AI FORA
artificial intelligence for assessment

Introduction to gamification & simulation as methods of the high-level strategy

Nigel Gilbert and Martha Bicket

University of Surrey

Guildford

UK

3 May 2023, Montserrat

http://cress.soc.surrey.ac.uk

cress
Centre for Research in **Social Simulation**

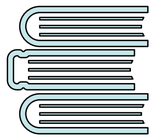
AI FORA
artificial intelligence for assessment

The steps in the modelling process

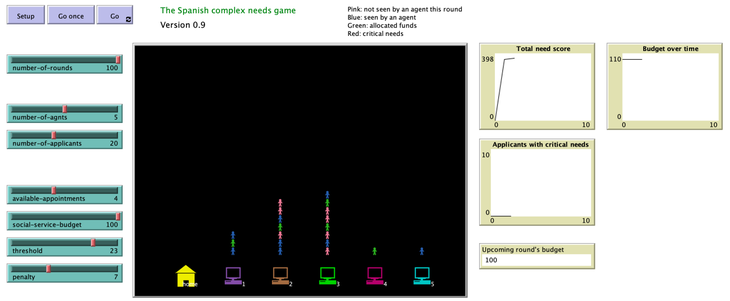
1. A workshop will be held to map out the overall existing case study system as a flow chart.
2. Rules for a game to be played with stakeholders will be written.
3. An Agent-Based Model (ABM) that models the current social assessment system, including the initial rule set and exemplar agent attributes, will be written.
4. The initial rule set will be checked and refined by running the ABM to become the 'current rule set'.
5. At a gamification workshop with the stakeholders, the current rule set will be gradually adapted by the stakeholders to become a more desirable assessment algorithm.
6. The 'better rule set' will be extracted using the records from the game play.
7. The ABM will be modified to incorporate the better rule set.
8. The ABM will be used to generate a data set that has a case ('row') for every permutation of the applicants' attributes. The ABM will be run using the better rule set for each case (possibly multiple times to deal with stochasticity) to see what the social assessment is for that combination of attributes. This will yield a dataset of 'inputs' (the attributes) and 'outputs' (the assessment).
9. Using this dataset, a neural network (NN) or other ML system will be trained to match the dataset. This NN is the 'better AI algorithm' for the case study.
10. A final stakeholder workshop will be held at which the better AI algorithm will be introduced and if possible tested against representative empirical data about applicants for social assessment (or if empirical data are not available, tested on the basis of plausibility).

