BioDynaMo: in-silico biomedical simulations coming of age

Marco Manca, MD

BioDynaMo
Biology Dynamics Modeller

The original definition

- BioDynaMo - A developmental biology simulator

A computational environment to run medium (10^5 cells) to large scale (10^9 cells) simulations of biological systems. Emphasis is put on the definition scale of the models, where cells are treated as agents (machines following rules dictated by current biological theories), rather than being described to the definition of molecules and genes. The goal is to offer a tool for comparison of competing theories, or benchmarking of theories against experimental results.

- GeneROOT

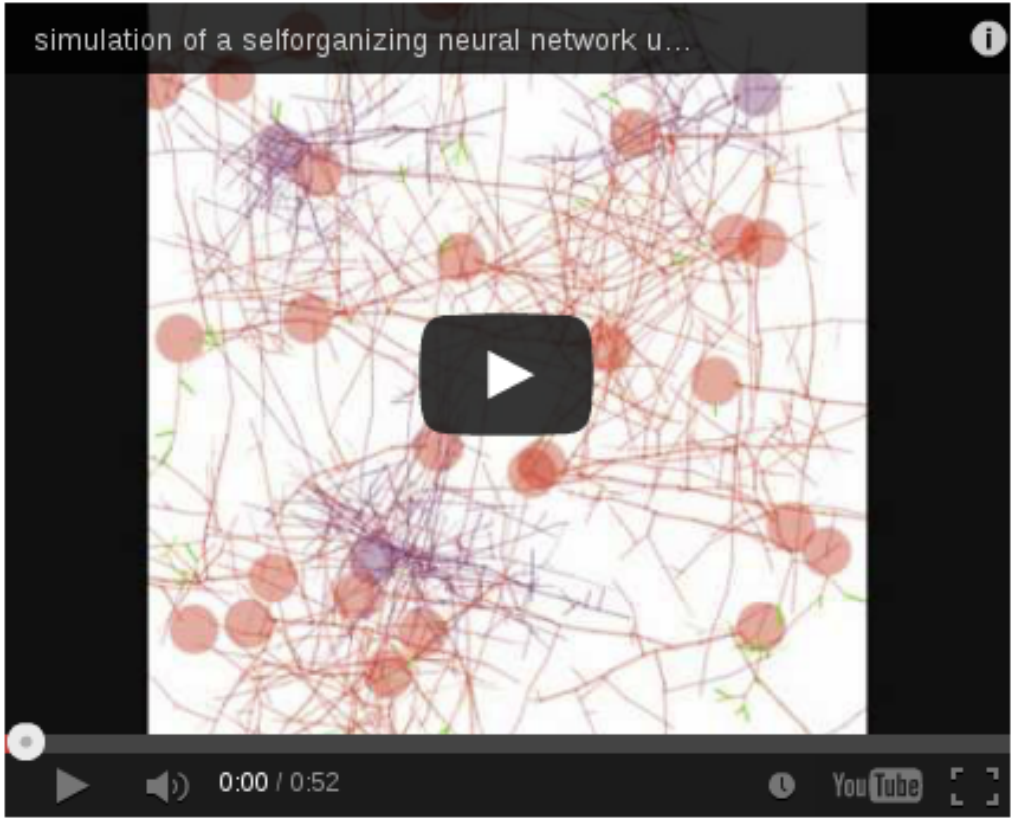
A bigdata-optimized statistical suite for genomics. Emphasis is put on handling large datasets of deep sequenced genomes for the discover and/or validation of structural variants, and their association to phenotypic data. The goal is to introduce a new model for computationally heavy distributed collaborations both optimised for the economies of fundamental and preclinical research, and aware of the sensitivities in biomedical data governance.

- ”Bigdata Epidemiology”

A machine learning solution for novelty and patterns discovery in epidemiological data. Emphasis is on the investigation of parallelisms across diverse domains of application of ML (industrial control, physics, medical epidemiology), and on opportunities of repurposing and crosspollination of innovation. The goal is to pilot a machine learning framework to analyze the volunteered occupational health and medical data of the communities at CERN.

