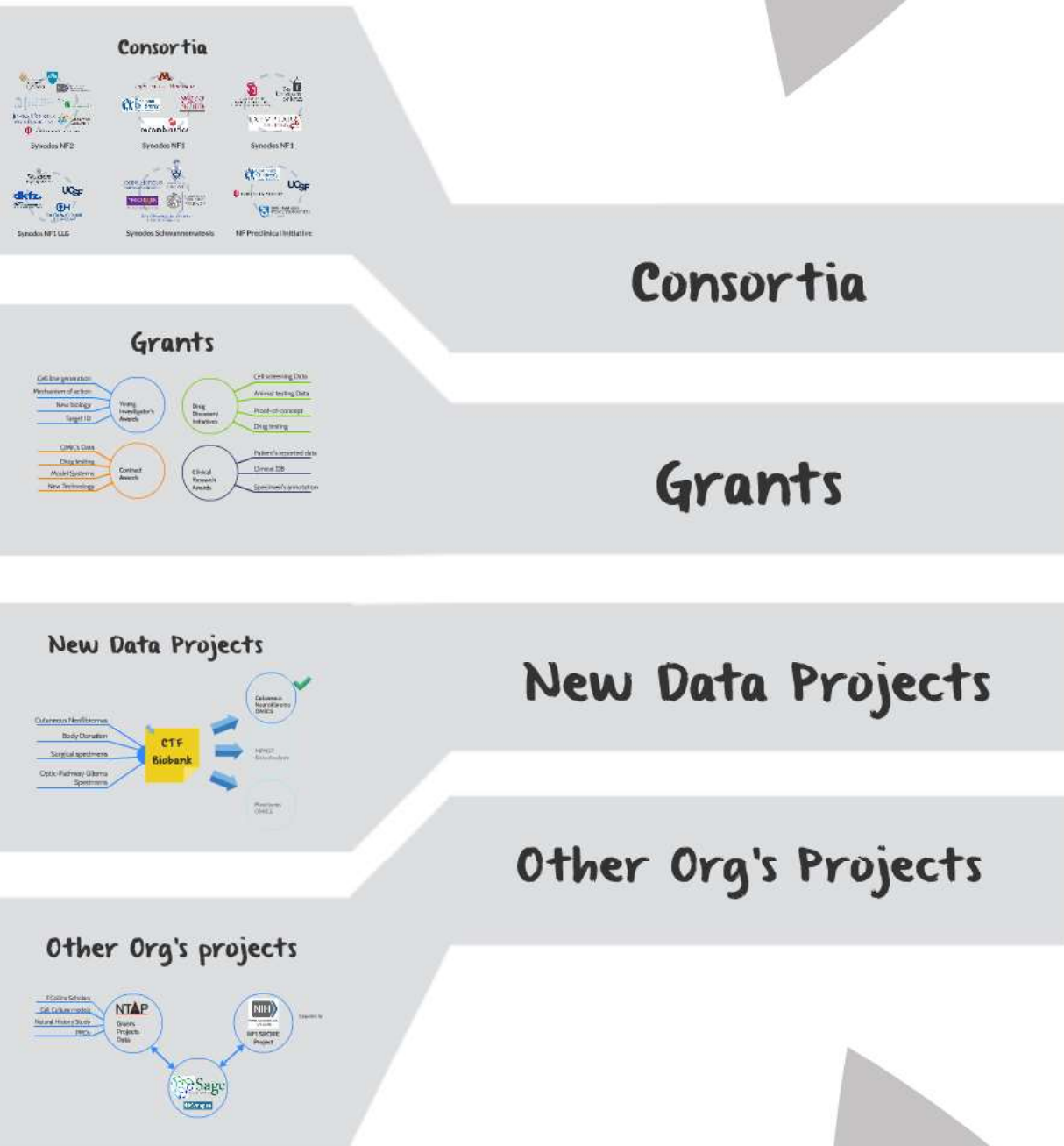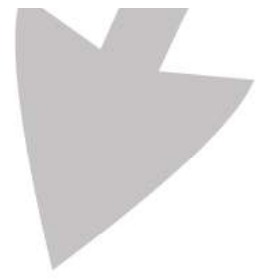
Synodos Schwannomatosis