The AI FORA process illustrated

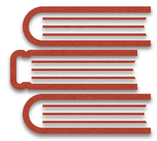
Workshop 1



Ruleset 1



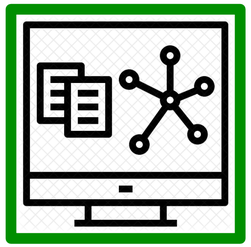
ABM v.1



Ruleset 2

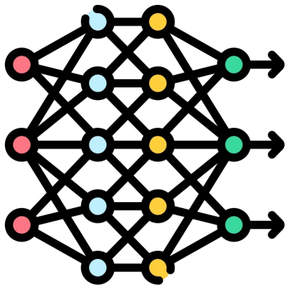
Workshop 2

Ruleset 3



ABM v. 2

Dataset



Neural network

Initial work

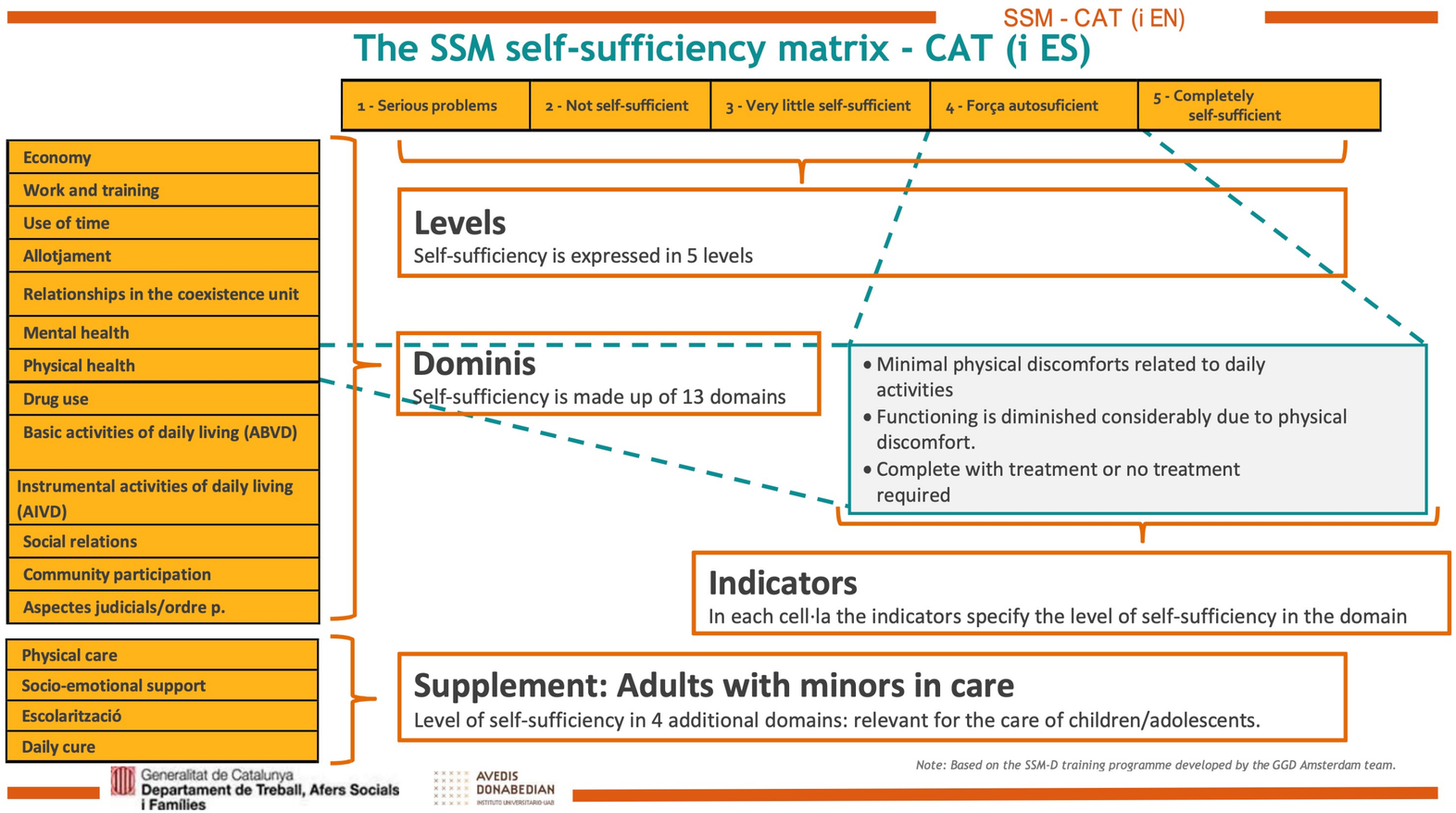
AI FORA
artificial intelligence for assessment

Meeting of stakeholders and experts to agree on the domain and its scope

Meeting of stakeholders for the Ukraine case study (November 2022)

Initial work

- Documentary review
- Output: a general idea of the use of AI in this social service delivery domain



Spanish case study: applicant categorisation proposed for Catalunya

Initial ideas for the game

- Who do the players represent?
  - Clients or Social Service personnel
- What are the players' characteristics ('attributes')?
  - e..g. age, gender, income, job history,...
- What is the scene?
  - 'Stations' - agents or clients move between stations during the game
  - Stations: service desks, home, town hall, etc.
- What do the players do?
  - Follow a prescribed set of rules, one at a time, again and again
  - Vote to change the rules if they think the rules are unfair
- How does it end?
  - When the money runs out, or whatever suits the game

Initial Rule Set: Spanish case study example

- Social Service Agents
  - The central characters in the game are social service agents working in a municipality who are seeking to allocate (limited) social service resources to deserving applicants, many of whom have multiple, complex needs.
  - The agents' aim is to allocate social service resources to applicants in order to maximise the sum of applicants' wellbeing.

