FRANCE MÉDECINE GÉNOMIQUE 2025

aviesan

DATA ANALYZER COLLECTOR (CAD): AN E-INFRASTRUCTURE FOR PERSONALIZED MEDICINE

1

C. CALVIN

CEA – Fundamental Research Division
Christophe.calvin@cea.fr

Forum Teratec 2018 19&20 JUIN JUNE 19&20
ECOLE POLYTECHNIQUE PALAISEAU / FRANCE

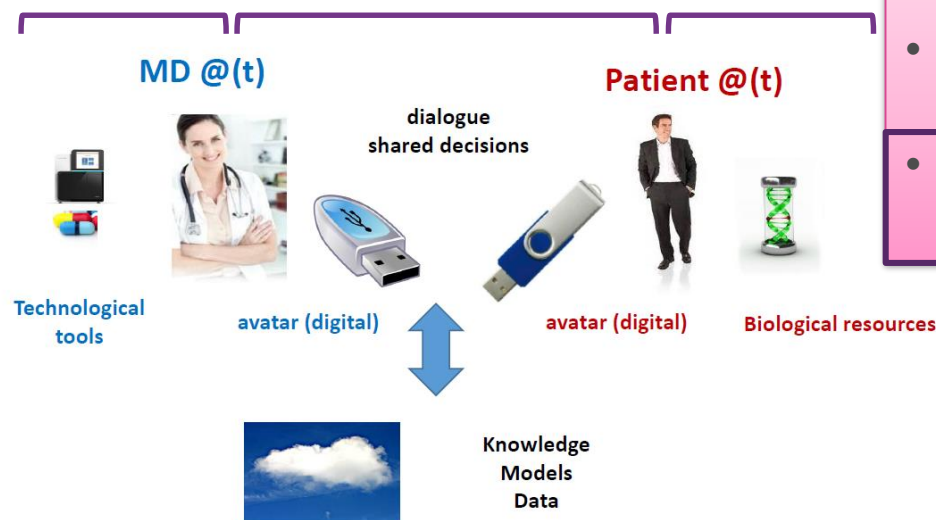# CONTEXT & BACKGROUND

WHAT MEANS "DIGITAL MEDICINE"?

Paradigm change :
- from optimized treatment based on « averaged » patient → personalized diagnostic and therapy

Medicine which uses algorithms :
- to interpret digital patients
- in order to make diagnoses and help to treat real patients

3 pillars

MD @(t)

dialogue
shared decisions

Patient @(t)

Technological
tools

avatar (digital)

avatar (digital)

Biological resources

Knowledge
Models
Data

- **Organic companion:** bio-tech, organ-on-chip
- **Interface life-electronic and robotics** : IoT
- **Digital patient**: use of all medical data → digital avatar

aviesan

FRANCE MÉDECINE GÉNOMIQUE 2025

INTERNATIONAL CONTEXT

Genomic medicine is a field of international competition supported by national plan where:
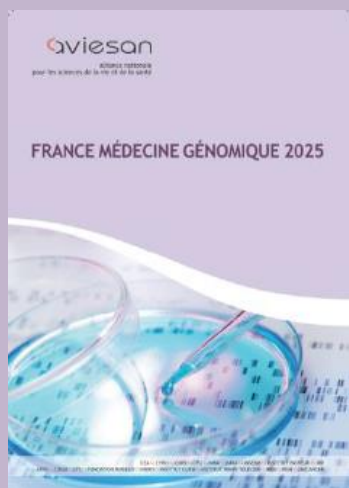
- Clinical application is a common goal

- The development of economic sectors is a strategic issue

PFMG 2025
« *Plan France Médecine Génomique 2025* »

TERMS OF ENGAGMENT LETTER FROM THE PRIME MINISTER

The Prime Minister entrusted Aviesan with the task of examining the current landscape for incorporating genome sequencing in the context of the healthcare pathway by touching on the following four points in his mission statement:

1. Defining the presence and importance of genomic sequencing in current medical practices as well as expected future developments in the coming 10 years.