NF Preclinical Initiative

Grants

Young Investigator's Awards
- Cell line generation
- Mechanism of action
- New biology
- Target ID

Drug Discovery Initiatives
- Cell screening Data
- Animal testing Data
- Proof-of-concept
- Drug testing

Contract Awards
- OMICs Data
- Drug testing
- Model Systems
- New Technology

Clinical Research Awards
- Patient's reported data
- Clinical DB
- Specimen's annotation

New Data Projects

Cutaneous Neofibromas

Body Donation

Surgical specimens

Optic-Pathway Glioma Specimens

CTF Biobank

Cutaneous Neurofibroma OMICS

MPNST Data Analysis

Plexiforms OMICS

Other Org's projects

F.Collins Scholars

Cell Culture models

Natural History Study

PROs

NTAP

Grants
Projects
Data

NIH
National Institutes
of Health

NF1 SPORE
Project

Supported by

CHILDREN'S
TUMOR
FOUNDATION
ENDING NF
THROUGH RESEARCH

Sage
BIONETWORKS

Synapse

CTF Data Source

Consortia

Grants

New Data Projects
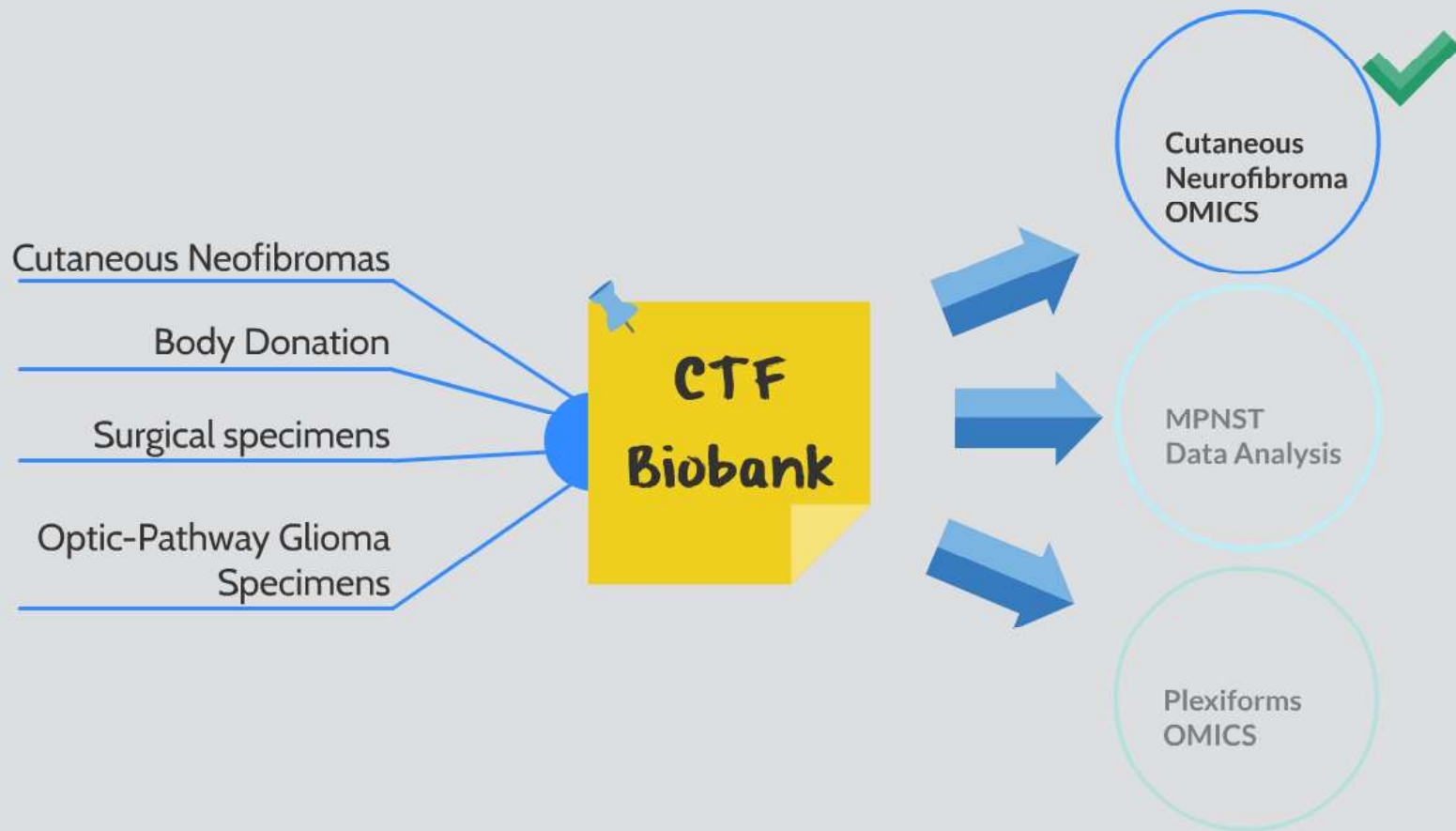
Other Org's Projects

Data layers

Genomic Data

Cell and Animal model data

Clinical phenotypes

Integrated Data

Our knowledge of how
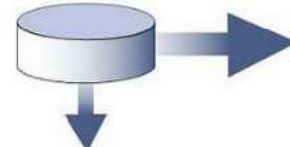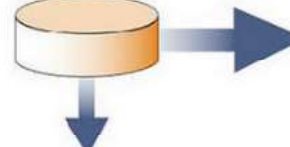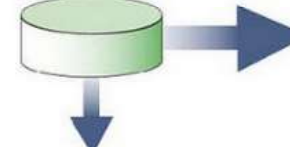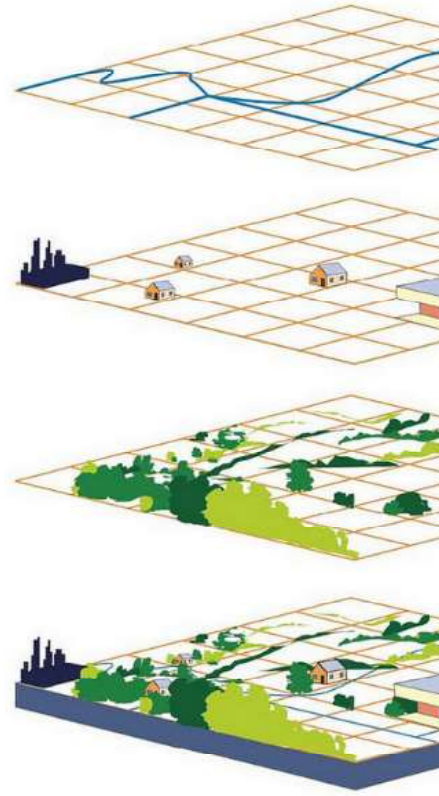
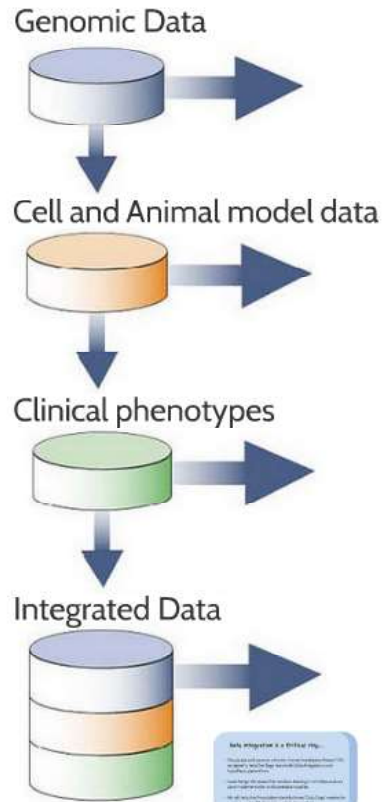# Data integration is a Critical step....