Initial Rule Set

- Applicants have:
  - Overall need score (highest scoring individual is deemed by the algorithm to be the most in need)
  - Household income
  - Number of dependents
  - Accommodation
  - Work and training
  - Mental health
  - Physical health

Initial Rule Set

- Resources
  - Total social service budget
- Locations
  - Agent home (end of round)
  - Social service desk
  - Office meeting
    - where social service agents can see the scoring algorithm and vote to change it and/or other elements of the game (e.g. the assessment process and resource allocation).

Game Play

- Agents deal with a limited number of applicants per round
- Before the round
  - Applicants are randomly allocated appointments with agents.
- During the round
  - Agents review and score applicants based on the scoring ruleset and applicants' attributes.
- At the end of the round
  - The social service budget is distributed to successful applicants in order of severity
    - The highest scoring applicant gets full payment, then the second highest applicant and so on until the budget for that round is used up
  - The applicants' needs are updated depending on their existing needs and whether they received support or not.
  - Some of the applicants' situations may then change at random. Some will improve if they received support. If they didn't receive anything, their situation may deteriorate further.

Scoring

For all attributes except household income and number of dependents:

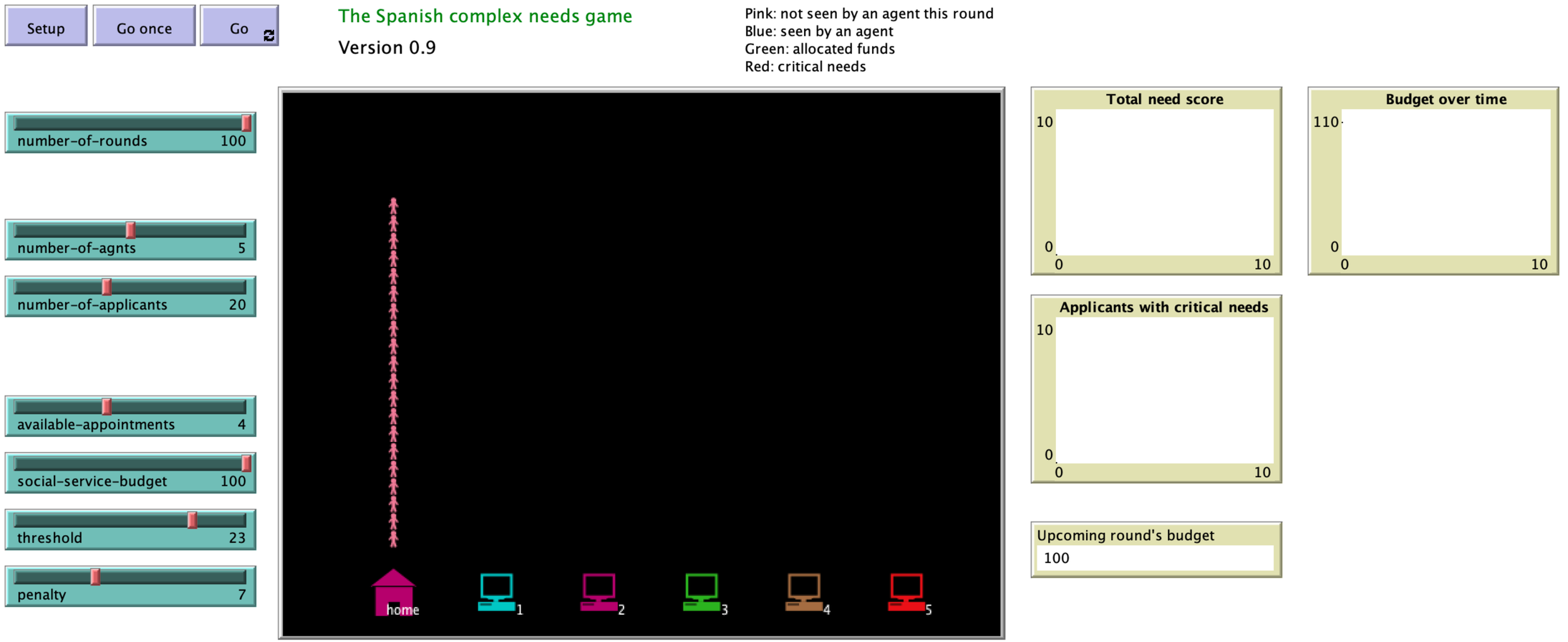
- One ‘need point’ is allocated according to which category the applicant’s current situation falls into, e.g.:

