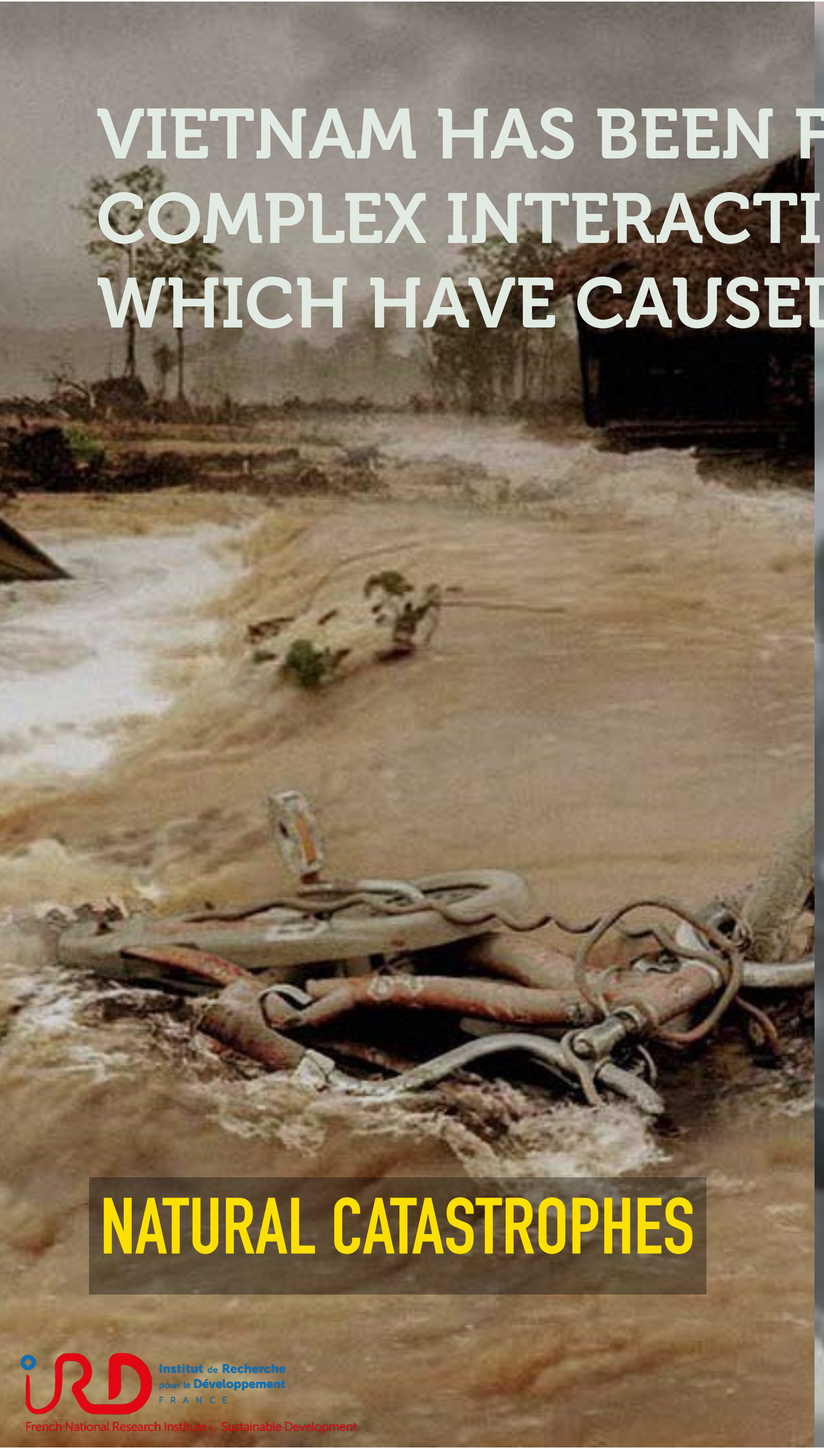
COMPUTER MODELING, OPEN SOURCING AND OPEN SCIENCE

A. DROGOUL, IRD, VIETNAM
alexis.drogoul@ird.fr
17 FEB 2021

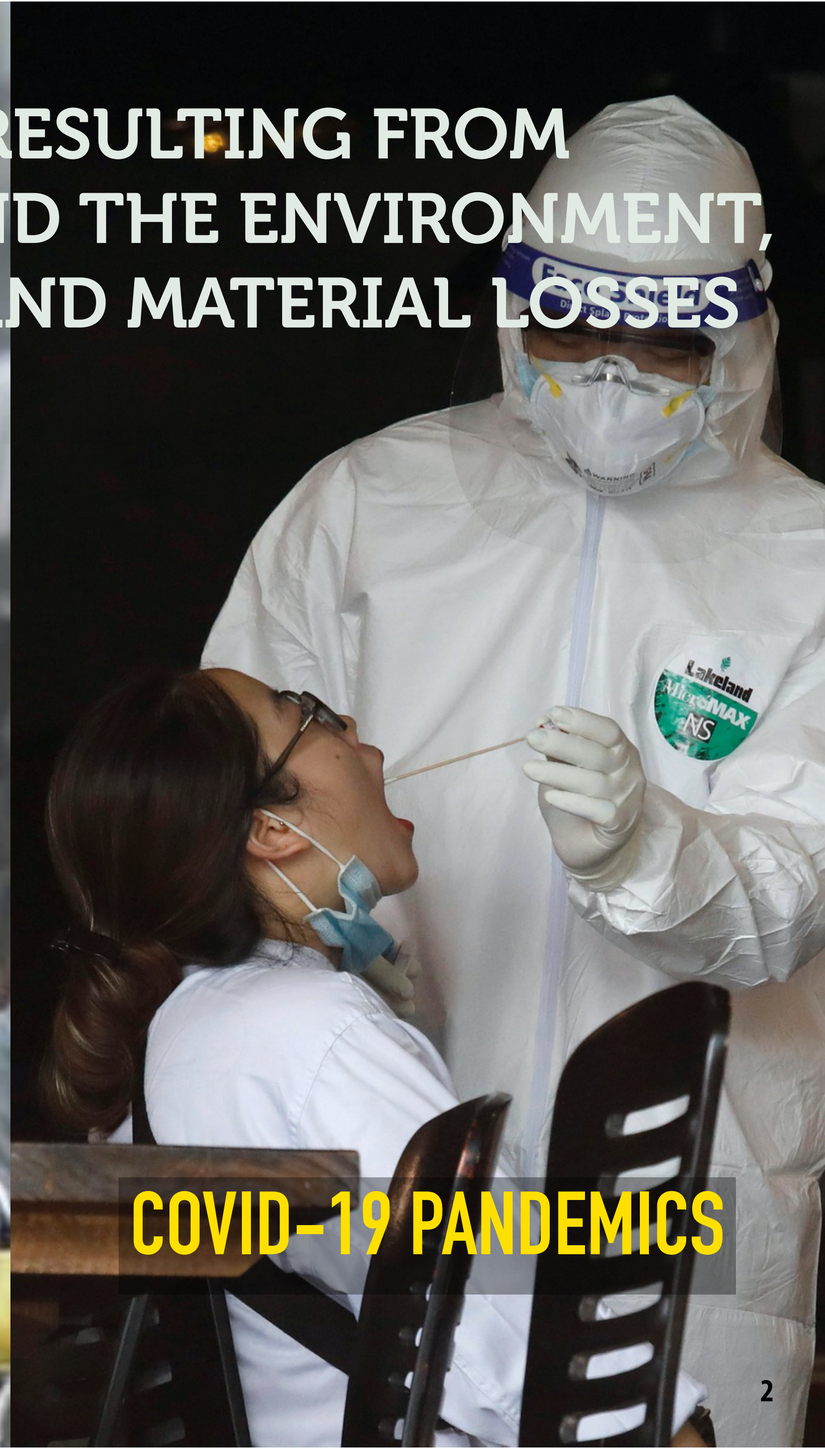
Institut de Recherche pour le Développement
FRANCE
French National Research Institute for Sustainable Development

VIETNAM HAS BEEN FACING, IN 2020, THREATS RESULTING FROM COMPLEX INTERACTIONS BETWEEN SOCIETY AND THE ENVIRONMENT, WHICH HAVE CAUSED CONSIDERABLE HUMAN AND MATERIAL LOSSES



NATURAL CATASTROPHES

ATMOSPHERIC POLLUTION



COVID-19 PANDEMICS

Institut de Recherche pour le Développement
FRANCE
French National Research Institute for Sustainable Development

THE SOCIAL AND POLITICAL DEMANDS ARE, OF COURSE, TO BETTER UNDERSTAND THESE THREATS SO AS TO ANTICIPATE AND MITIGATE THEM