2. Evaluating France's positioning in the field of genomic research and its role in current health plans and priorities to be implemented in line with national strategies for health and research

3. Evaluating the issues related to innovation and technology transfer and economic growth, taking into account technological aspects, management of large data sets and ethical implications.

4. Proposing a long-term medico-economic model integrating coverage by medical insurance and the establishment of an industrial sector to support such an initiative.
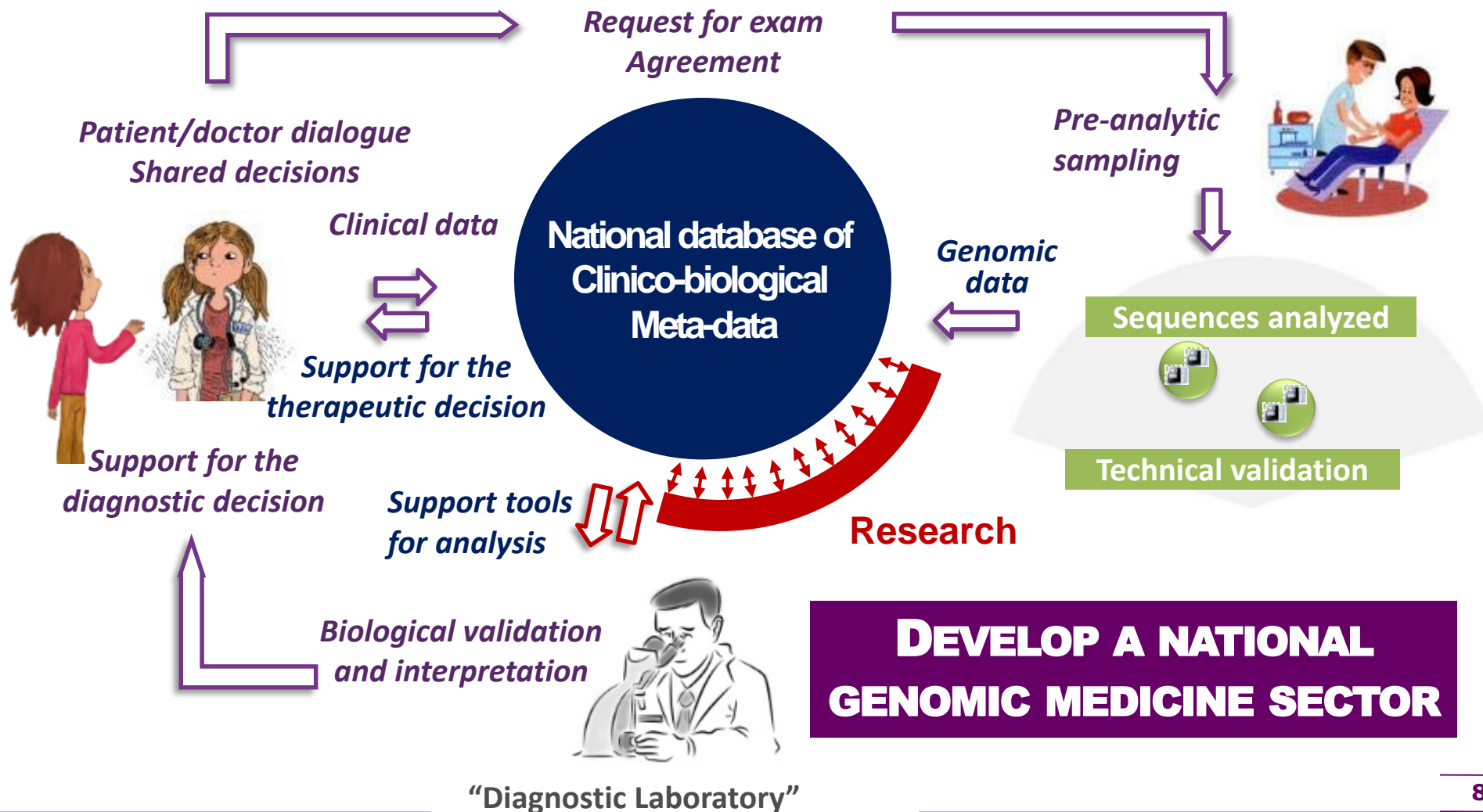
A Co-Constructed National Plan