Need points	5	4	3	2	1
Applicant’s situation	Serious problems	Not self-sufficient	Minimally self-sufficient	Sufficiently self-sufficient	Completely self-sufficient

For household income and number of dependents:

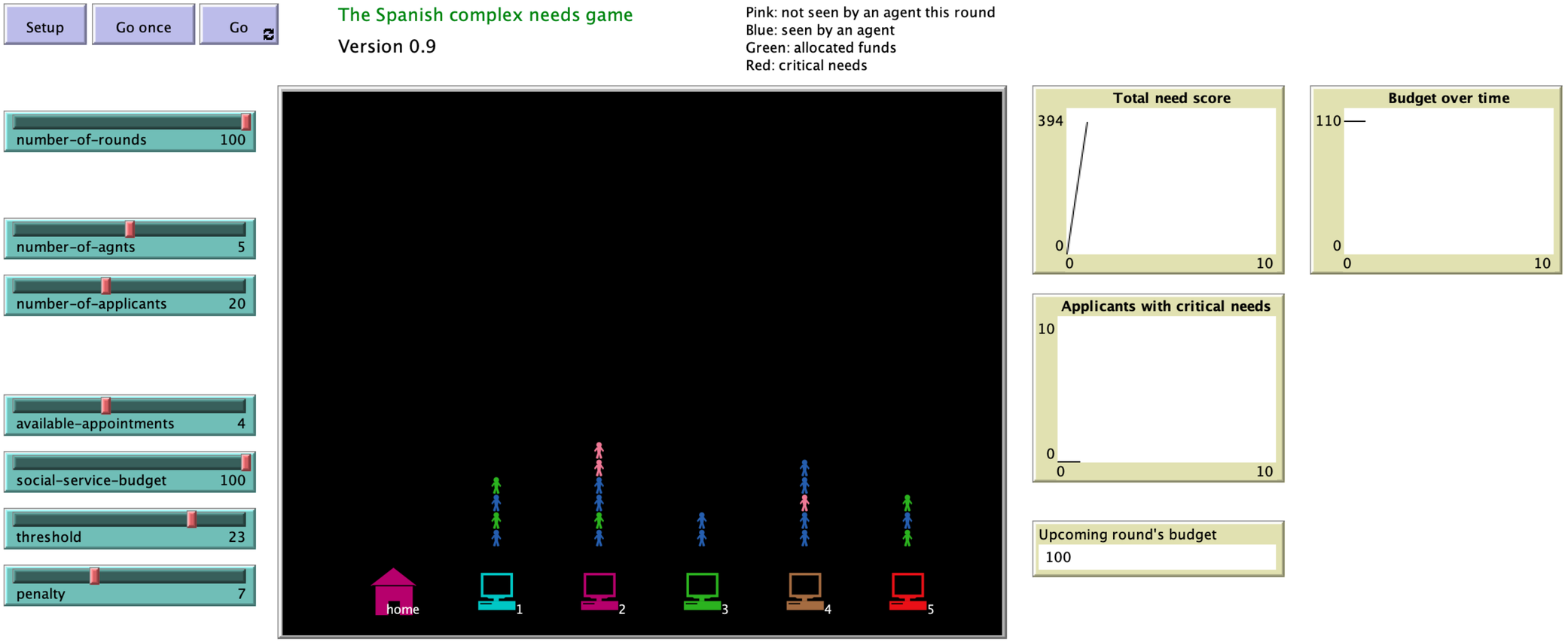
- applicants are ranked against each other and given a need point based on their position relative to other applicants in that round.