SCIENTISTS (WORKING ON CLIMATE CHANGE, AIR POLLUTION, EPIDEMIOLOGY, ETC.) ARE THEREFORE PUT ON THE FRONTLINE

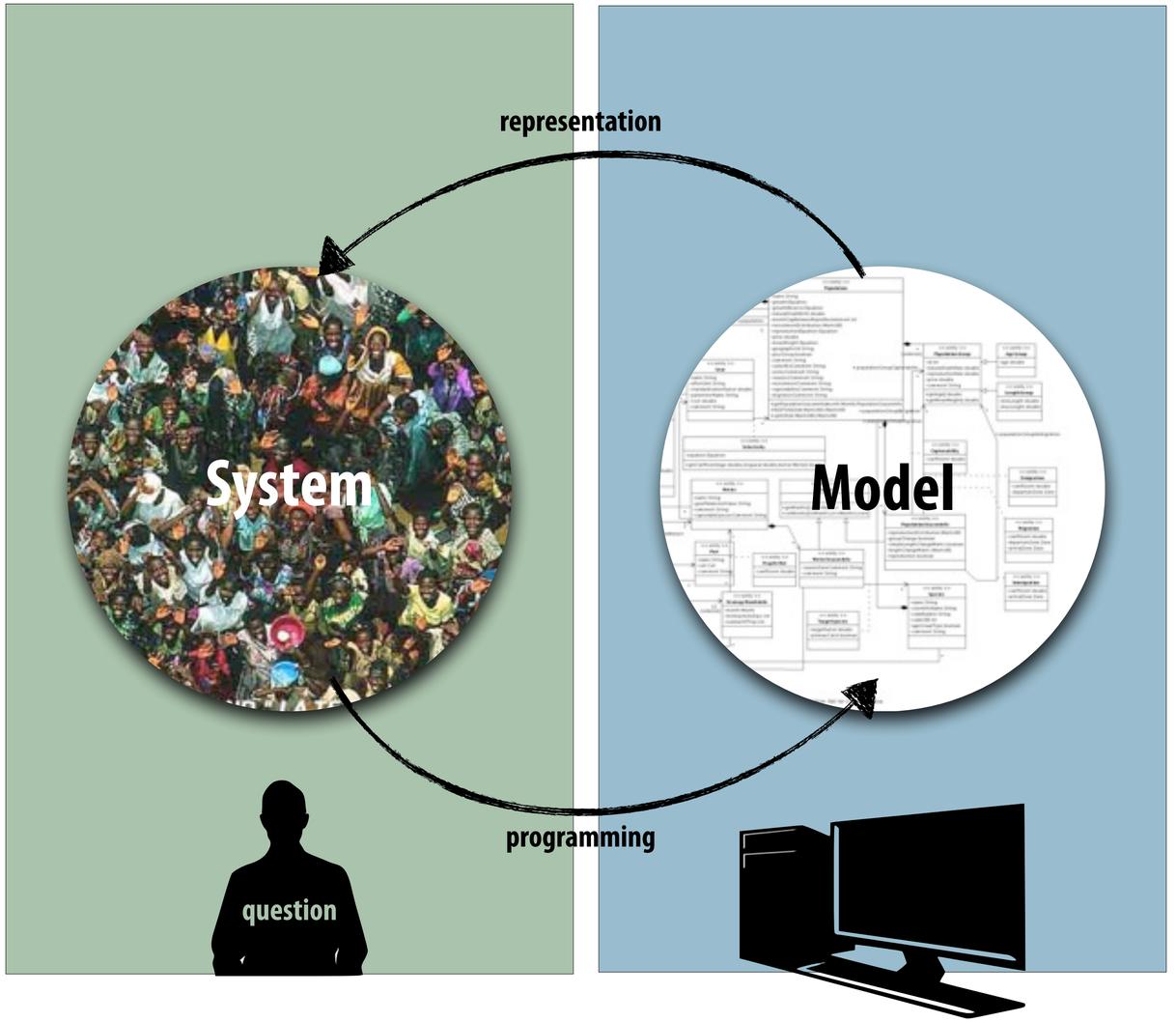
HOWEVER, THESE ARE PHENOMENA FOR WHICH NO EXPERIMENTAL APPROACH IS ETHICALLY AND PRACTICALLY CONCEIVABLE

THE ONLY POSSIBLE APPROACH IS TO WORK AND REASON ON MODELS

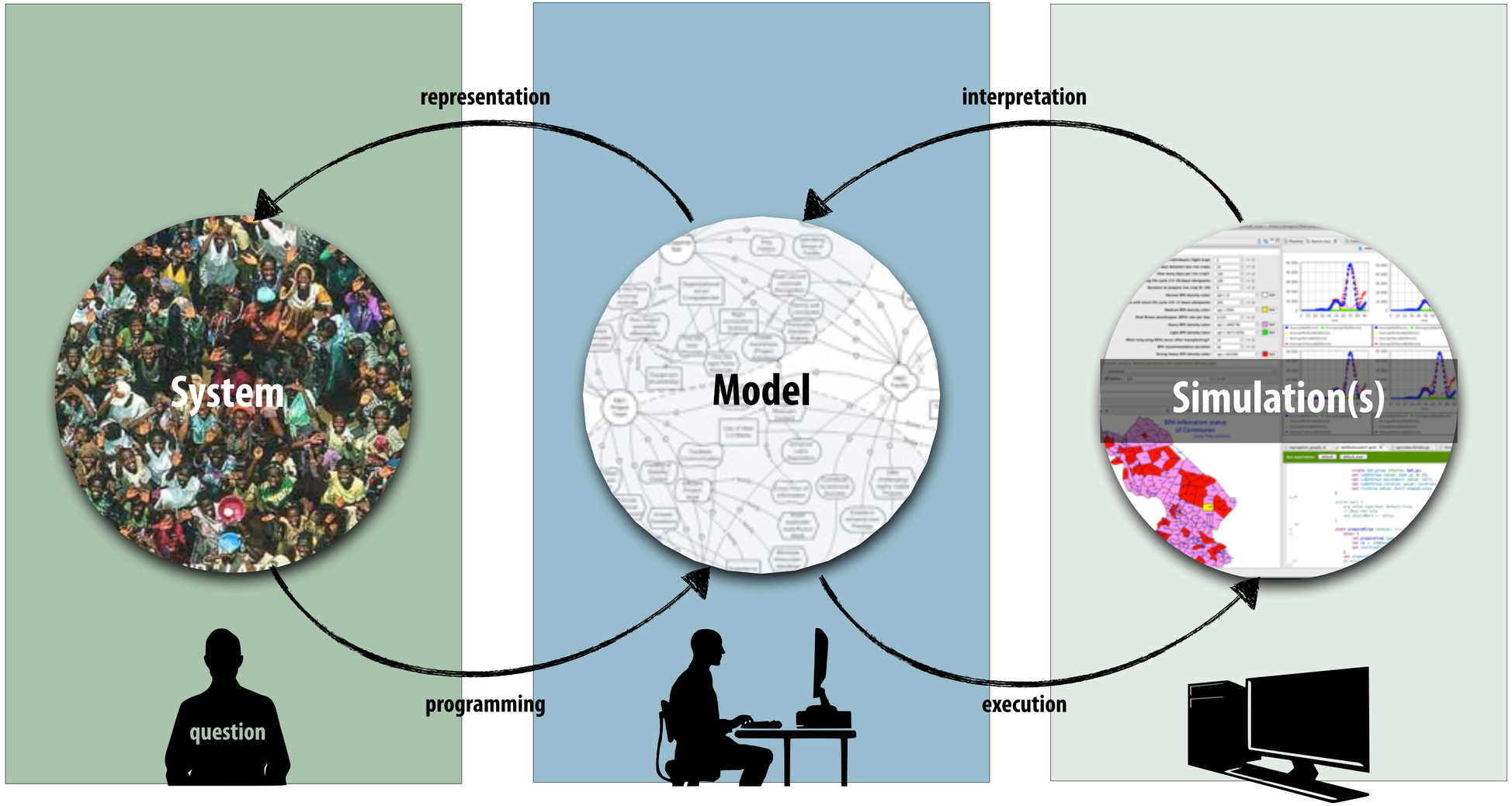
Institut de Recherche pour le Développement
FRANCE
French National Research Institute for Sustainable Development

A MODEL IS A SIMPLIFIED AND ABSTRACT REPRESENTATION OF A REFERENCE SYSTEM, WHICH HELPS TO ANSWER A QUESTION ABOUT THIS SYSTEM.