- We placed and sponsor a former Young Investigator Award (YIA) recipient to help the Sage team with Data Integration and hypothesis generation.

- In exchange the researcher receives training in rich data analysis, data implementation and correlation studies.

- He will help the Foundation identify those 'Data Gaps' needed to create meaningful datasets ready to be queried.
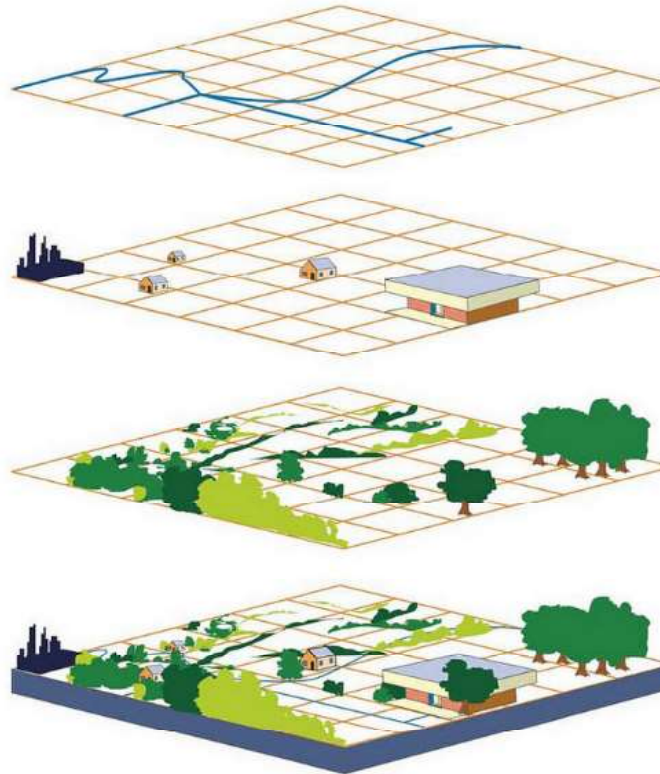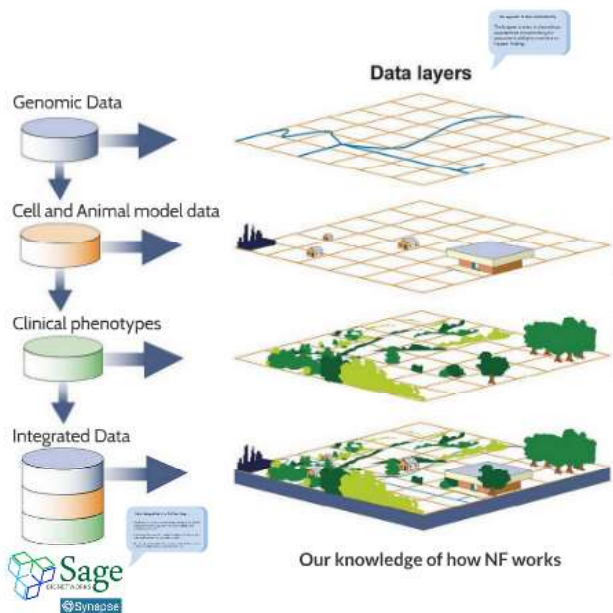
Consortia