More than 160 people mobilized under an Aviesan presidency

- Institutional representatives

- transversal competences across fields of research, health and industry

- Research and health agencies

- Central administrations of ministries

- Representatives from the industry (Ariis, IT,…)

- CNAM and HAS

- CGI,

- Toulouse School of Economics…

aviesan
alliance nationale
pour les sciences de la vie et de la santé

GENOMIC MEDICINE FRANCE 2025

CEA ○ CHRU ○ CNRS ○ CPU ○ INRA ○ INRIA ○ INSERM ○ INSTITUT PASTEUR ○ IRD
ARIIS ○ CIRAD ○ EFS ○ FONDATION MERIEUX ○ INERIS ○ INSTITUT CURIE ○ INSTITUT MINES-TELECOM ○ IRBA ○ IRSN ○ UNICANCER

A PLAN ORGANIZED AROUND 3 MAJOR OBJECTIVES – 14 ACTIONS

I- Set up the tools for a genomic healthcare pathway

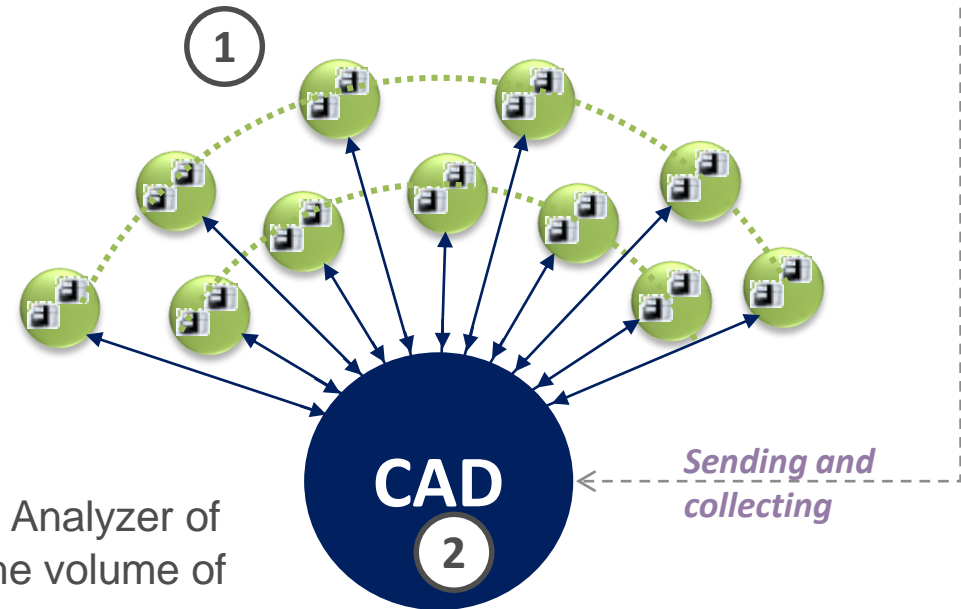II- Ensure these developments in a safe technical & ethical framework

III - Implement monitoring and management tools

FRANCE MÉDECINE
GÉNOMIQUE 2025
aviesan

Action ③ - Allow the integration and use of patient data in the healthcare pathway