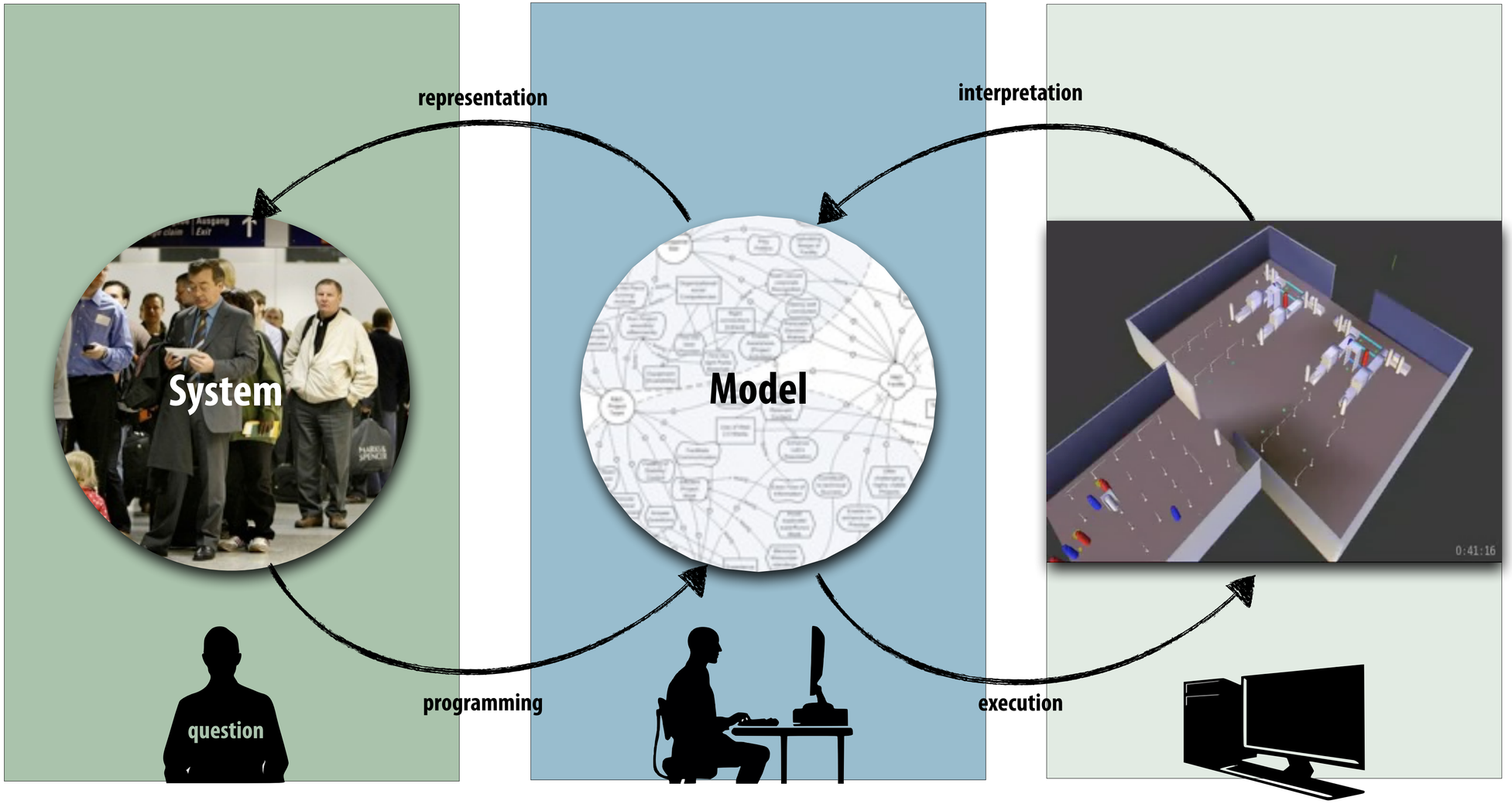
A COMPUTER MODEL USES A PROGRAM AS A REPRESENTATION



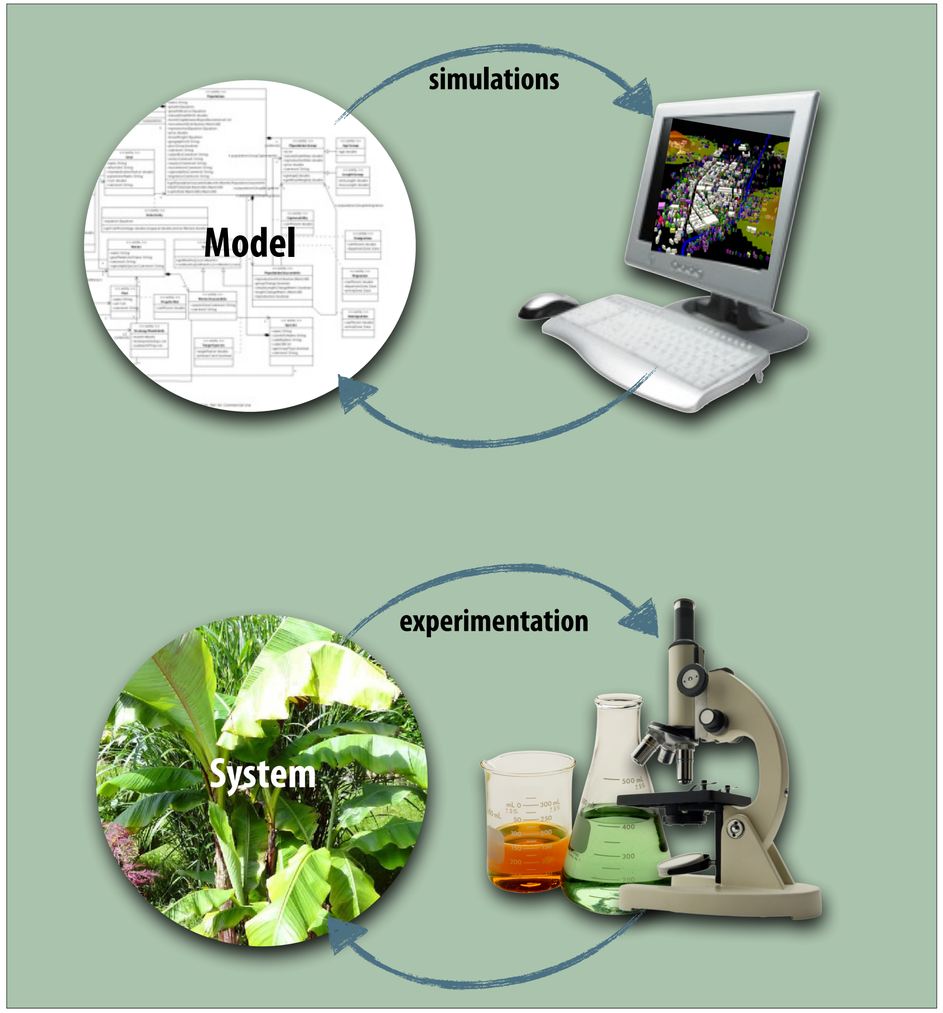
THE EXECUTION OF A COMPUTER MODEL IS CALLED A SIMULATION

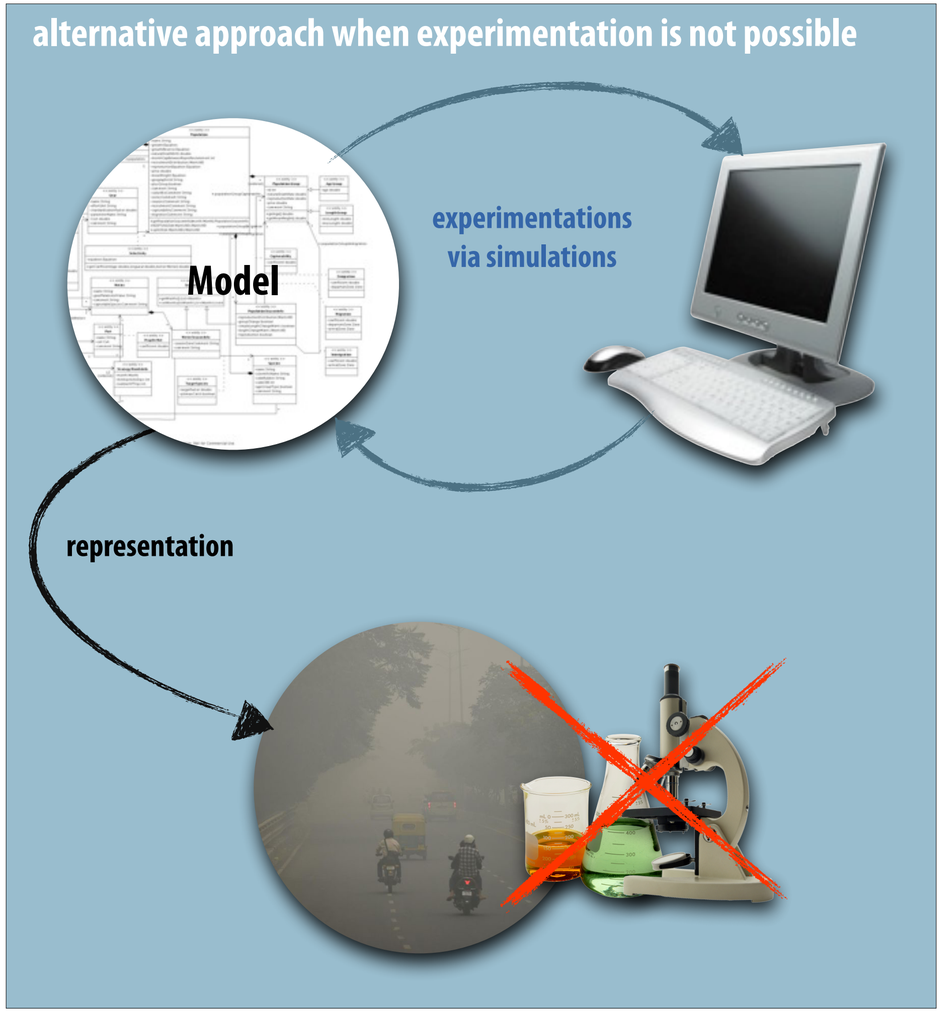


SIMULATIONS CAN SERVE MANY PURPOSES: VISUALIZATION, TRAINING, CONTROL, FORECASTING, DECISION SUPPORT, EVALUATION…

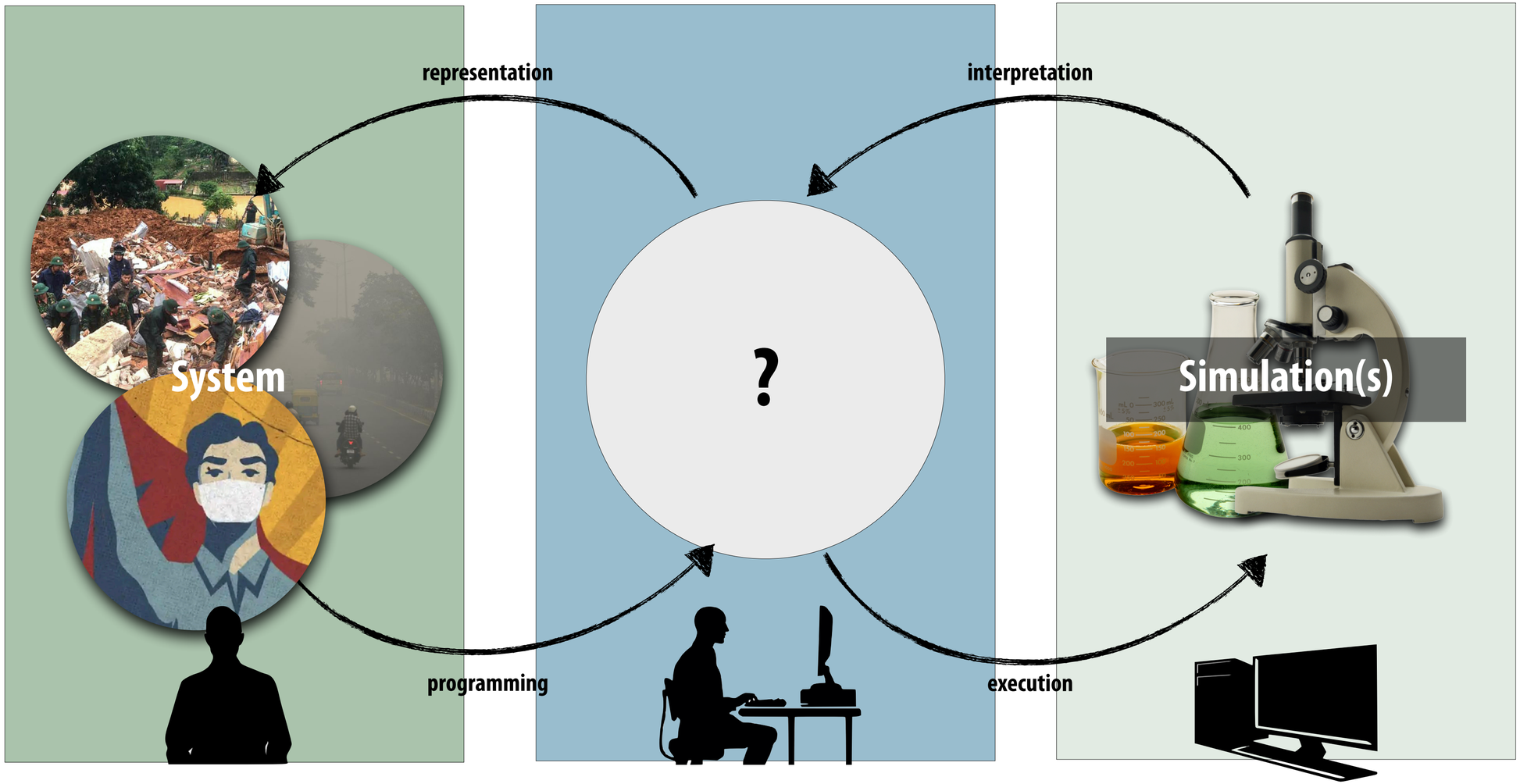


IN A SCIENTIFIC APPROACH, SIMULATIONS ARE TO THE MODEL WHAT EXPERIMENTATIONS WOULD BE TO THE SYSTEM





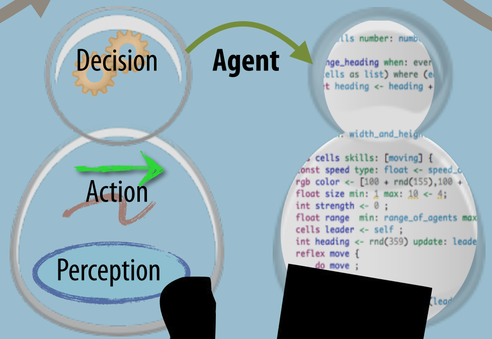
WHICH MODELING/SIMULATION METHODS ARE USED TO REPRESENT AND EXPERIMENT WITH THE COMPLEXITY OF THE INITIAL EXAMPLES?