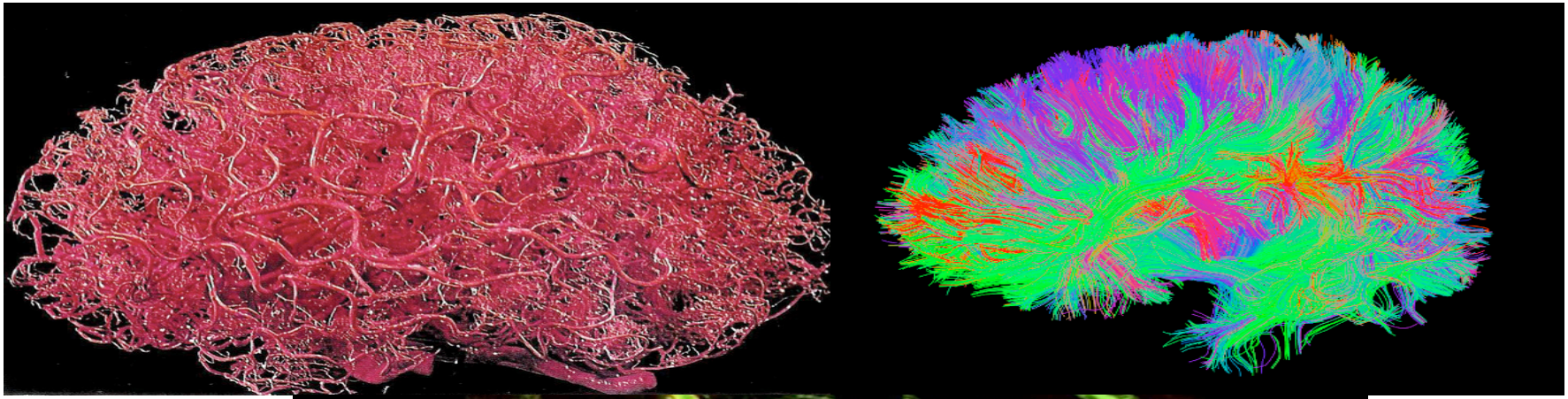
Cx3D – What provoked us

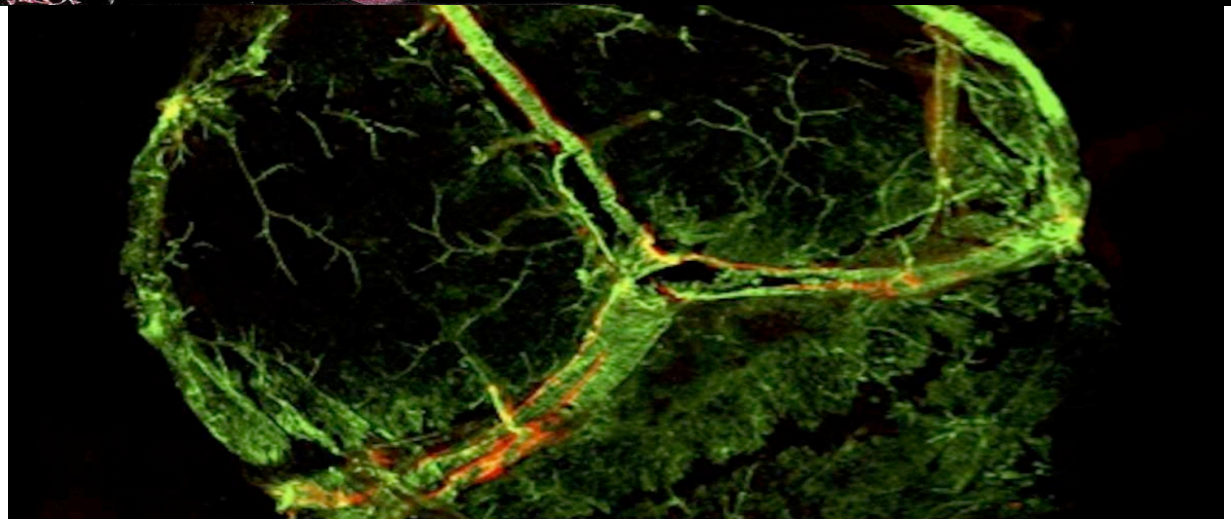
https://youtu.be/il2uc-ZUZQ4



Dwayne Godwin is a neuroscientist at the Wake Forest University School of Medicine. Jorge Cham draws the comic strip Piled Higher and Deeper at www.phdcomics.com.

What is the brain?





Judgement of the heart

- 4th Century B.C. - Aristotle considered the brain to be a secondary organ that served as a cooling agent for the heart and a place in which spirit circulated freely. Interestingly he designated this space in which all the spirits came together as the "common sense."
- 1st Century A.D. - Alexandrian anatomists (e.g. Rufus of Ephesus) provided a general physical description of the brain, identifying some basic structures such as the *pia mater* and *dura mater,* in addition to proposing a basic division of the brain itself.
- 2nd Century A.D. - Roman physician Galen concluded that mental actively occurred in the brain rather than the heart, based on the effects of brain injuries on mental activity. He concluded that the brain was the seat of the animal soul -- one of three "souls" found in the body, each associated with a principal organ. The brain though was thought a cold, moist organ formed of sperm.
- In the Middle Ages, the anatomy of the brain had consolidated around three principal divisions, which came to be called ventricles. Each ventricle hosted a different mental activity, traditionally: imagination in the anterior ventricle, memory in the posterior one, and reason in between.
- 11th Century A.D. - Avicenna wrote that “common sense” was housed in the "faculty of fantasy," receiving "all the forms which are imprinted on the five senses." Memory preserved what common sense received.
- 14th Century A.D. - Mondino de' Liuzzi wrote in his *Anatomy* that common sense lay in the middle of the brain. Aware of the contractions that had proceeded him, he affirmed that "there is only the *sensus communis* which is variously called fantasy and imagination.”
- 16th Centuries A.D. - Leonardo da Vinci drew and dissected the brain. He began to examine the relationship between the brain and the olfactory and optical nerves through experimenting with wax injections that helped him to model the ventricles. He sketched the brain from many different perspectives, looking closely at the ventricles and the origins of the nerves in the medulla. **The more Leonardo looked, the less he was sure about the function of each *ventricle***.
- However, contemporaries such Berengario wrote “In the walls of the ventricles also there is some portion of the *pia mater* that carries blood and spirit, the first to nourish the parts nearby to it, the latter for the operations of the soul”. Alessandro Achillini claimed that the sutures of the cranium allowed the vapors of the brain to escape periodically.
- 17th Century - Thomas Willis published his *Anatomy of the Brain* (1664) and Nicolaus Steno his *Lecture on the Anatomy of the Brain* (1669), rebutting the nonsense of “common sense” by describing the design of the ventricles as the accidental result of the complication of the brain, and by rejecting the “spirits” as non-sense.