Action ① - Creation of a network of sequencing platforms

1

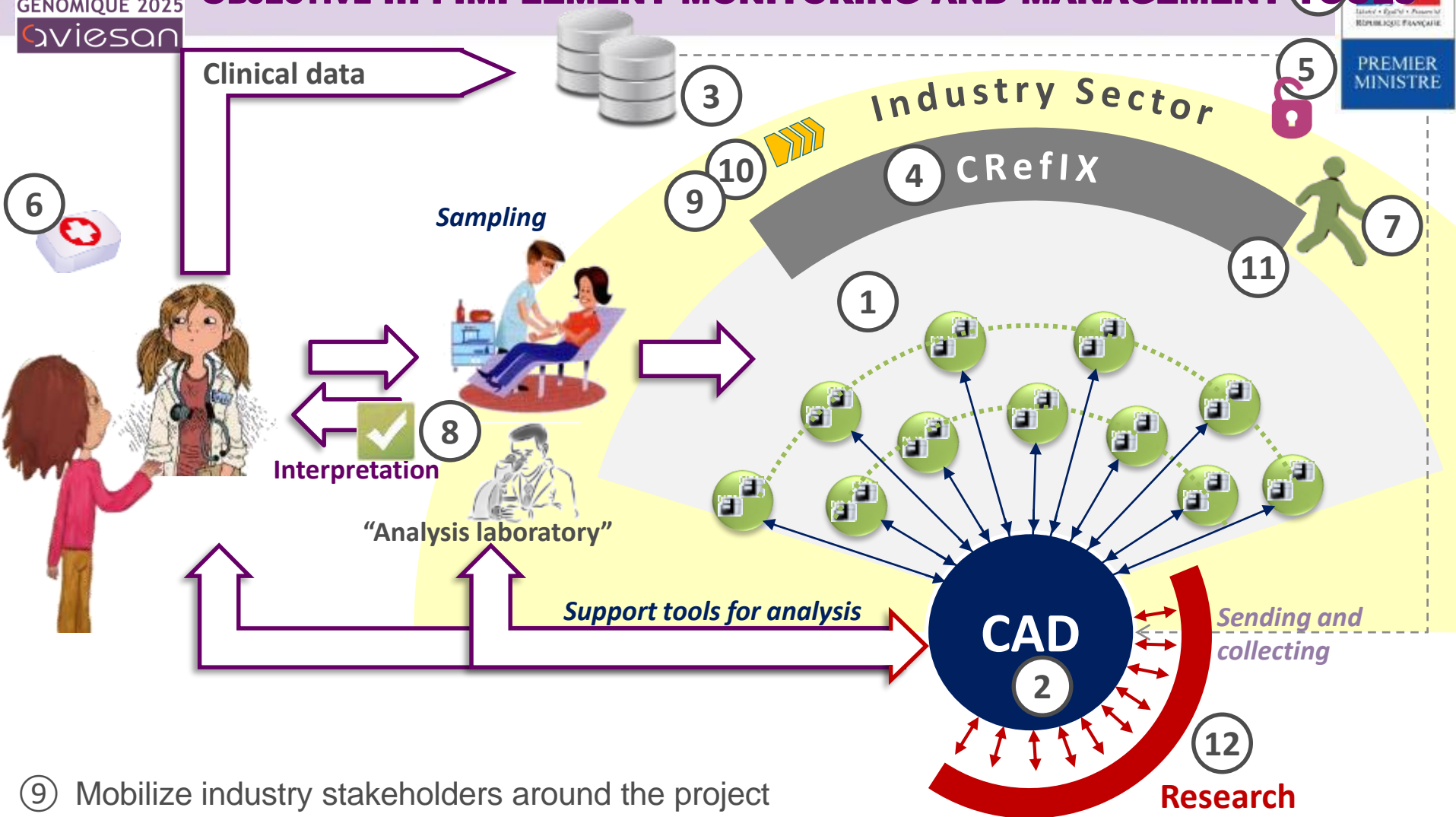*Sending and collecting*

CAD

2

Action ② - Creation of a Central Analyzer of Data (CAD) to process and use the volume of data generated

<u>**Healthcare system: Equal access to sequencing**</u> (110 000 patients 310 000 analyses) rare diseases, cancers (~60 000 patients per year)