Agent-based model



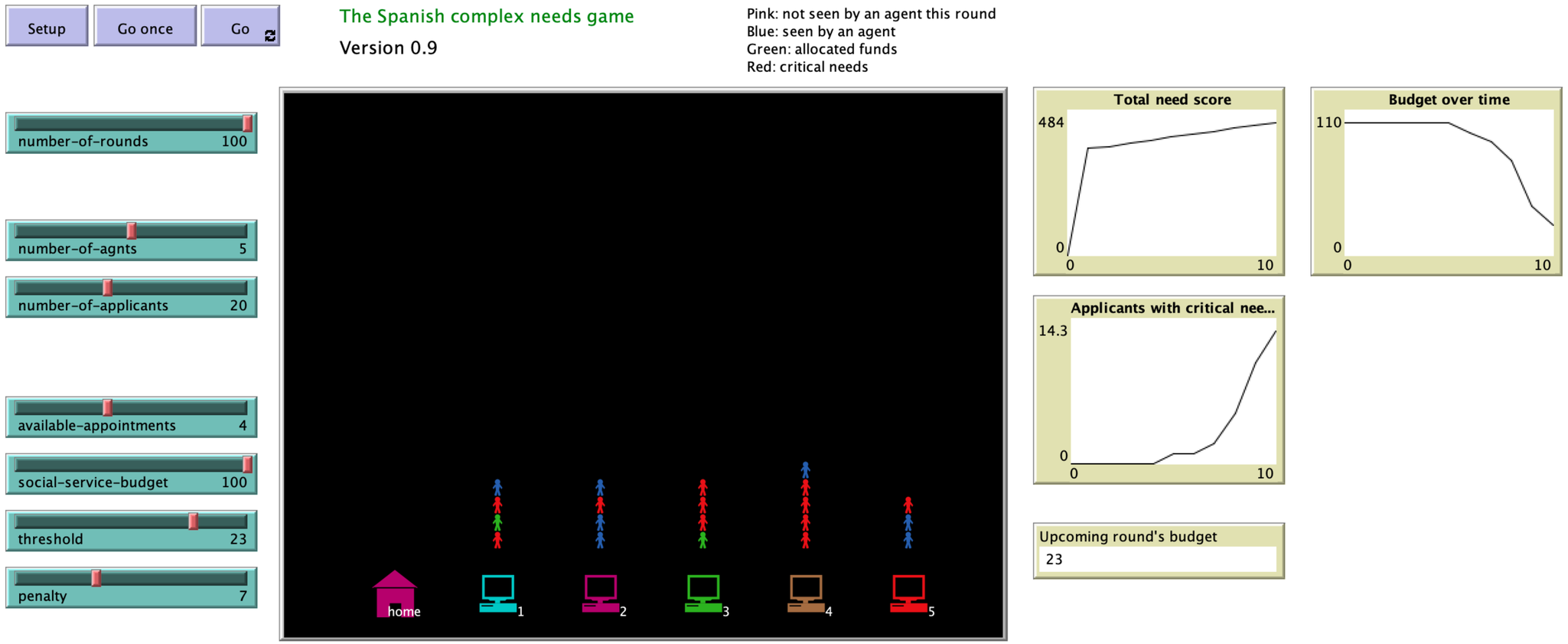
At the *start* of round 1

Agent-based model



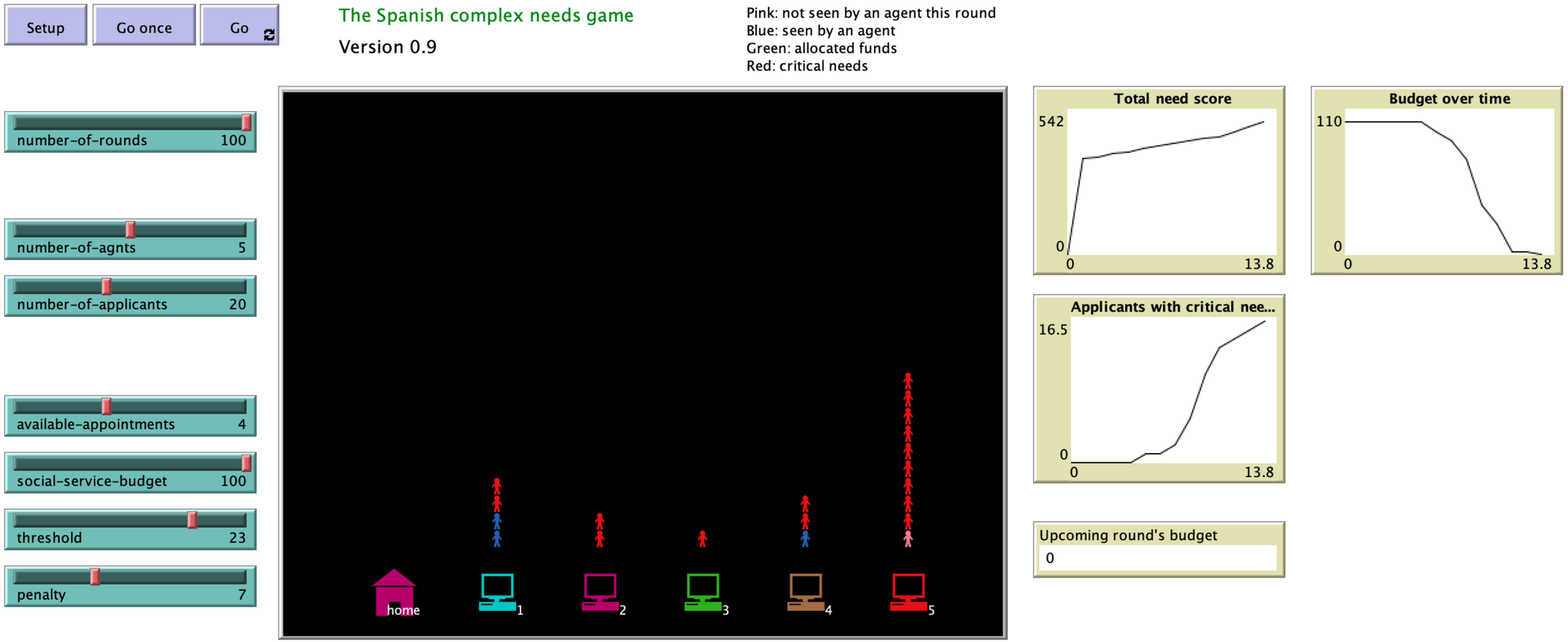
At the *end* of round 1

Agent-based model



At the end of round 10

Agent-based model



At the end of round 13, the end of the run

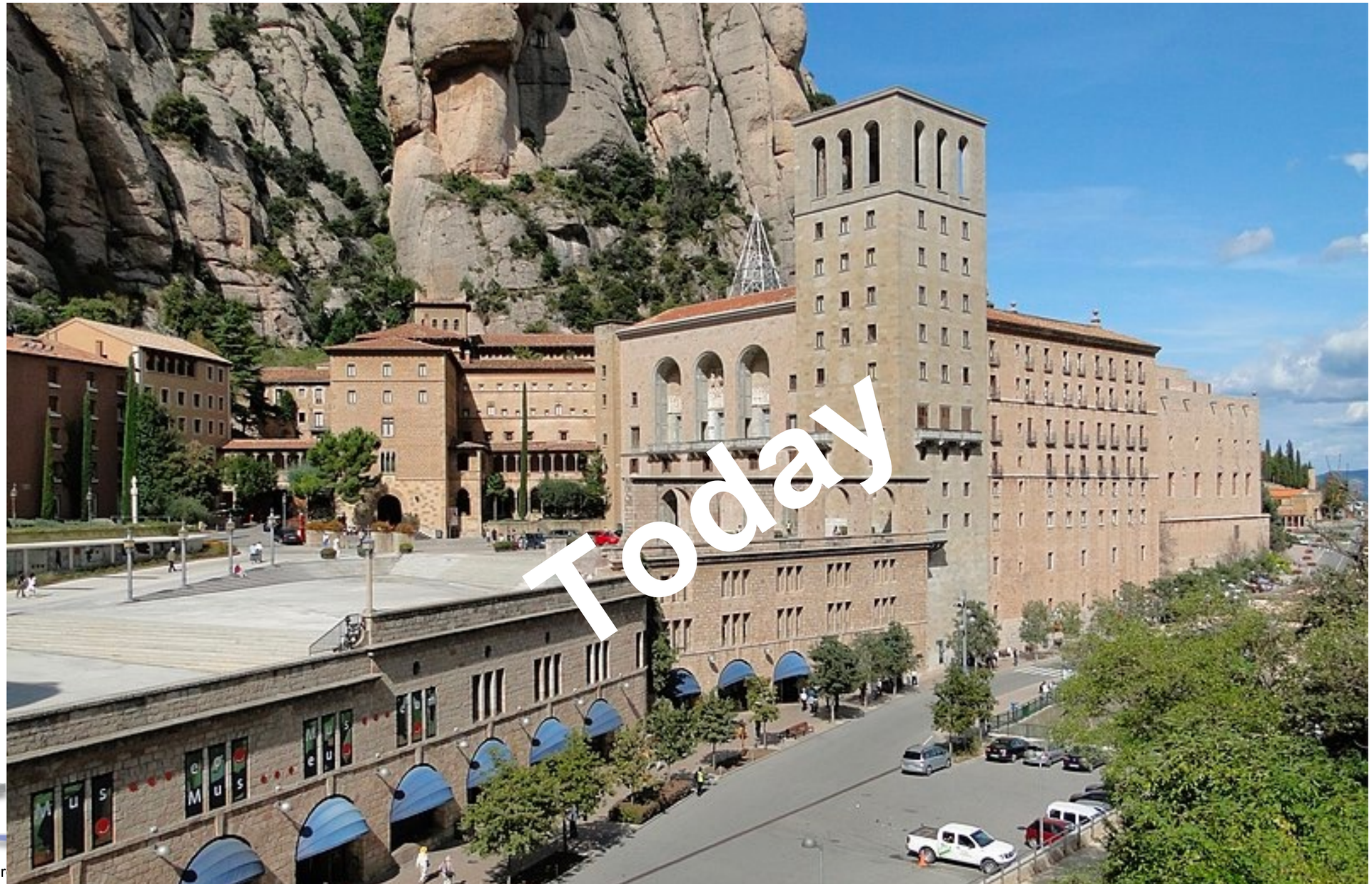
Ruleset 1

- What did we learn about the initial rule set from the ABM?
  - The number of categories against which applicants' needs are assessed was reduced (from 17 categories in the SSM-CAT self-sufficiency matrix to 6 categories)