Each metaphor reflected the most advanced thinking of the era that spawned it

Bible – humankind made from “clay” infused with spirit (which would explain intelligence)

With the invention of hydraulics, around III cent. BC, human peculiarities started being explained by fluids, “humours”, and this vision persisted for more than 1600 years

In 1500s AD the diffusion of mechanical automata coincided with great thinkers (e.g. Descartes) declaring humans complex machines.
Hobbes suggested that thinking arose from small mechanical motions in the brain

Modern chemistry and discovery of electricity inspired their own analogies… von Helmholtz described the functioning of the brain as a telegraphic network

With the advent of modern computers the analogy switched again: the brain would play the hardware and our thoughts (the mind) would be the software.
In “Language and Communication” (1951) George Miller established the cognitive sciences, which would exploit information theory, computation and linguistics…

In “The Computer and the Brain” (1958) von Neumann flatly stated that the brain function is “prima facie digital”

In Our Own Image (2015), George Zarkadakis

Sleep paralysis through history…

Cell

Hematopoietic Origin of Pathological Grooming in *Hoxb8* Mutant Mice

Shau-Kwaun Chen,[1] Petr Tvrdik,[1] Erik Peden,[1] Scott Cho,[2] Sen Wu,[1] Gerald Spangrude,[2] and Mario R. Capecchi[1,*]
[1]Howard Hughes Medical Institute, Department of Human Genetics, University of Utah School of Medicine
[2]Department of Medicine and Pathology, University of Utah
Salt Lake City, UT 84112, USA
*Correspondence: mario.capecchi@genetics.utah.edu
DOI 10.1016/j.cell.2010.03.055

Review

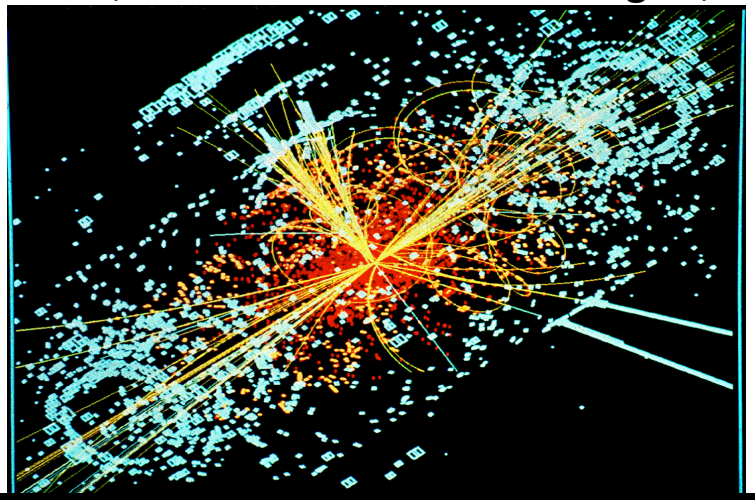
Switch to Standard View

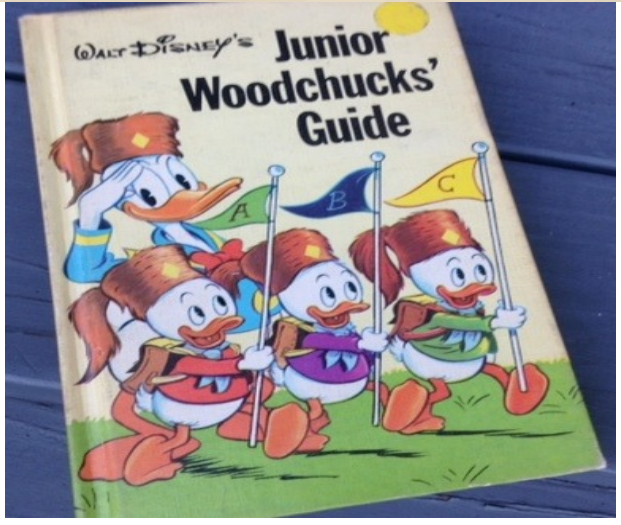
Innate Immunity in the CNS: Redefining the Relationship between the CNS and Its Environment

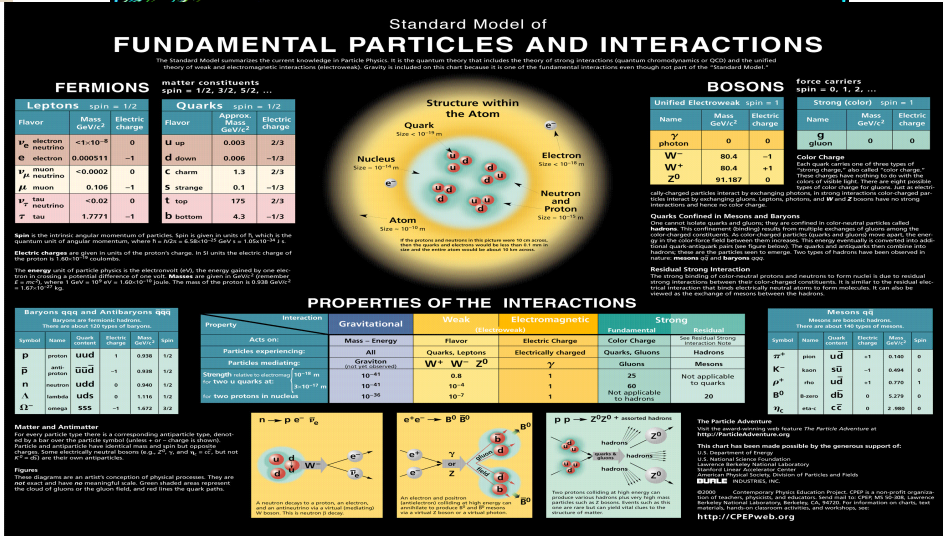
Antoine Lampron, Ayman ElAli, Serge Rivest

DOI: http://dx.doi.org/10.1016/j.neuron.2013.04.005

Every new insight into the brain functioning so far, came from new technologies, it came not from new ideas about the brain.