④ Set up a center of reference, innovation, expertise and transfer : CREFIX
⑤ Overcome technological, clinical and regulatory barriers encountered along the pathway
⑥ Set up an evaluation and validation system of new indications for access to genomic diagnosis
⑦ Foster new skills and personnel capable of meeting the challenge of analyzing and interpreting the data
⑧ Integrate ethical aspects related to the processing of clinical & genomic data

Clinical data

Industry Sector

CRefIX

Sampling

Interpretation

"Analysis laboratory"

Support tools for analysis

CAD

Sending and collecting

Research

⑨ Mobilize industry stakeholders around the project
⑩ Guide sector activities to address industry issues in the genomic healthcare pathway
⑪ Monitor the developments at the international level
⑫ Implement a program dedicated to health economic aspects
⑬ Organize information, consultation, and involvement of concerned stakeholders
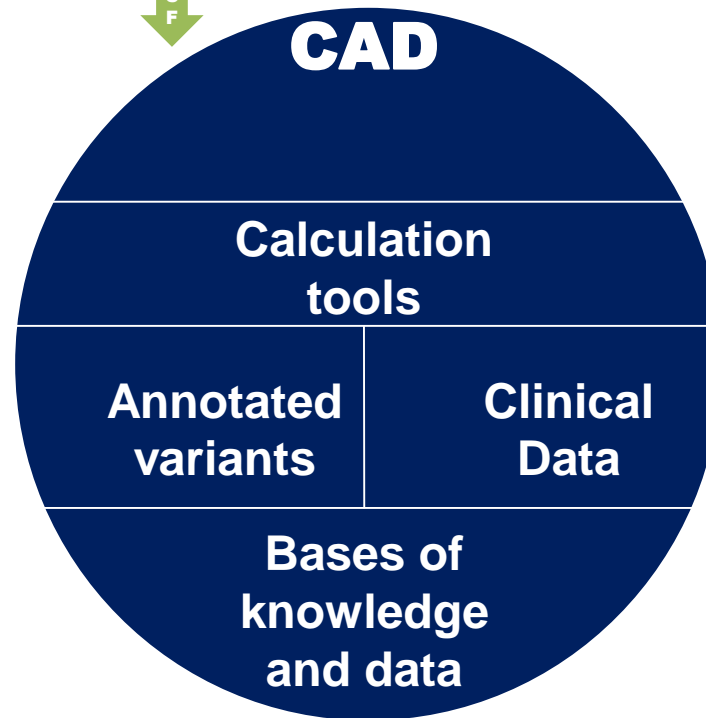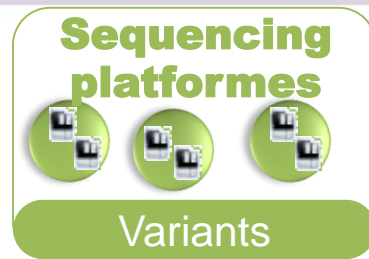⑭ Governance of the Plan
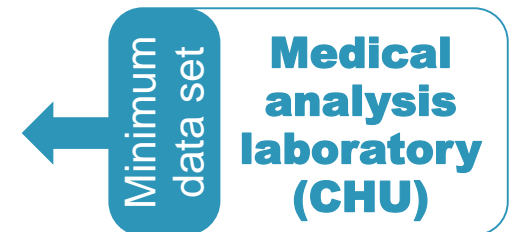
CAD
« Collecteur Analyseur de Données »
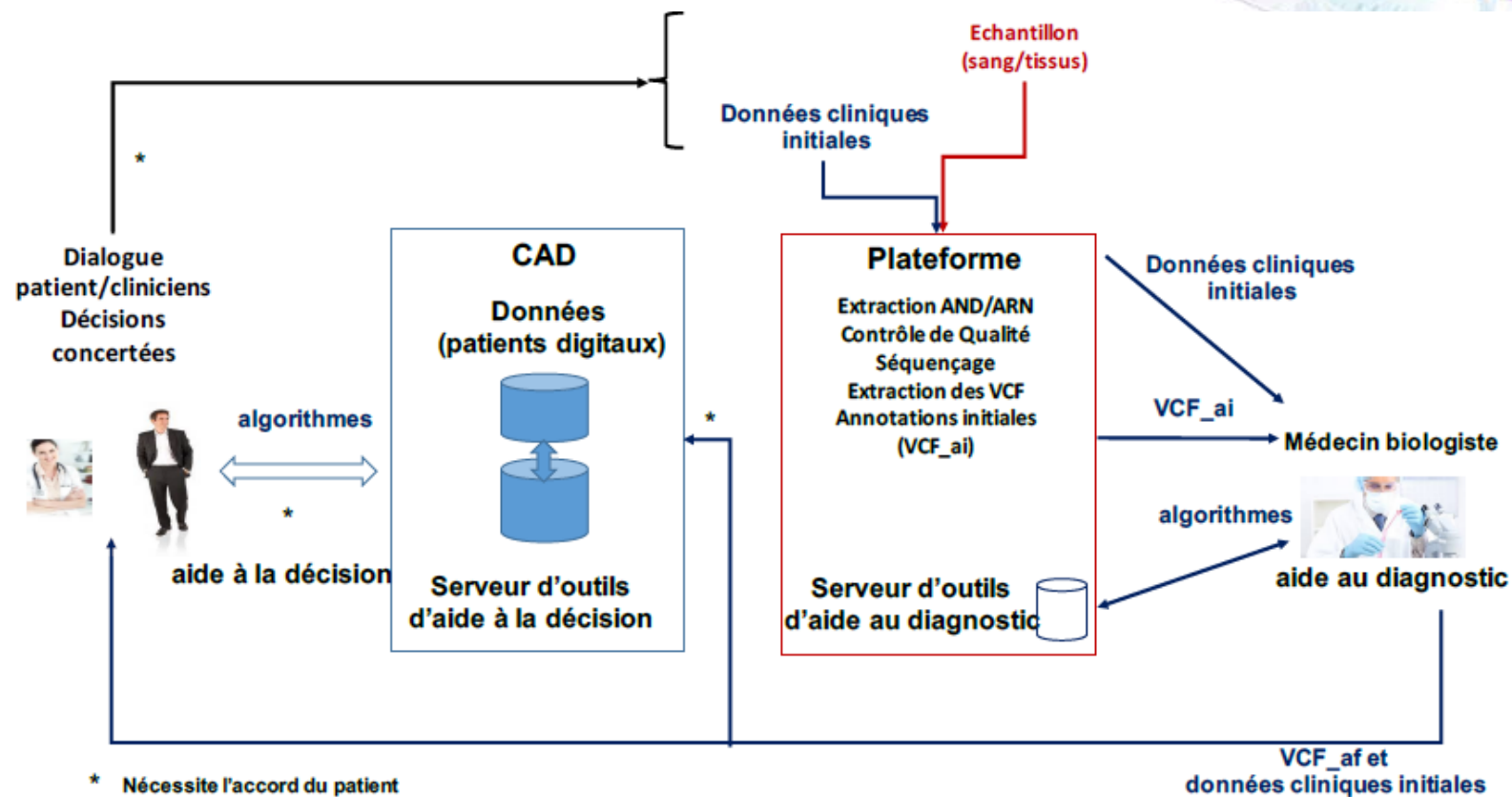
- Three major key points

– From patient to sequence variations list (pre-analytic, sequencing, alignment and variants, annotation)