Grants

Data Projects

Org's Projects

Genomic Data

Cell and Animal model data

Clinical phenotypes

Integrated Data

Sage
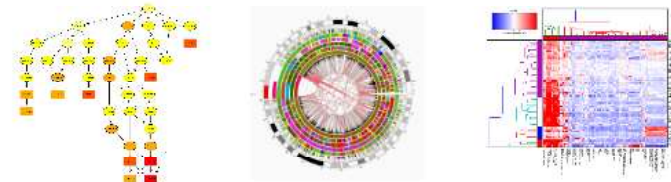BIONETWORKS

Synapse

Data layers

Our knowledge of how NF works
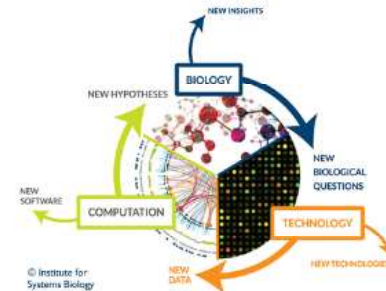
Our approach to Data confidentiality

The balance is to try to disseminate asap without compromising the researcher's ability to maximize on his own findings

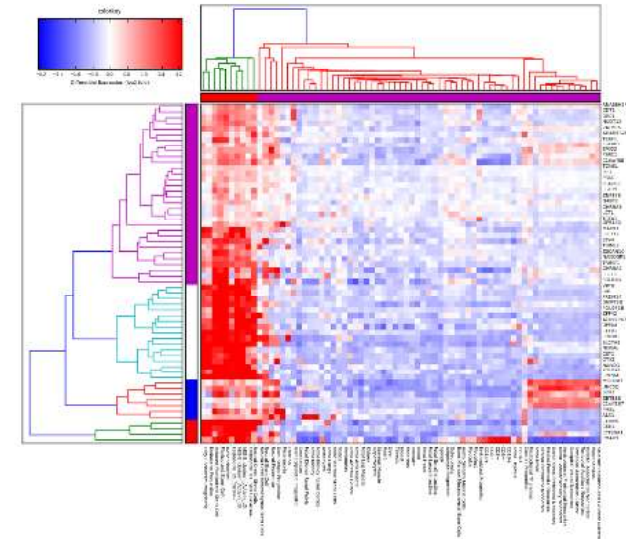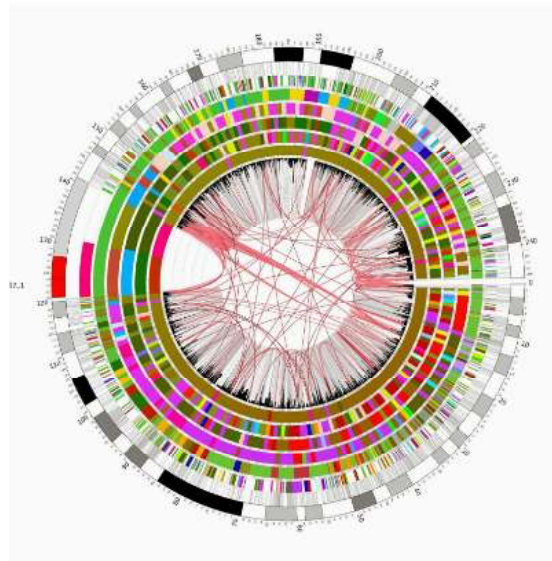New analyses... more understanding...
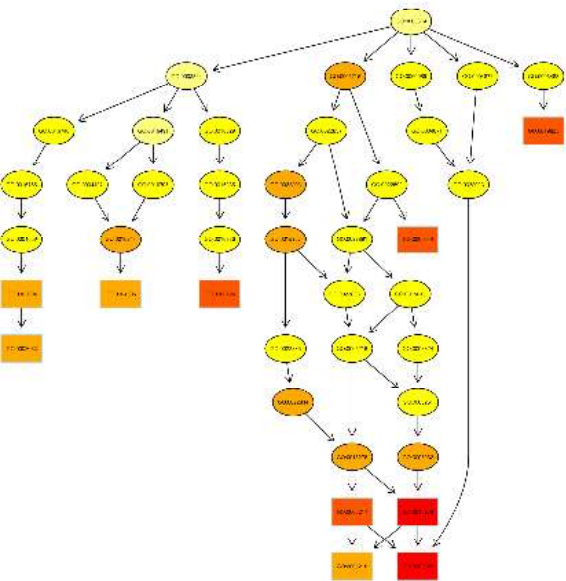
...new treatment options!