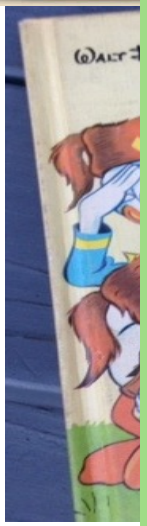






Every ne

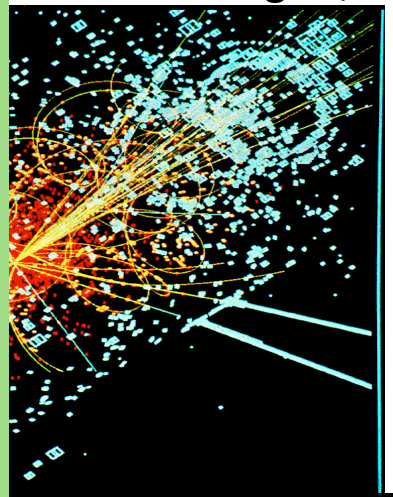
it came

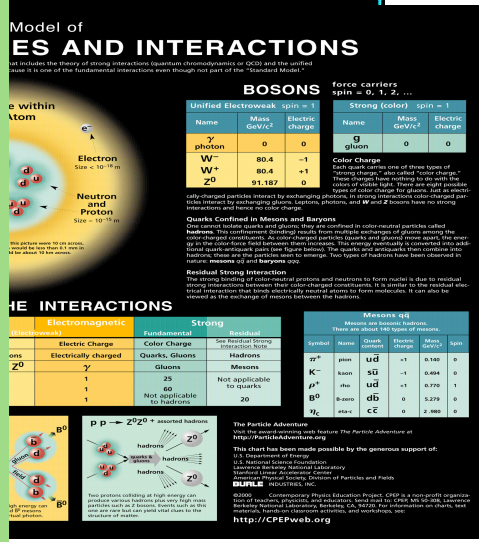


Lab anxiety

facebook.com/pedromics

new technologies,





A brief history of … BioDynaMo

BioDynaMo was originally inspired by Cortex3D, a java agent-based simulator of multiplying and migrating neurons. The original project, brainchild of Prof Rodney Douglas, had long exceeded its original scopes. Developed mostly as a didactic project, it had been successfully exploited to produce experimental data, e.g. supporting the idea that “winner takes all” rules alone can drive the emergence of observed patterns of neuronal arrangement in optical cortex. However, it could not escape its architectural debt and failed to scale up beyond 10^5 cells and 10^6 connections.

It all started as an effort to translate to C++ and modernize the code of Cortex3D. However, as early reflections and conversations with the biomedical community unfolded, it became evident that both on the technical and design level the challenges were such that a complete new project should be born.

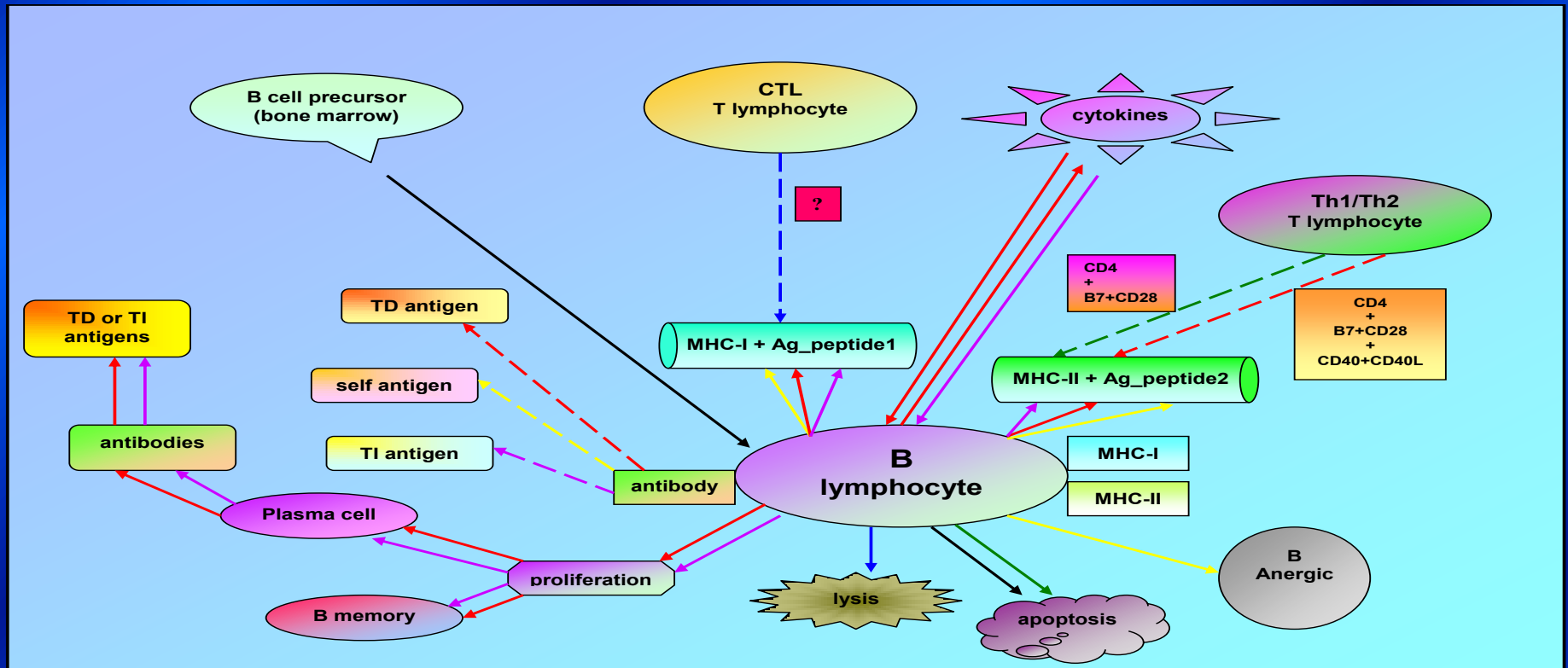
BioDynaMo is a simulation-environment-as-a-service on the cloud. It allows users to define arbitrary (both for shapes, and behavioral rules) agents, and their environments (or to upload one or both of them in an appropriate .xml format), and to run simulations of the development of inputted scenarios without having to worry about the underlying operations of management of computing resources.