  - The game needed a termination condition - added:
    - » Before the next round
    - » When critically needy applicants' needs aren't met in a given round, this reduces the upcoming round's available budget
    - » The assumption is that these applicants will end up drawing on social services elsewhere in the system.
    - » The game ends if there is no budget left at the beginning of a round to allocate to any of the applicants.
  - An 'overall wellbeing score' was removed as it was redundant

Ruleset 2

- What did we learn about what should be in ruleset 2 from running the ABM?
  - Have agents assigned different numbers of applicants
  - Rather than having one overall budget, give each agent an individual budget to allocate
  - Separate the critical need score from the ordinary need score, and if the applicant is not seen by any agent, increase the critical need score by one
  - Even when an applicant does receive support, it may not be effective in reducing their need

Stakeholder workshop



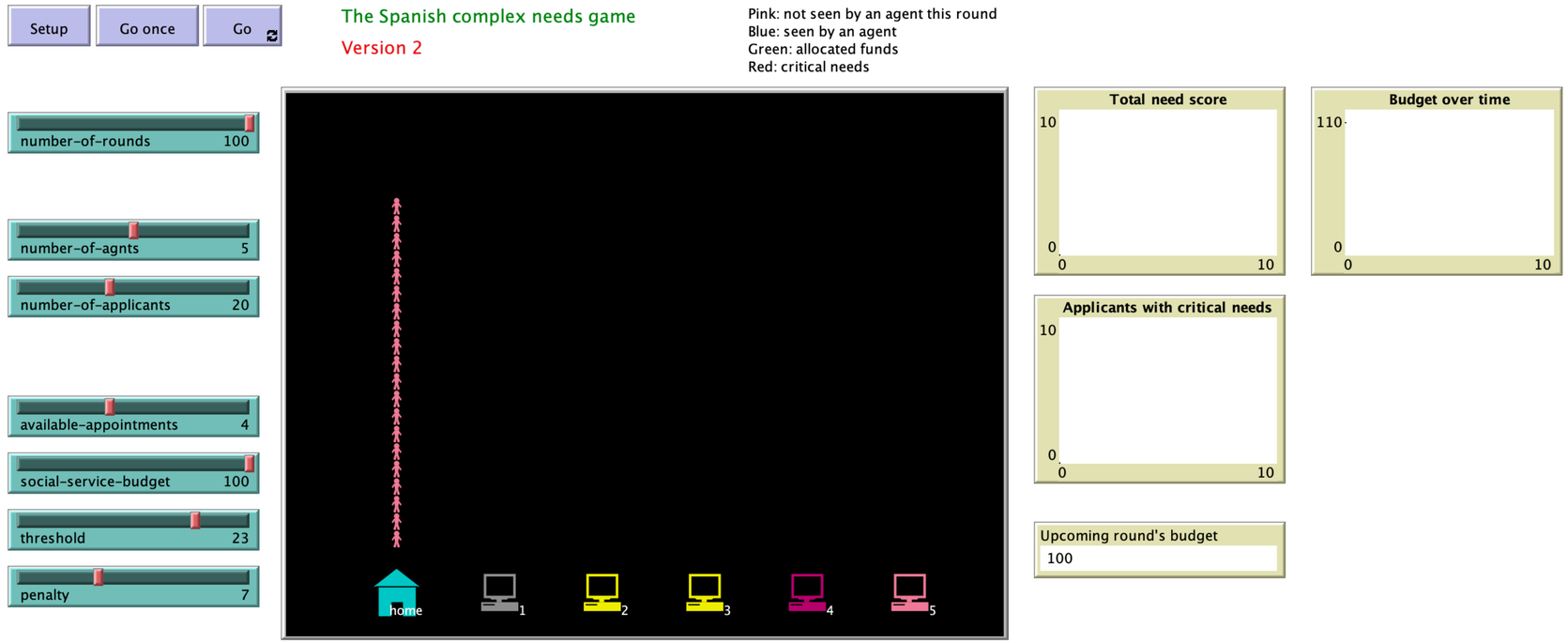
AI FORA
artificial intelligence for assessment