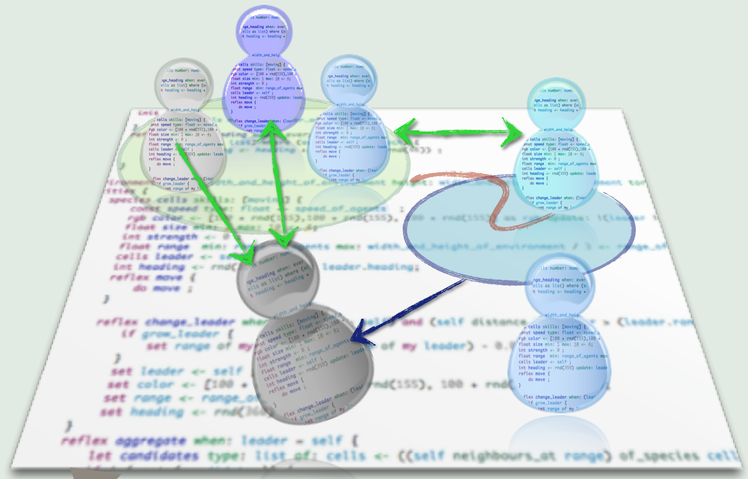
ONE IS AGENT-BASED MODELLING: BASED ON AN "INDIVIDUAL-CENTERED" REPRESENTATION THAT ALLOWS TO RECONSTRUCT AND SIMULATE "VIRTUAL WORLDS" ON A COMPUTER.

Real system actors

… represented by computer programs called agents…

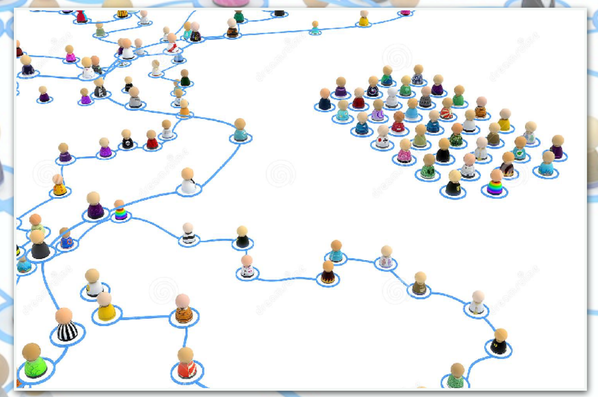


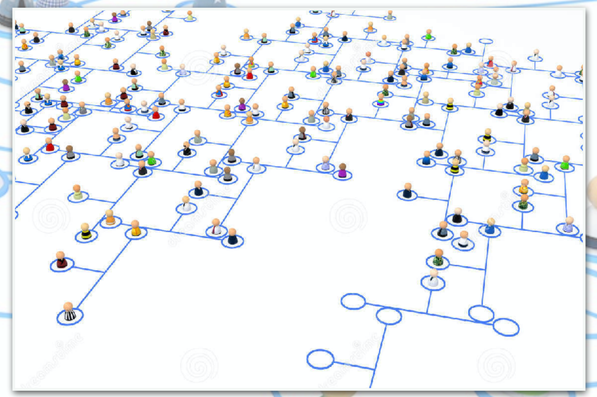
… gathered together and run on a computer

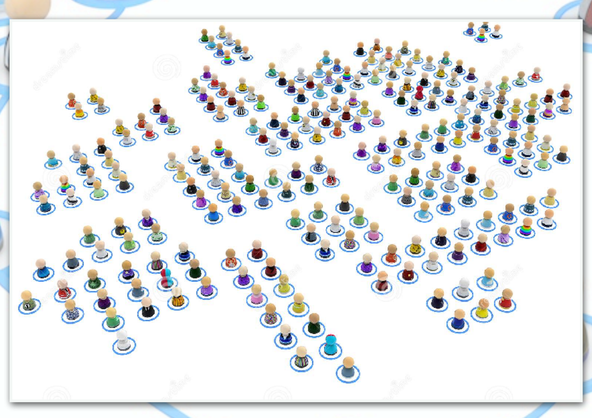


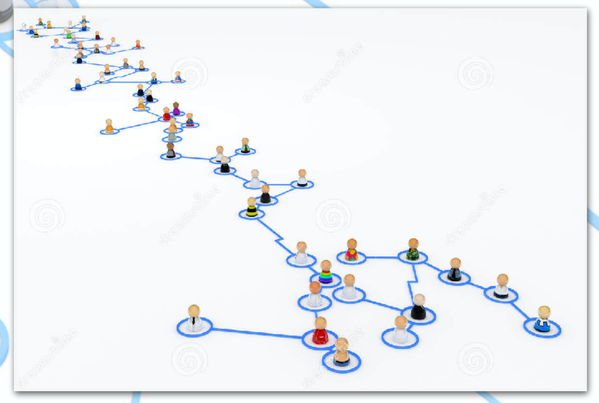
Institut de Recherche pour le Développement
FRANCE
French National Research Institute for Sustainable Development

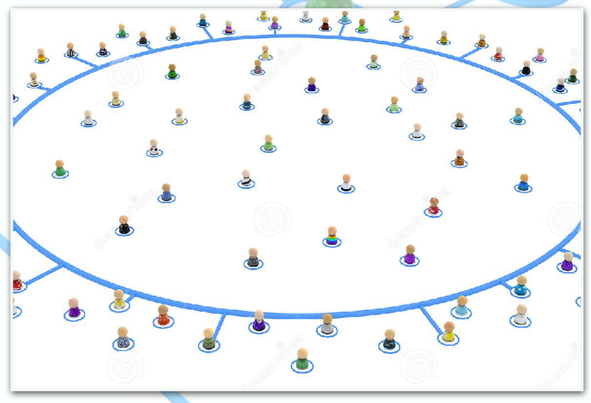
AGENT-BASED SIMULATIONS ARE MINIATURE LABORATORIES THAT SUPPORT EXPERIMENTATION ON THESE "VIRTUAL WORLDS" AND TAKE ADVANTAGE OF THE POWER OF COMPUTERS TO EXPLORE AS MANY SCENARIOS AS NECESSARY.







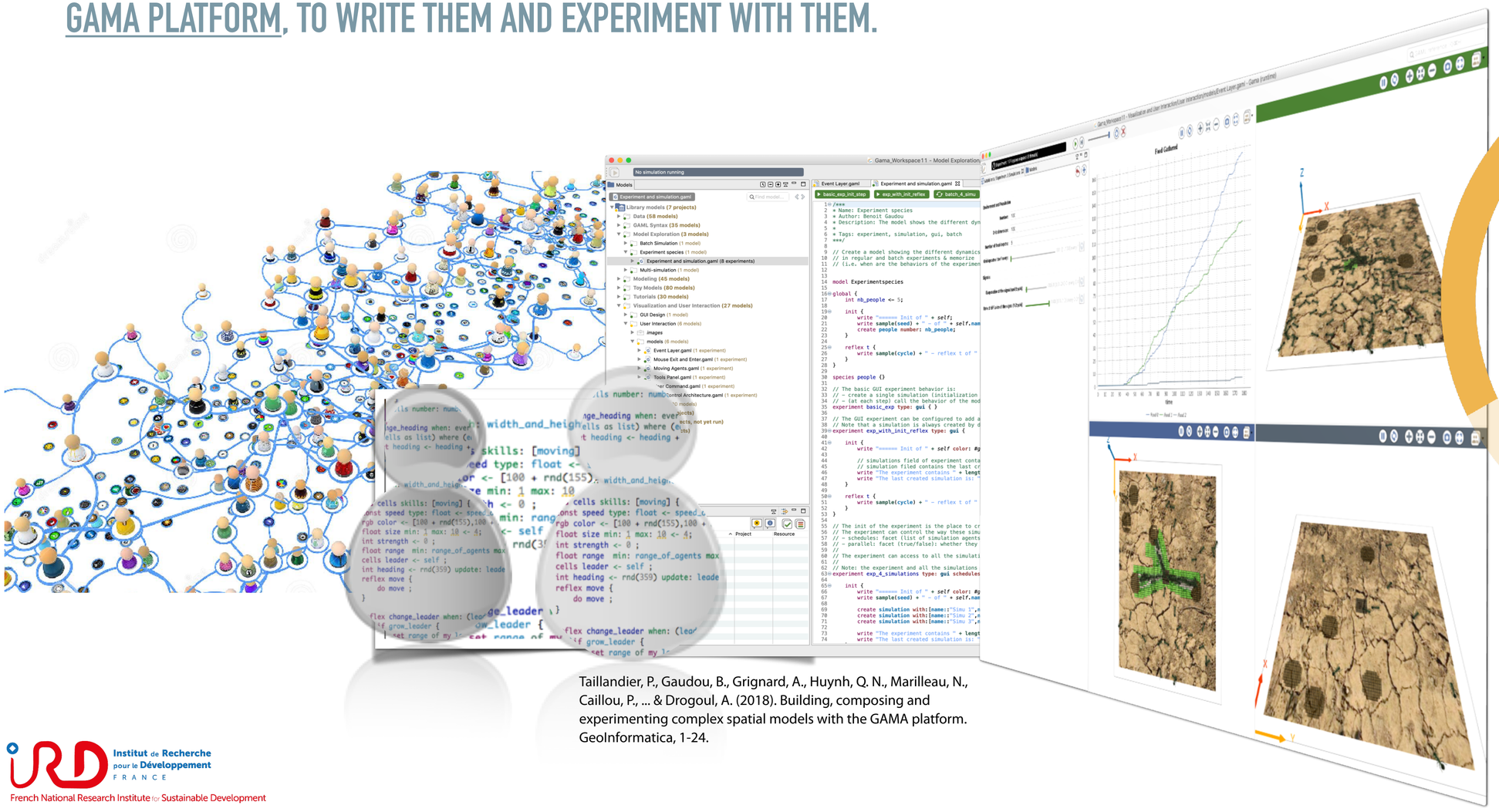




French National Research Institute for Sustainable Development

ABM ARE VERY DETAILED MODELS, AND THEREFORE COMPLEX PIECES OF SOFTWARE, WHICH USUALLY RELY ON COMPLEX DATASETS (GIS DATA, QUALITATIVE AND QUANTITATIVE SURVEYS, DEMOGRAPHICAL AND STATISTICAL DATA…)