ABISS

Agent Based Immune System Simulator

At the very foundation of our simulator of the IS there is the modelling approach based on the concepts of

- (multi-)agency
- structure
- interaction



Pathways of functional status changes for B lymphocytes as implemented in ABISS. Different arrow colours refer to different evolutionary pathways. Dotted lines refer to steric molecular recognition by antibodies or TCRs. Costimulatory molecules and related receptors are indicated when necessary for state changes. The cytotoxic activity of CTL T lymphocytes against B cells is a working hypothesis of the model.

Competing for the moon…

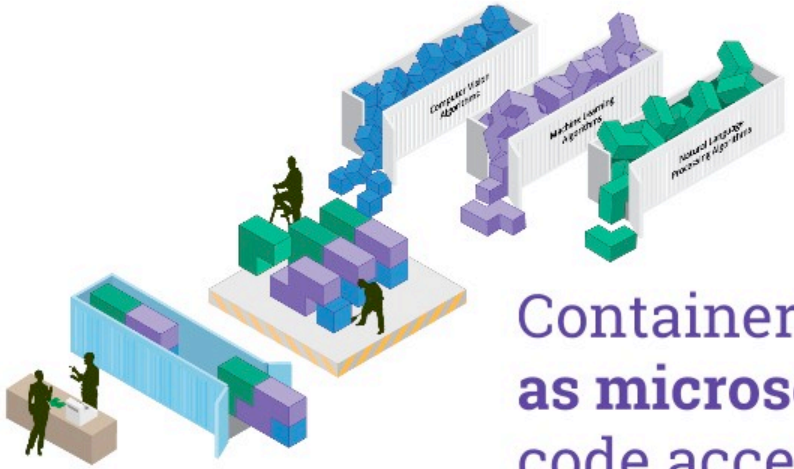
A non exhaustive list of software solutions for computational neurology…

- BRIAN, a Python based simulator
- Budapest Reference Connectome, web based 3D visualization tool to browse connections in the human brain
- DigiCortex, DigiCortex project implementing large-scale simulation and visualization of biologically realistic cortical neurons, synaptic receptor kinetic, axonal action potential propagation delays as well as long-term and short-term synaptic plasticity.
- EDLUT, a simulation software for large-scale neural networks and real-time control systems.
- Emergent, neural simulation software.
- GENESIS, a general neural simulation system.
- ModelDB, a large open-access database of program codes of published computational neuroscience models.
- MCell, Particle-based Monte Carlo simulator of microphysiology and cell signaling.
- NeMo, a C++/CUDA-based, high-performance spiking neural network simulator, intended for large-scale real-time simulations, and with APIs for C++, C, Python and Matlab.
- Nengo, a Python scriptable, GUI simulator for large-scale neural models
- NEST, a simulation tool for large neuronal systems.
- Neuroconstruct, software for developing biologically realistic 3D neural networks.
- NEURON, a neuron simulator also useful to simulate neural networks.
- SNNAP, a single neuron and neural network simulator tool.
- ReMoto, a web-based simulator of the spinal cord and innervated muscles of the human leg.
- TopoGraphica, Topographica is a software package for computational modeling of neural maps. The goal is to help researchers understand brain function at the level of the topographic maps that make up sensory and motor systems.Topographica is intended to complement the many good low-level neuron simulators that are available, such as Genesis and Neuron. Topographica focuses on the large-scale structure and function that is visible only when many thousands of such neurons are connected into topographic maps containing millions of connections.