Centre for Resear

VolkswagenStiftung

The ‘Better AI’ ABM

- The stakeholders at the workshop should generate ‘better’ rules, i.e. ones that they feel more comfortable with and that yield assessments that they think are ‘fair’.
- We will then program these rules (Ruleset 3) into the ABM



Better AI

- The revised ABM, with ruleset 3, will be simulate lots of applicants and lots of assessments (in principle, the outcomes of assessing every possible different applicant).
- This generates a data set:

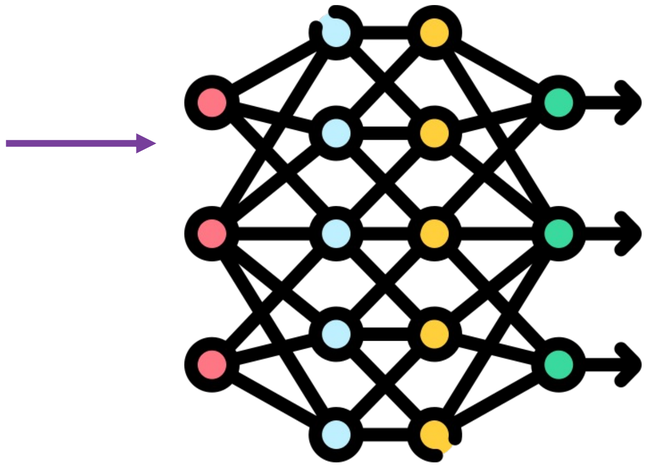
The number of rows is the product of number of categories for each attribute

Applicant #	Income	Number of dependents	Work and training	Mental health	Physical health	Needs score	Assessment
1	1	1	1	1	1	1	Award
2	2	1	1	1	1	1	Award
3	1	2	1	1	1	1	No award
4	2	2	1	1	1	1	Award
5	1	1	2	1	1	1	No award
6	2	1	2	1	1	1	No award
7	1	2	2	1	1	1	No award
8	2	2	2	1	1	1	Award
9	1	1	1	2	1	1	Award
…	…	…	…	…	…	…	…

Neural net

- The data set is used by a machine learning algorithm to create a neural network that, given a particular applicant (i.e. a set of attributes of that applicant), produces an assessment.
- This assessment should be same as the Stakeholders would have given (and the same as the revised ABM would have given) for that applicant.

Applicant #	Income	Number of dependents	Work and training	Mental health	Physical health	Need score
2	2	1	1	1	1	1



Applicant #	Assessment
2	Grant

AI FORA
artificial intelligence for assessment

What next?

- Complete the process for the Spanish case study
  - Stakeholder workshop, Revised ABM, Neural network
- Validate the result for the Spanish case study at a final Stakeholder workshop
- Create ABMs and follow the process for more case studies
- Refine the process as we gain experience
- Publish what we find