IT IS NECESSARY TO USE MODELLING PLATFORMS, SUCH AS THE GAMA PLATFORM, TO WRITE THEM AND EXPERIMENT WITH THEM.



Taillandier, P., Gaudou, B., Grignard, A., Huynh, Q. N., Marilleau, N., Caillou, P., ... & Drogoul, A. (2018). Building, composing and experimenting complex spatial models with the GAMA platform. GeoInformatica, 1-24.

http://gama-platform.org

Institut de Recherche pour le Développement FRANCE
French National Research Institute for Sustainable Development

GAMA: A FREE AND OPEN-SOURCE MODELLING AND SIMULATION PLATFORM CREATED IN VIETNAM IN 2007, DEVELOPED BY AN INTERNATIONAL CONSORTIUM

France
IRD (PI), CNRS, INRAE, University of Toulouse 1 Capitole, University of Paris Sud, University of Rouen

A. Drogoul, B. Gaudou, P. Taillandier, N. Marilleau, K. Chapuis, P. Caillou, A. Brugière, Nguyen Huu Tri, Y. Sklab, JD Zucker

Netherlands
Delft University of Technology

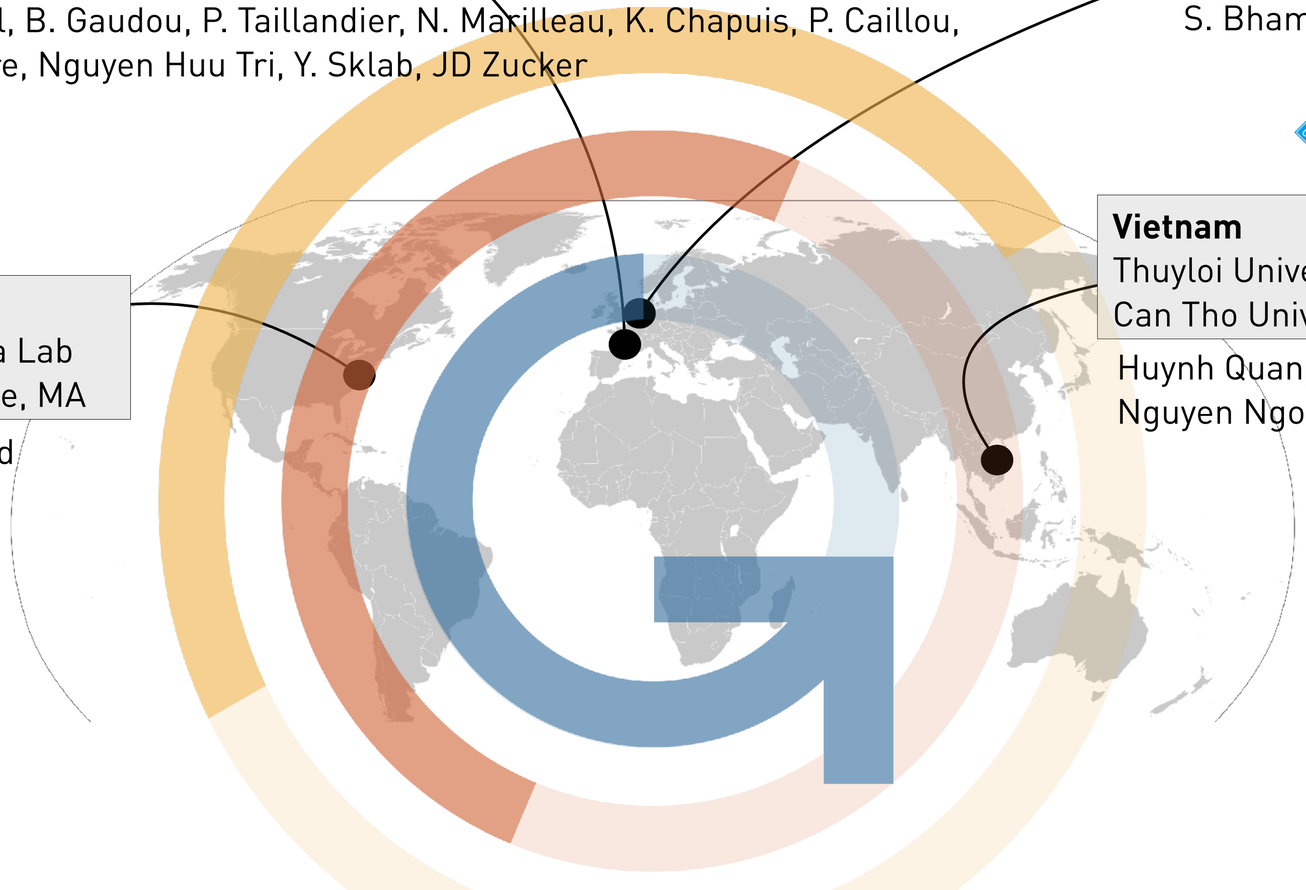
S. Bhamidipati

USA
MIT Media Lab
Cambridge, MA

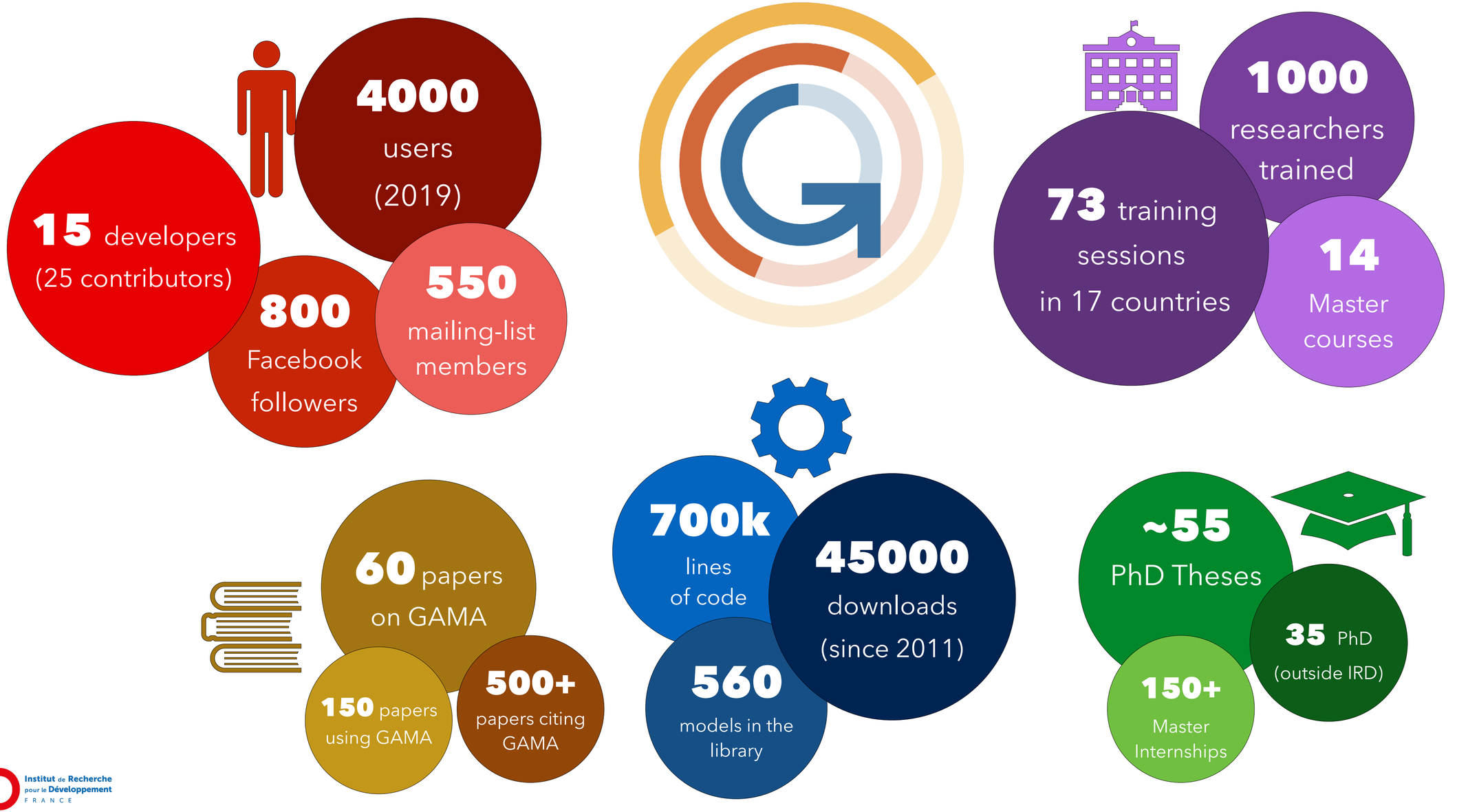
A. Grignard

Vietnam
Thuyloi University,
Can Tho University

Huynh Quang Nghi,
Nguyen Ngoc Doanh

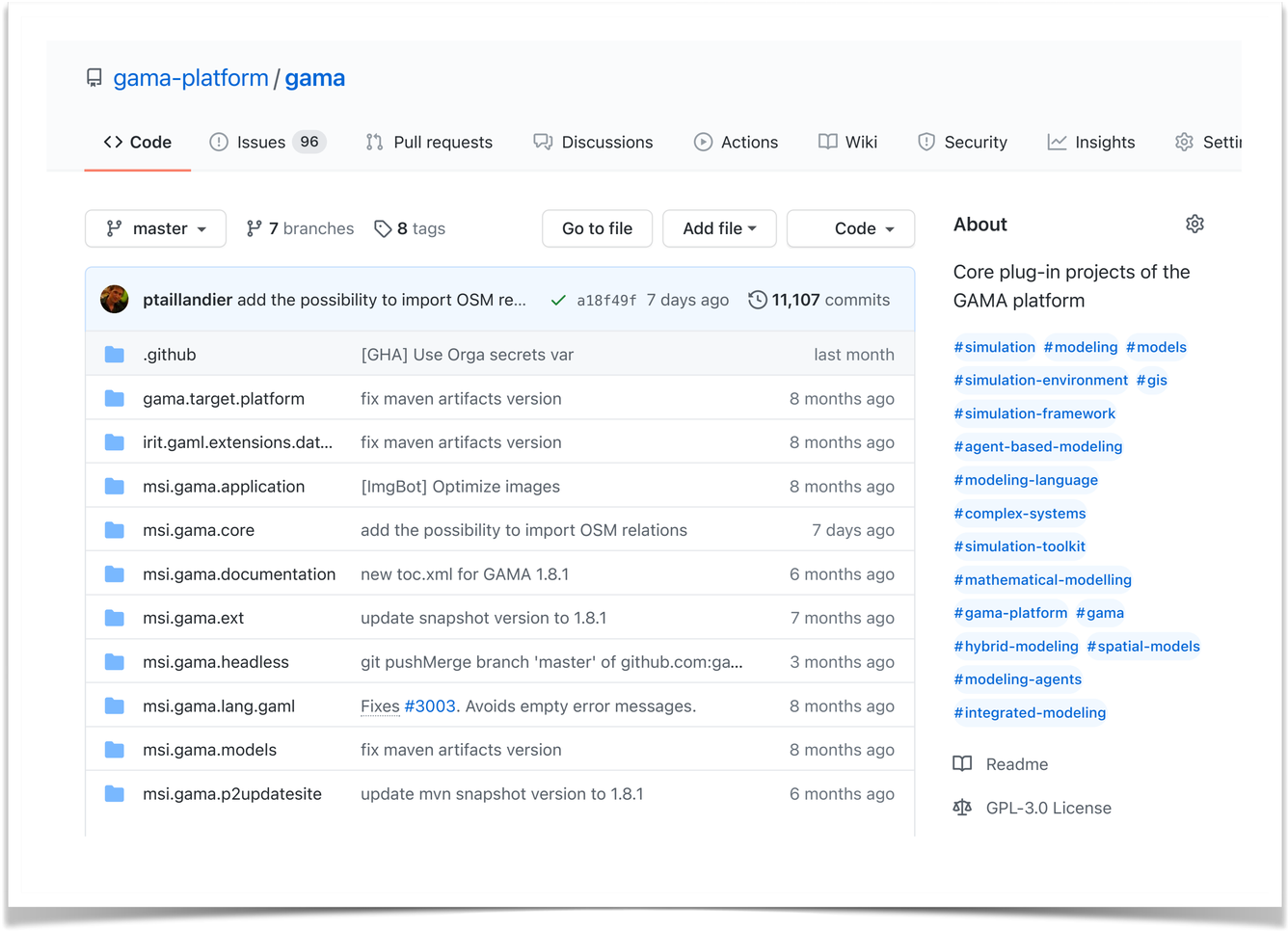


GAMA, IN A FEW FIGURES…