…and let’s not get started with MatLab, Octave, and simply custom javas

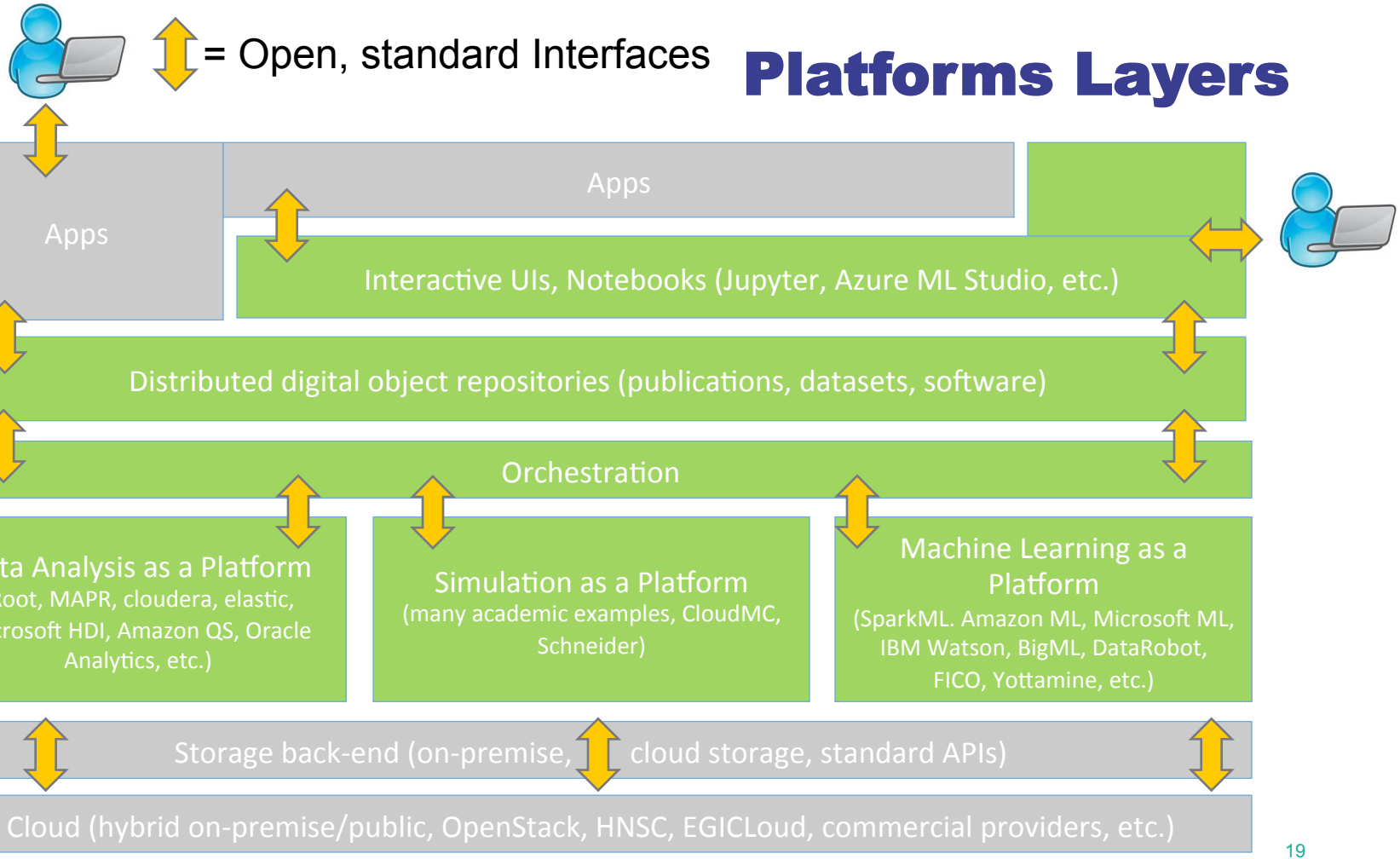
Chose your poison

Let’s play LEGO



Containerizing **algorithms as microservices** makes code accessible via an API, and hosted on scalable, serverless infrastructure in the cloud

Platforms Layers



Challenges become opportunities… BioDynaMo

BioDynaMo faces several challenges, some are technological, others are of design of use, e.g.:

1. Balancing the load of computation to maximize performances should exploit the locality of agents, but be able to identify and deal with long distance dependences;
2. Models can expand or deflate during the simulations, and the behaviors to be simulated can show sudden bursts in/for short time windows;
3. The instructions associated to the model can be rather long and heterogeneous, or short simple and repeated;
4. Computing precision should be optimized to speed up computation and reduce divergence and hardware dependencies, while keeping the simulation meaningful;
5. Models should systematically be tested for parameters sensitivity;
6. Homologous models should be benchmarked and profiled, heterologous ones should be clustered and highlighted to offer educational opportunities;
7. Relevant data and publications should be easily accessed/acquired and interacted with, for verification, correlation, or any other post-analytical purpose;
8. Models and their results should be searchable, and explorable according to permissions;
9. Models should be monitored in “real time” both at the numerical level and in visualization, to be able to interact with them, or to decide to halt them for divergence.

Phase I and prototype…

BioDynaMo will be prototyped by developing a standalone installable application for desktop workstations and containers.

This prototype will already boast many features that are planned for the final product:

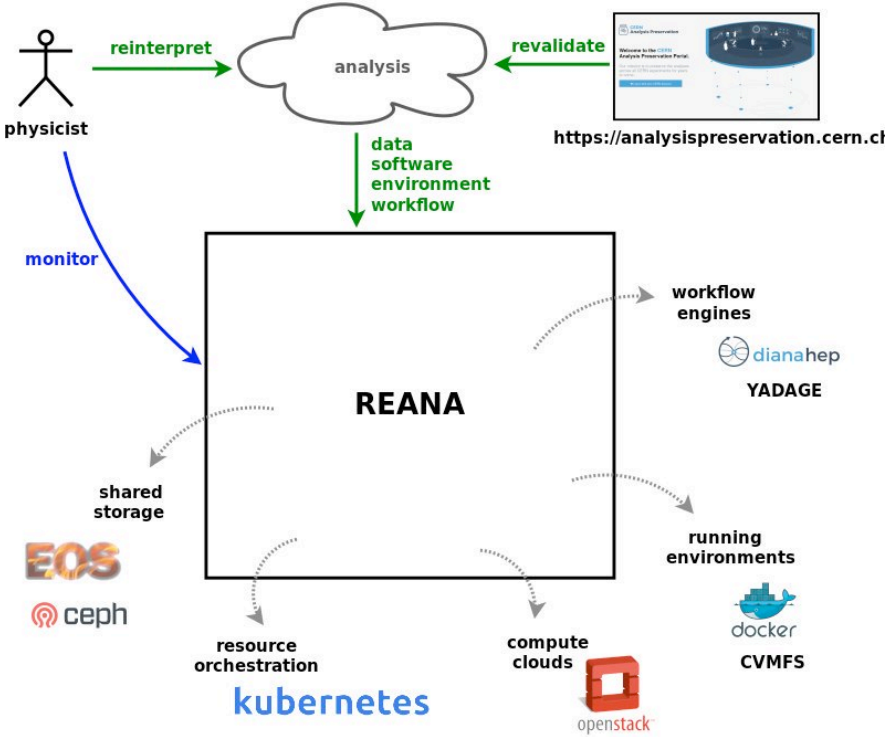
1. Optimized load balancing for multi-core machines and multi-CPU distributed clusters
2. Near real-time visualizations of the models based on ParaView libraries
3. C++ interpreter (CLING), and I/O inherited from ROOT
4. Octree optimization for volumes cutting and parallelization on distributed clusters
5. Handler (import/export, editor) for NeuroML models
6. APIs/wrappers to interact with complementary software (e.g. NEST)
7. Minimal GUI