Data layers

Genomic Data

Cell and Animal model data

Clinical phenotypes

Integrated Data

Our knowledge of how NF works

NEW INSIGHTS

NEW HYPOTHESES

BIOLOGY

NEW
SOFTWARE

COMPUTATION

NEW
BIOLOGICAL
QUESTIONS

TECHNOLOGY

© Institute for
Systems Biology

NEW
DATA

NEW TECHNOLOGIES

· Identify the gaps in the data
· Launch specific studies to generate the missing data
· Create incentivies for researchers to use the data
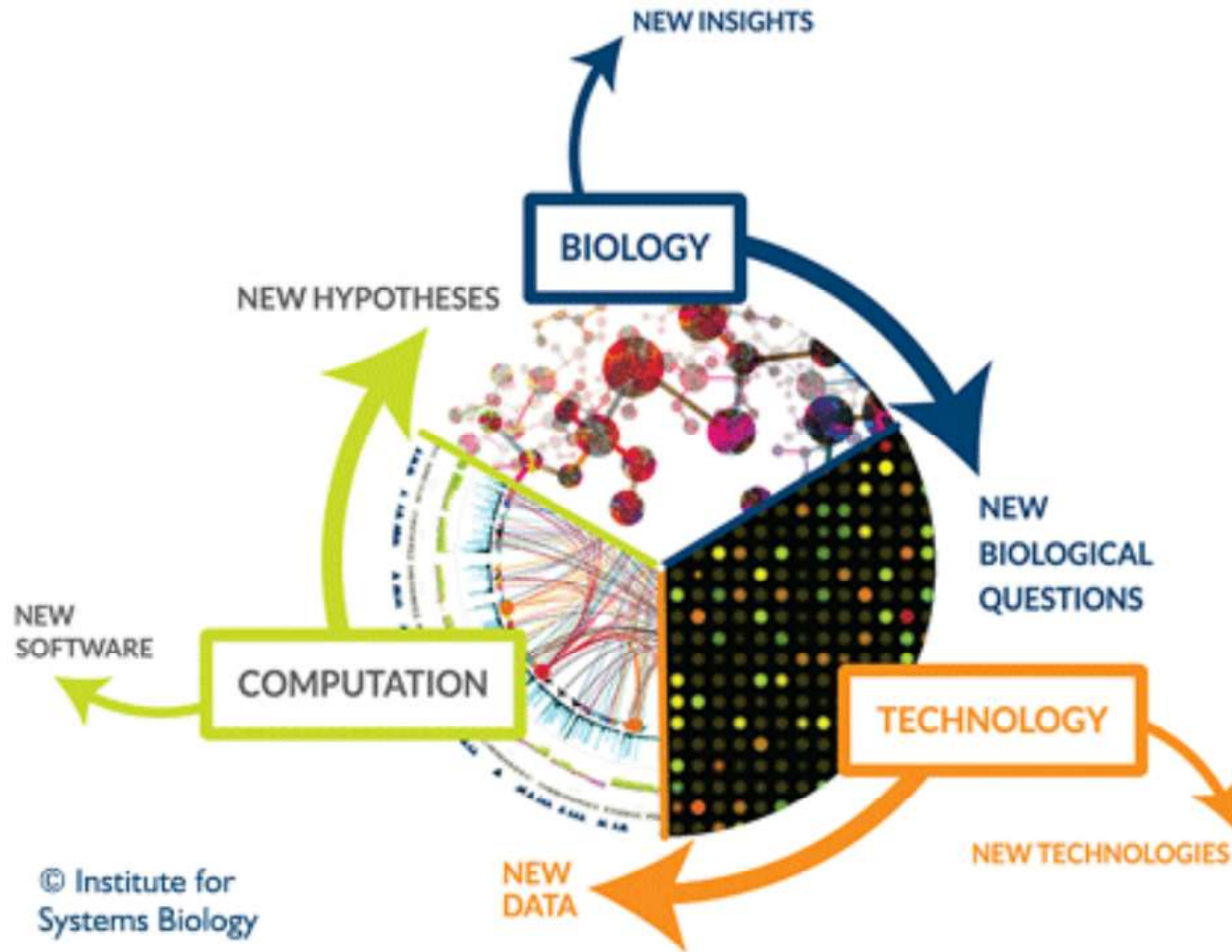 (workshops, challenges, specific RFAs)