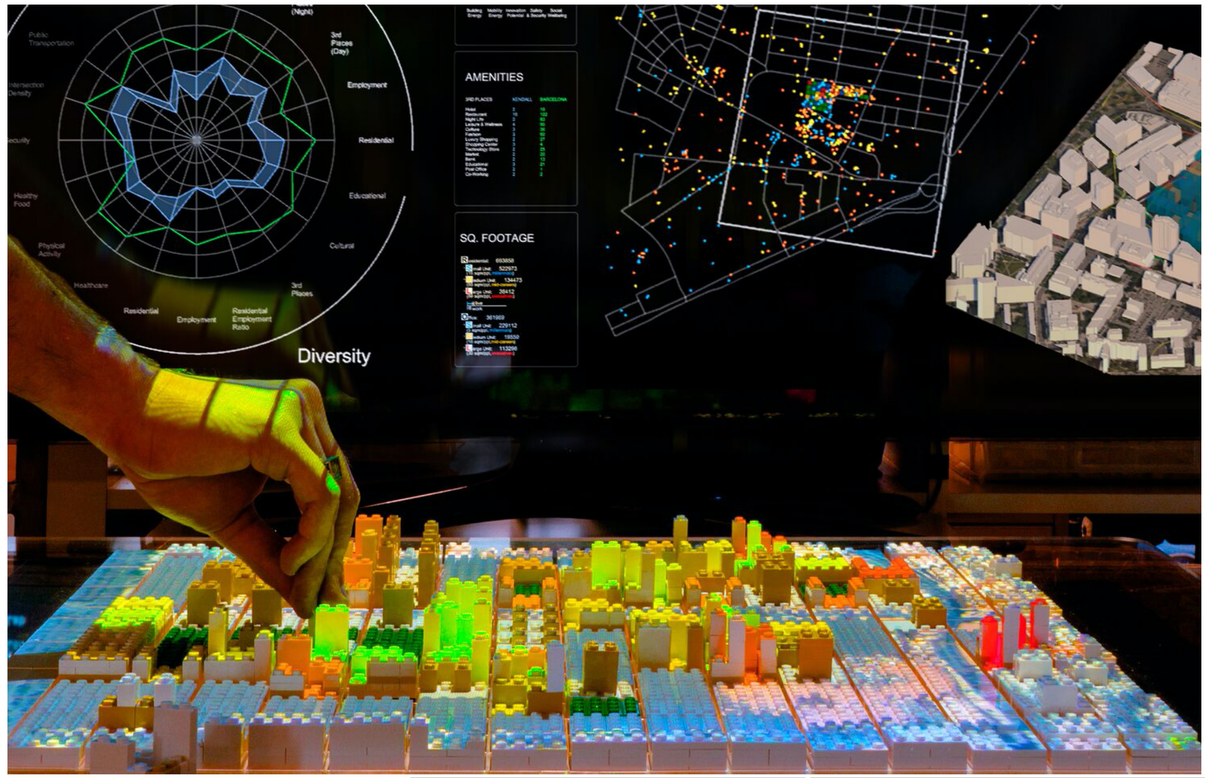
GAMA IS OPEN-SOURCED SINCE 2008

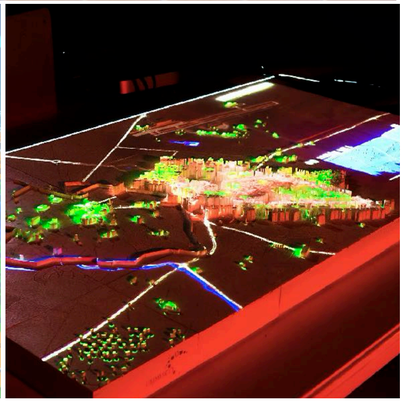
- Well-known advantages of open-source (robustness, maintenance, evolution…) in software development
- Possibility for an institute like IRD, present in 37 countries, to support distributed cooperative work
- Allows researchers to open the "black box" of simulations: random number generators, scheduling algorithms, primitives, etc. => which can have a strong influence on the outcomes of simulations



BUT OPEN-SOURCING SIMULATION PLATFORMS IS NOT ENOUGH: MODELS NEED TO BE OPEN AS WELL

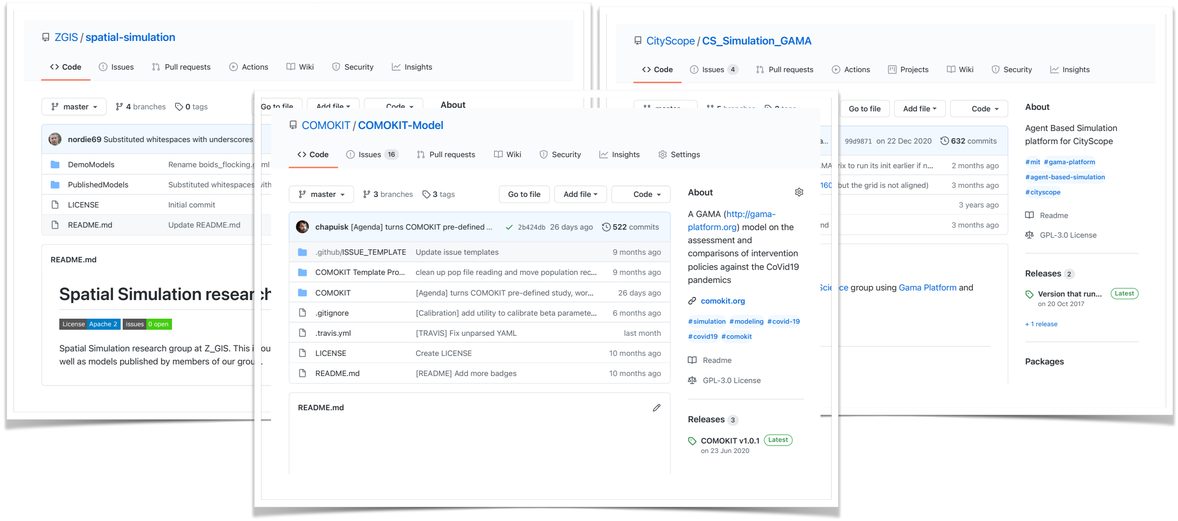
- As their influence on policy-making becomes widespread, the design of very detailed and realistic models like agent-based models raises different challenges, among which :
  - the necessity to be as transparent and manipulable as possible in order to support multidisciplinary contributions;
  - the necessity to remain understandable by stakeholders invited to participate in the design and assessment of strategies and policies;
  - the necessity to be trustable by users: let people understand that there is no magic behind a model, and be able to unveil untold hypotheses