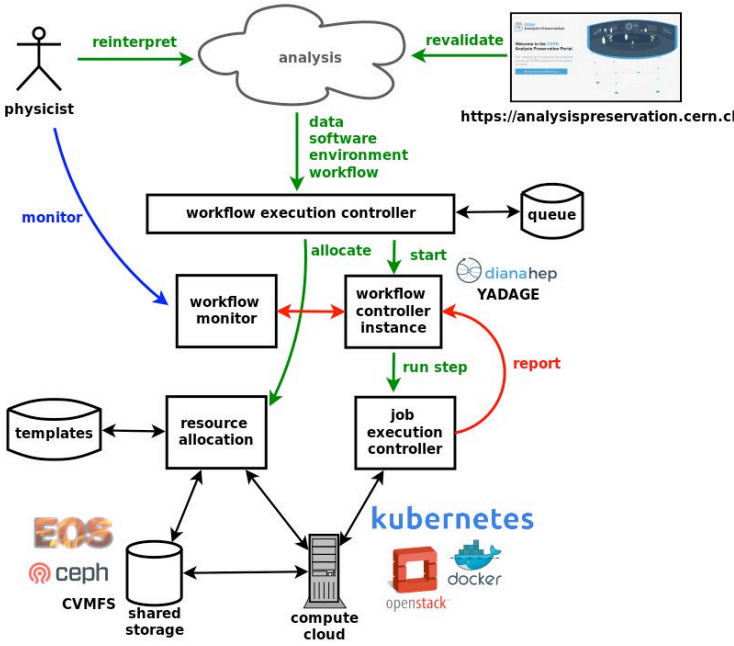
Developing this prototype will offer the opportunity to start working on some of the interesting scientific and technical challenges:

- Optimized parallelization on heterogeneous hardware
- Optimized clustering and allocation of instructions for fast/slow elements of the models
- Static vs Dynamic precision of computation
- Benchmarking with external data

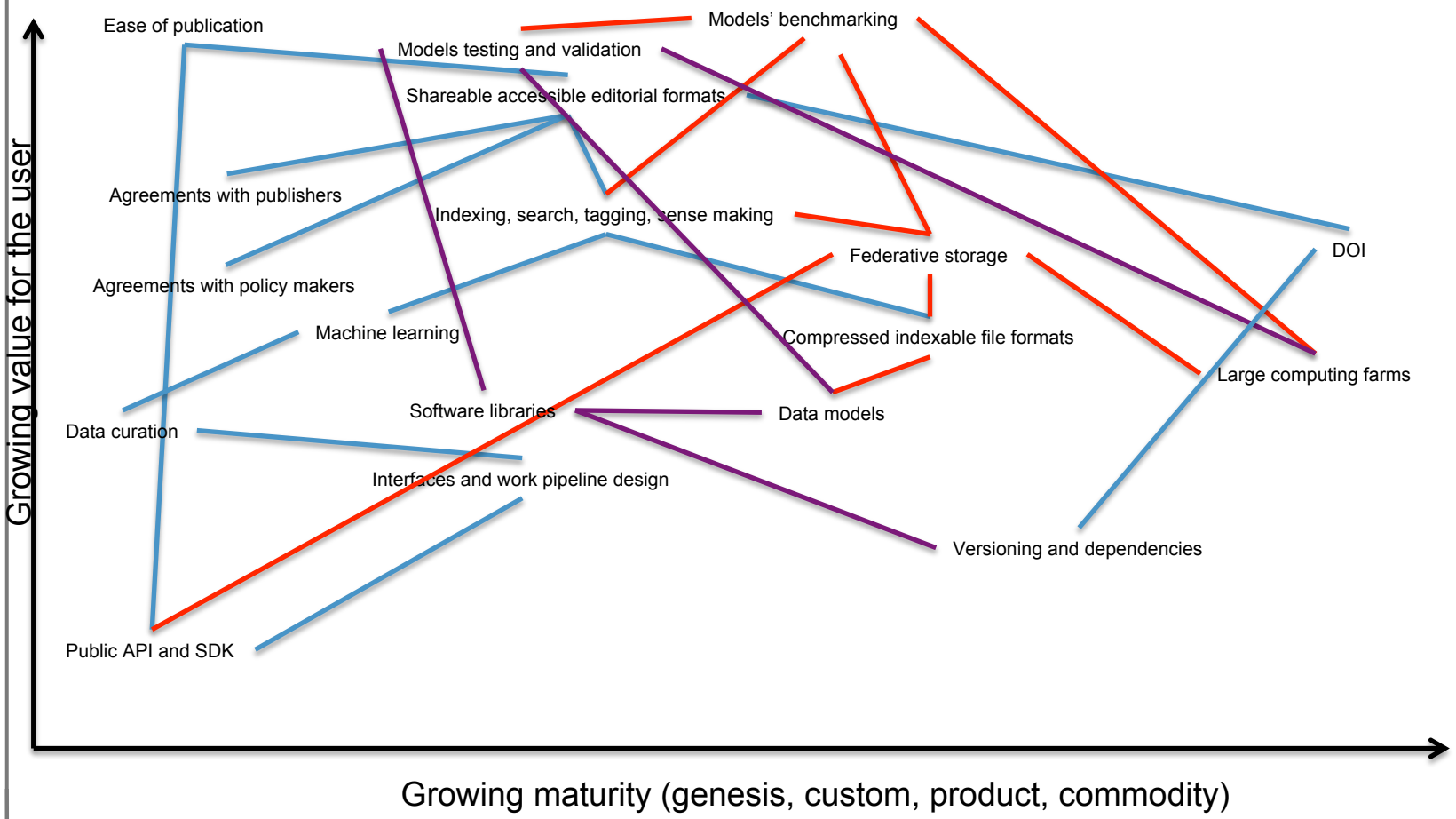
Maybe most importantly, it will offer the opportunity and tools to nurture a growing engagement of the user community.