New analyses... more understanding...

...new treatment options!
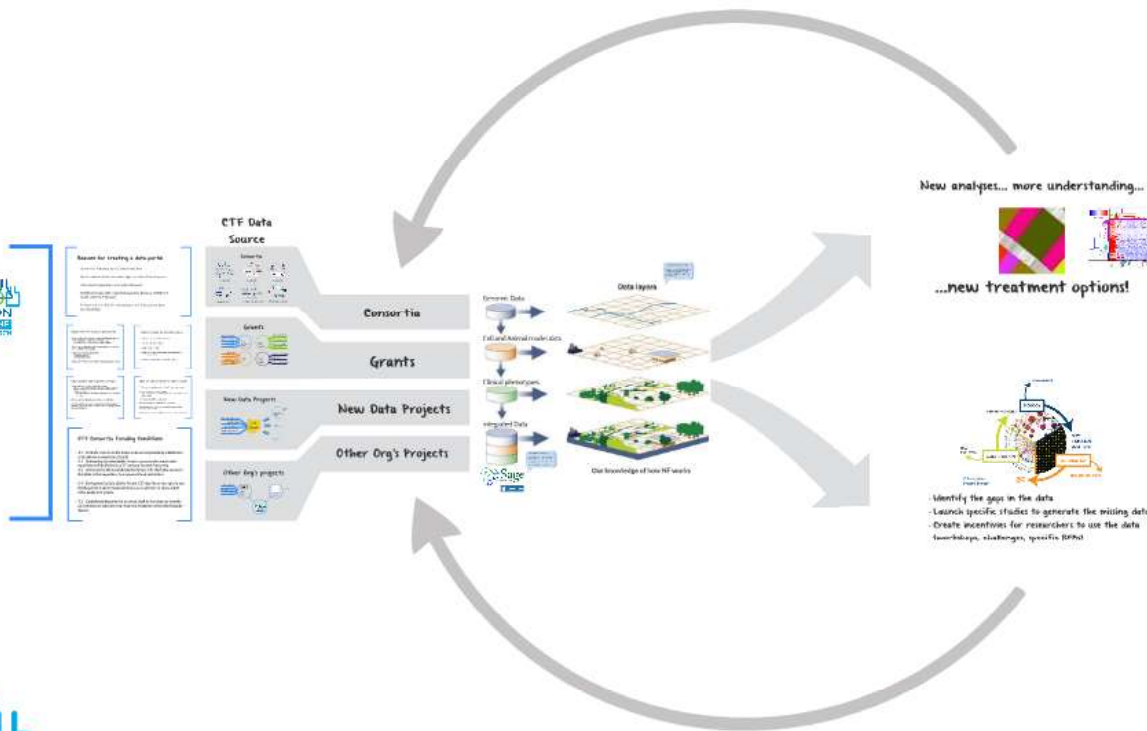
© Institute for Systems Biology

- Identify the gaps in the data
- Launch specific studies to generate the missing data
- Create incentivies for researchers to use the data (workshops, challenges, specific RFAs)

Our aim is to create useful resources and build an open-data ecosystem to promote NF research

in collaboration with

Thank you!