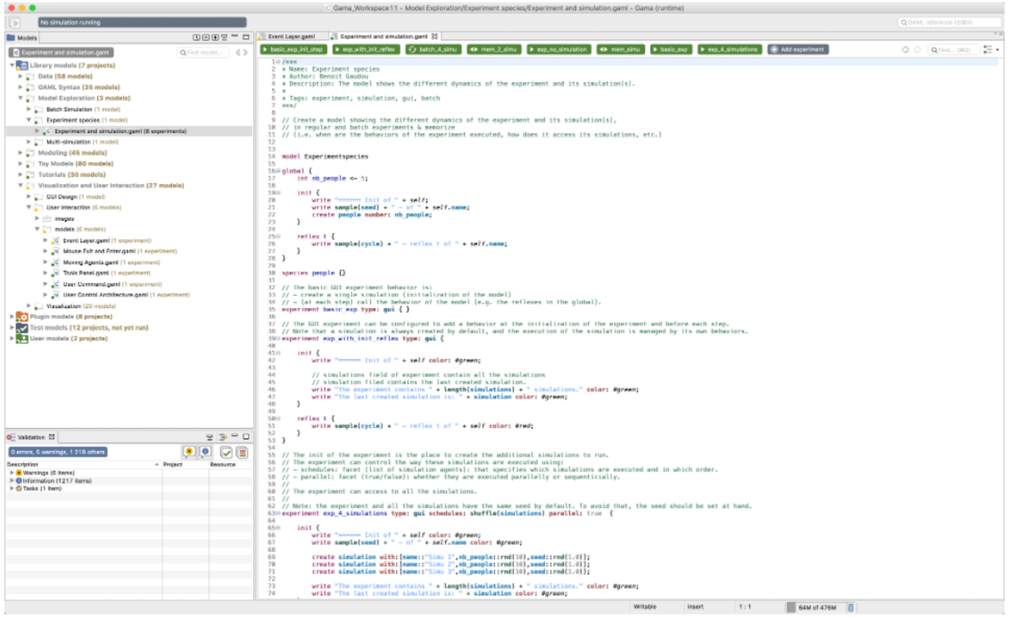




SINCE THE BEGINNING, AUTHORS OF MODELS IN GAMA ARE STRONGLY ENCOURAGED TO MAKE THEIR MODELS OPEN-SOURCE

- Also to rely on open data and publish the outcomes of experiments as open data (cf. the initiative to open a Dataverse repository/warehouse at IRD)
- A radical (and discussed) choice in GAMA is to always let the source code available for viewing and editing by "users", even in "demo" mode.



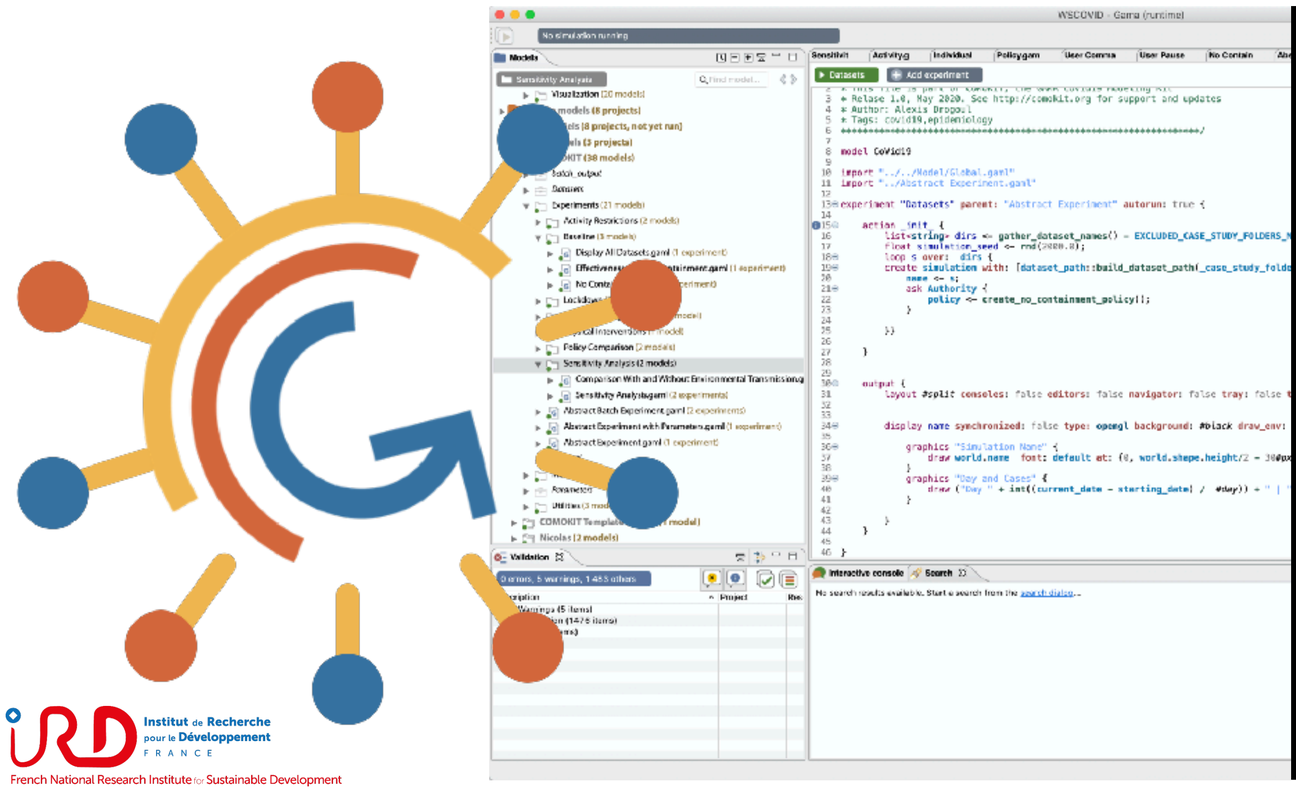


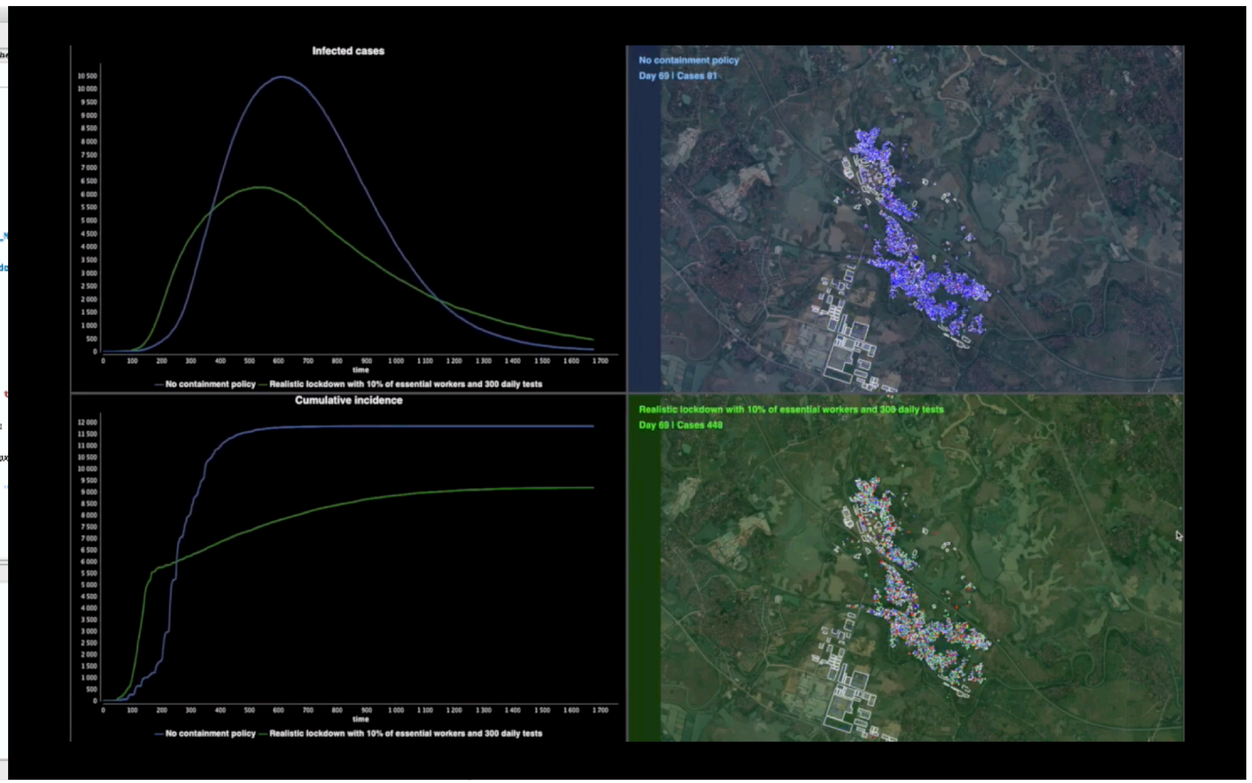
EXAMPLE OF THE COMOKIT MODEL

- Initial goal: support Vietnamese authorities (esp. the Rapid Response Team of the National Steering Committee) in their fight against COVID-19
- A generic model to assess and compare mitigation policies and interventions at the level targeted by these authorities (i.e. commune, ~10.000 inhabitants)
  - Developed since March 2020 by IRD, Thuyloi University, Can Tho University, NIHE (Pasteur Institute), Hong Kong University, Oxford University (OUCRU), Toulouse 1 University, INRAE
- Quickly open-sourced to allow people to (1) apply it to other case studies; (2) extend it

HYPOTHESES BEHIND COMOKIT

- Classical models lack the level of individual detail needed for applications to small populations. At the commune scale, spatial aspects and individual and social heterogeneities are key factors to take into account.
- These individual properties can be represented by agent-based models where people and households are explicitly represented and serve as a basis for describing the dynamics of the epidemic.
- Hourly time scale - No transmission/infection during transportation