Using, reusing, misusing…





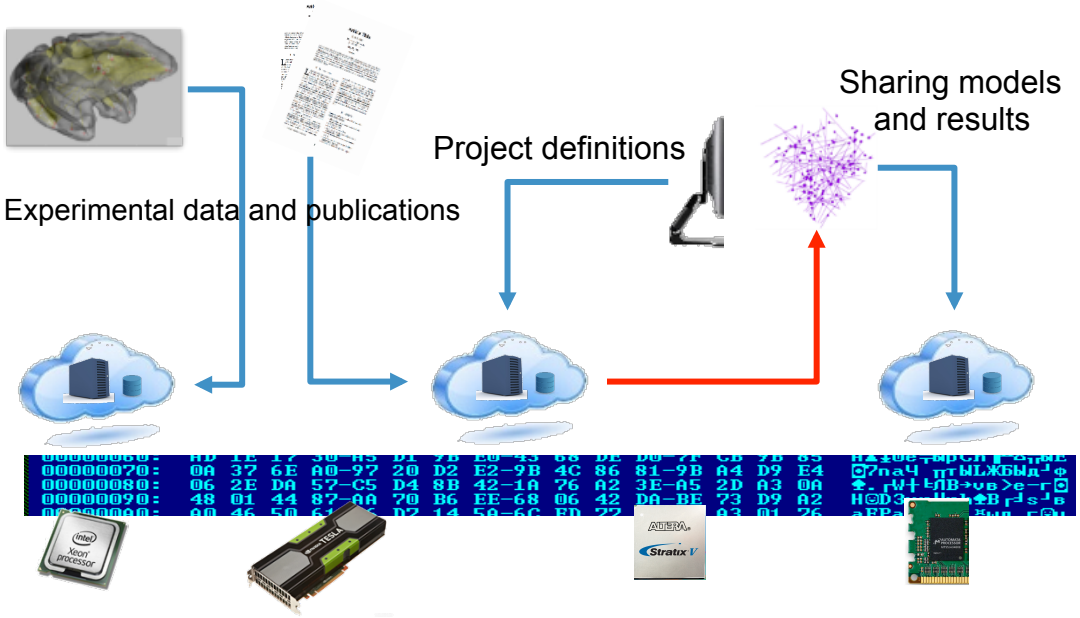

HELP!



BioDynaMo

Founding Members:

CERN
Newcastle University (Neuroinformatics Institute)
Innopolis University
Kazan University
Intel



BioDynaMo

Design and implementation of an architecture-neutral, large-scale biological development simulation platform

Cloud backend able to dynamically scale or shrink with the simulated model

Integration and sharing of experimental data (e.g. functional MRI images), textual publications, models, results

Cheaper to operate compared to more sophisticated HPC solutions and more expandable

Current status: Prototype, initial development of a software stack optimized for various accelerated architecture

Seeking collaborations: developers, domain experts, testers, use cases

CERN openlab

Ecosystem Effect

The Garden as a Metaphor for Platforms

Concerts, games, suggestion box (Events, feedback)

People walking, playing, fishing (Large communities)

Kiosks (Third-party services and extensions, Marketplace)

Opening and closing times, don't walk on the grass, no fishing on Tuesdays (Policies)

How many people came today? What did they do? (Statistics, trends, predictions)

The Mayor (Funding Agencies)

Park manager, guards (Governance)

Different plants and flowers (Commodity services)

Insects and animals (Small communities, micro services)

Seeds (Pilot projects)

Weeds (Failed projects)

Sellers of fertilizers, equipment, furniture, electricity, water (Infrastructure providers)

Gardeners, botanists (Technical experts)

System architecture
Management

Alberto Di Meglio
CERN

Fons Rademakers
CERN

Technical coordination
Development and optimization

Manuel Mazzara
Innopolis

Data structures
DataViz

Roman Bauer
NCL

Marcus Kaiser
NCL

Scientific coordination

Beta testing and design

Lukas Breitwieser
CERN

Ahmad Hesam
CERN

Leonard Johard
Innopolis

Vasileios Vavourakis
UCL

Marek Oropallo
UniBO

Jean De Montigny
NCL

System architecture
Management

Manuel Mazzara
Innopolis

Roman Bauer
NCL

Marcus Kaiser
NCL

Alberto
CERN

Scientific co dinati n

Techn
Develo nd optimization

Dat tructures
Dat Viz

et testing and design

BioDynaMo

Biology Dynamics Modeller

Lukas Breitwieser
CERN

Ahmad Hesam
CERN

Leonard Johard
Innopolis

Vasileios Vavourakis
UCL

Marek Oropallo
UniBO

Jean De Montigny
NCL

A New Interpretation of Information Rate

By J. L. KELLY, JR.

(Manuscript received March 21, 1956)

If the input symbols to a communication channel represent the outcomes of a chance event on which bets are available at odds consistent with their probabilities (i.e., "fair" odds), a gambler can use the knowledge given him by the received symbols to cause his money to grow exponentially. The maximum exponential rate of growth of the gambler's capital is equal to the rate of transmission of information over the channel. This result is generalized to include the case of arbitrary odds.

Thus we find a situation in which the transmission rate is significant even though no coding is contemplated. Previously this quantity was given significance only by a theorem of Shannon's which asserted that, with suitable encoding, binary digits could be transmitted over the channel at this rate with an arbitrarily small probability of error.

Questions are welcome…

Everything that happens during question time, stays in the question time ;)