– Help with diagnosis and medical decision

– Information system global architecture, data security

Sequencing platformes

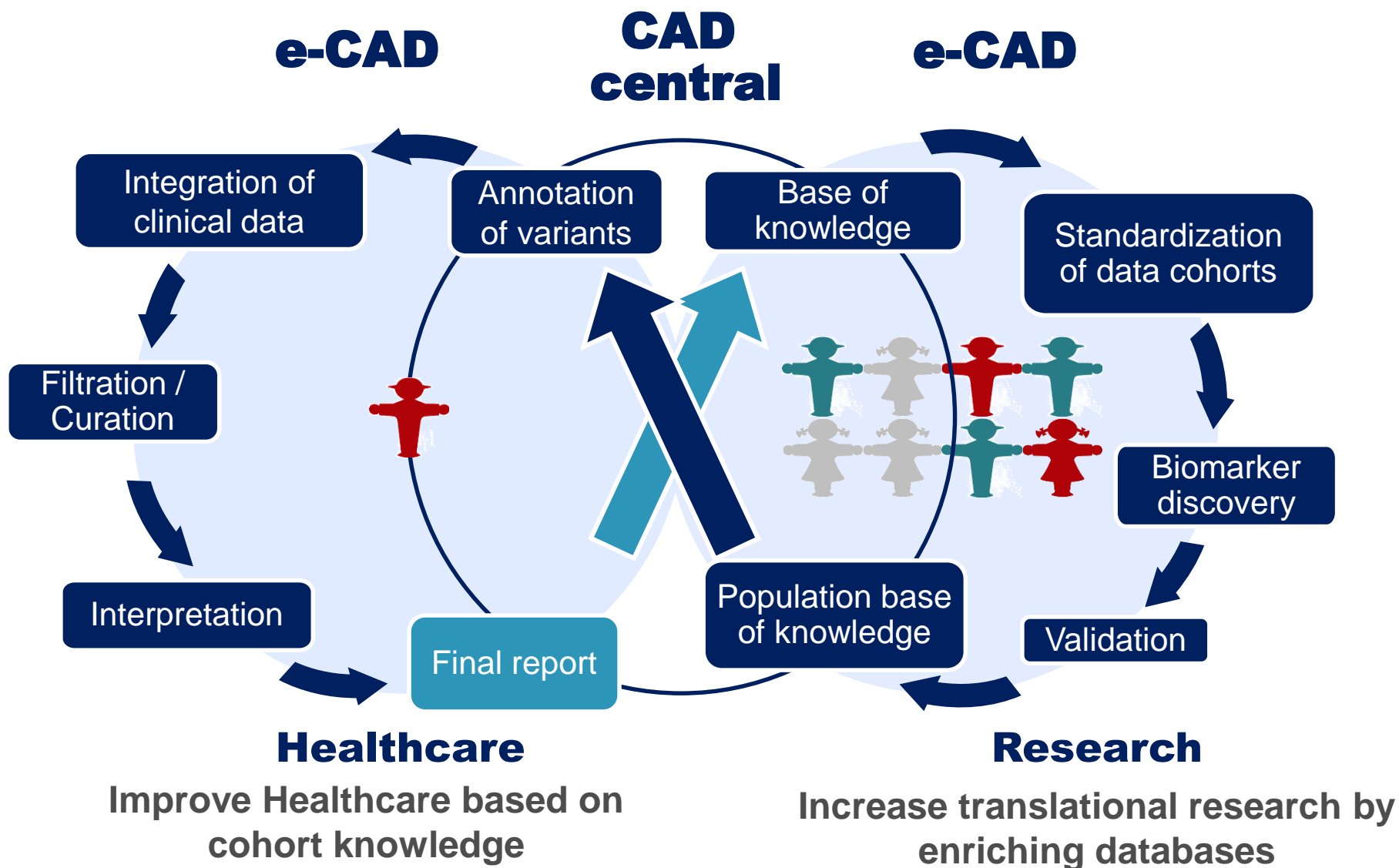Variants

VCF

CAD

Calculation tools

Annotated variants | **Clinical Data**

Bases of knowledge and data

- **Collection and annotations** of genomic data
- Association with **clinical data**