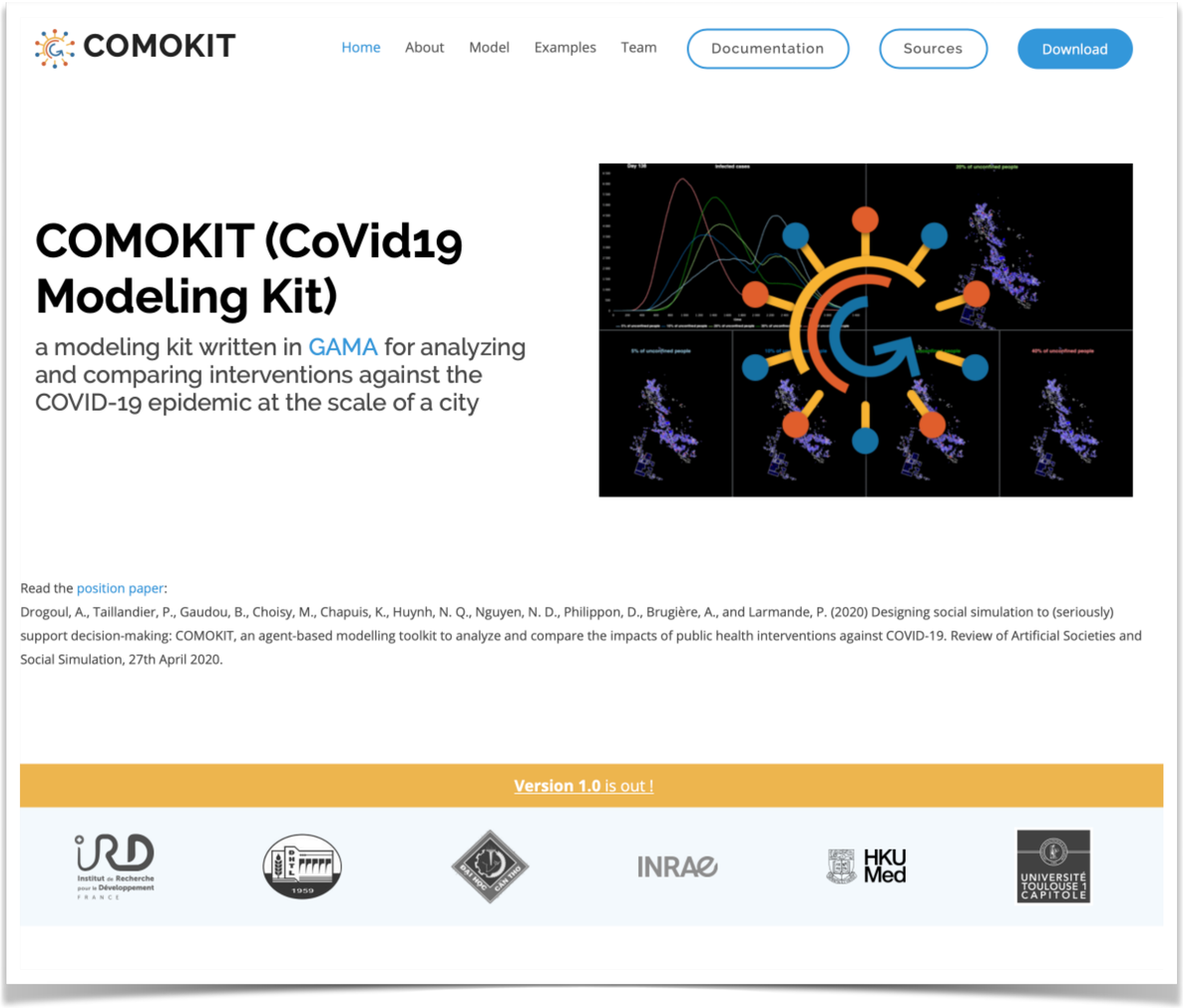




COMOKIT 1.0 IS AVAILABLE AT HTTP://COMOKIT.ORG.

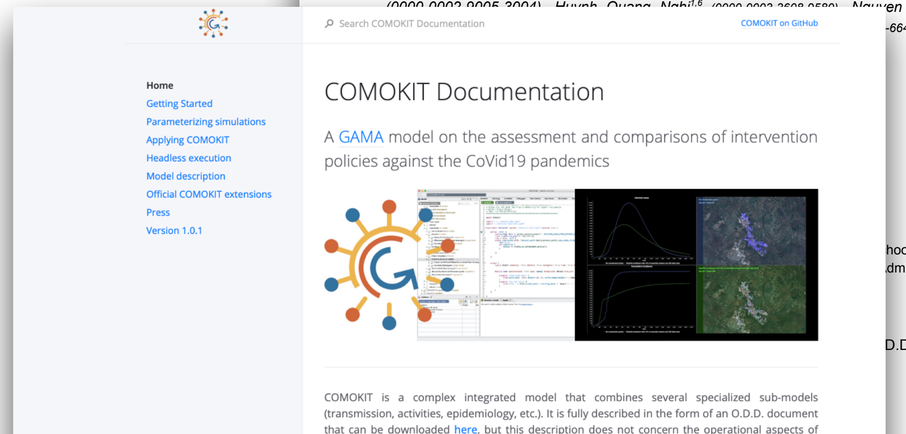
- Already applied to different case studies, among which:
  - Refugees camps in Turkey
  - Nice city center
  - Netherlands

- Open to contributions and extensions, among which:
  - Upscaling (population size, case study extent, …)
  - Addition of realistic urban fluxes (transportation, …)
  - Addition of more detailed activities
  - Extension to other infectious diseases
  - Coupling with existing models (urban planning, …)



THE EFFORT OF OPEN SOURCING SUCH MODELS IS HOWEVER VERY IMPORTANT, REQUIRES LOTS OF RESOURCES, AND GOES BEYOND "SIMPLY" OPENING THE CODE

- To be able to effectively open and share the model, we also had to
  - Make several deposits at https://www.comses.net/ for keeping trace/history of the running versions
  - Write and maintain an ODD documentation (http://jasss.soc.surrey.ac.uk/23/2/7.html) to provide a readable description of the logic of the model and the questions it answers
  - Build and maintain a dedicated website for documenting how to extend it

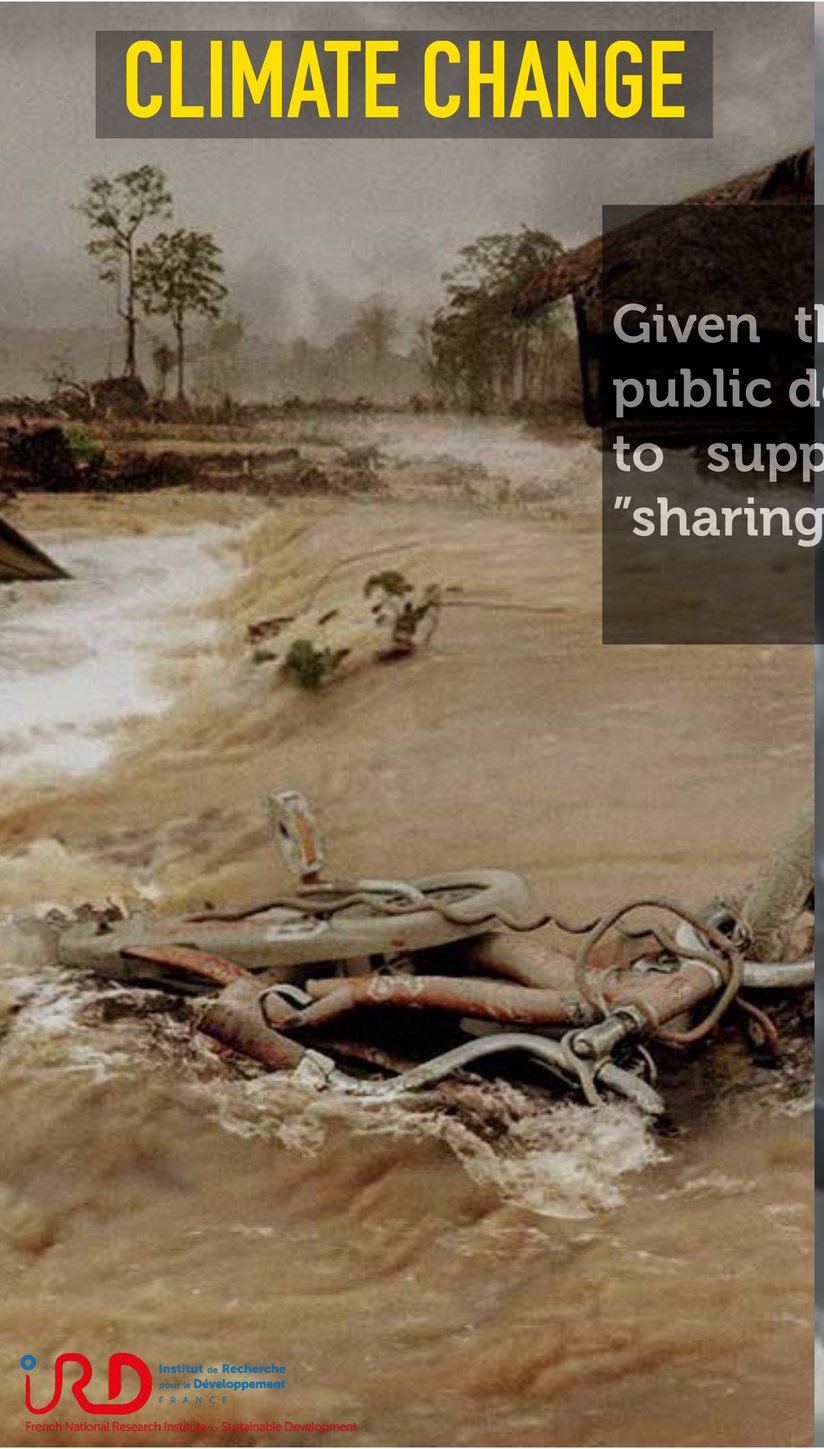


CLIMATE CHANGE

AIR POLLUTION

FUTURE PANDEMICS

Given the growing importance of models in the public debates, a public initiative would be welcome to support researchers in better "opening" and "sharing" them to deciders and the general public



Institut de Recherche pour le Développement
FRANCE
French National Research Institute for Sustainable Development