Minimum data set → **Medical analysis laboratory (CHU)**

- **Analytical methods, calculation** associated with large volumes
- **Constitution and management of Healthcare-Research Knowledge Bases**

e-CAD

CAD central

e-CAD

Integration of clinical data

Annotation of variants

Base of knowledge

Standardization of data cohorts

Filtration / Curation

Biomarker discovery

Interpretation

Population base of knowledge

Validation

Final report

Healthcare

Improve Healthcare based on cohort knowledge

Research

Increase translational research by enriching databases

CAD : A SERVICE PROVIDER INFRASTRUCTURE

Providing services to the Platforms

- versioning variant call software

e-CAD orchestrator

- Orchestrator software (monitoring the whole process)
- Tools for diagnostic and therapeutic choices assistance
- Interconnection to databases of open clinical trials
- Interoperability with existing tools and methodologies (DCC, PNMR)
- Portability

Sequencing platforms

Variants

CAD central

Calculation tools

| Annotated variants | Clinical Data |

Bases of knowledge and data
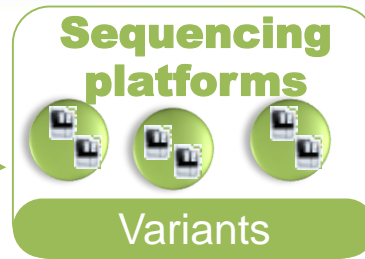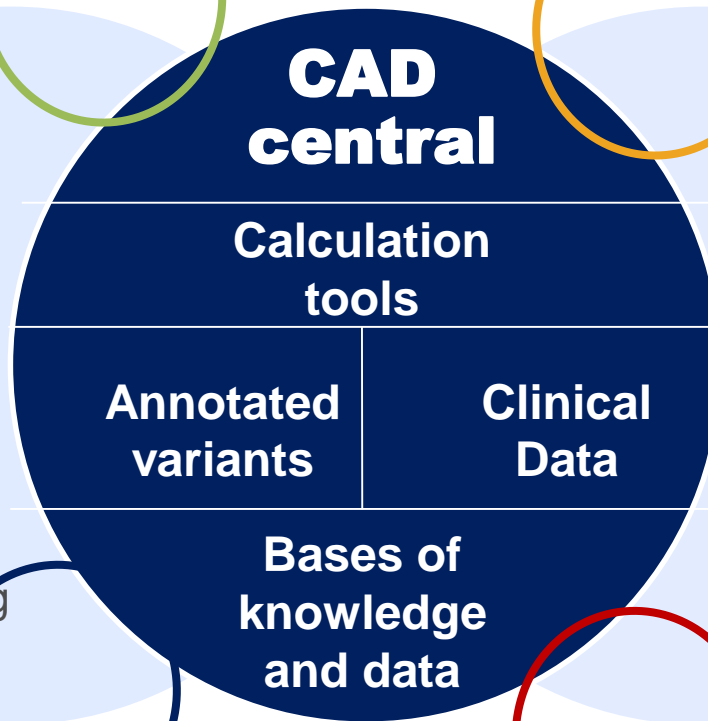
Providing services to industry

e-CAD portal

- Reception desk
- Assistance and advice on project design
- Access to datasets in response to legally validated queries
- Access to tool libraries
- Provision of virtual machines
- Hosting / storage of genomic data and associated data
  - Support and methodological tools
    - Statistical analyses
    - Data transfers

Providing services to clinic

Providing services to Research

NEEDS FROM HARDWARE TO SOFTWARE

1st guess

Storage	Computing	Software tools
> ~70Pb/year	(2048 cores + 8 To)* x10	**Level 1 : Acquisition of varients** • Managment/follow-up • Quality control • Matching of varients • Varient calling • Management of workflow
CAD > ~70Pb/year	**2x** + growth	**Level 2 : Annotation and validation** • Storage • Automatic annotation • Validation / Reporting • Exchange of information **Level 3 : Methodology** • Digital models, data mining
CRef IX	Innovative computing	**Level 4 : Research and development** • Software support (anonymization, coding, compression, parallelization, visualization, certification) • Digital models

Thank you for your attention