Aims

- To develop a framework for semi-automated and automated knowledge discovery by teams of human and robot scientists.
- This collaborations will produce scientific knowledge more efficiently than either could alone.
- To make scientific research more efficient: cheaper, faster, better.
The Concept of a Robot Scientist

Computer systems capable of originating their own experiments, physically executing them, interpreting the results, and then repeating the cycle.

- Background Knowledge
- Hypothesis Formation
- Analysis
  - Experiment selection
  - Robot
  - Results Interpretation
- Final Theory
Eve: Robot Scientist
Application area: the diauxic shift

- Yeast (S. cerevisiae).
- Transition between growth phases:
  1) turn glucose into ethanol.
  2) turn ethanol into CO₂.

- Cancer
- Ageing
Key Challenges

- Scientific knowledge is inherently uncertain.

- Within the AdaLab framework we are developing Bayesian methods that make inferences and plan experiments under uncertainty. We developed an ontology for representing uncertain knowledge.

- Systems biology deals with extremely complex phenomenas, and integrates computational modelling, experimentation, and biological interpretation.

- AdaLab, as an AI system, is designed to automate such cycles of investigations.
AdaLab partners & structure

1. Semantic bioinformatic database
   - Brunel
2. Computational biomodels
   - Evry
3. Leuven, Paris-Nord
4. Manchester
5. Decide on hypotheses
6. Form new hypotheses
7. Plan experiments to test hypotheses
8. Knowledge about experiments
   - Brunel, Manchester
10. Yeast experiment
Partners

- Brunel University London (UK): Coordination, knowledge representation, a communication mechanism between robot and human scientists.

- The University of Manchester (UK): Scientific leadership, Eve experimentation, the application.

- KULeuven (Belgium): Machine learning (ILP), probabilistic inference.

- Universite Paris-Nord: Machine learning (active learning), cost minimisation, model revisions

- Universite Evry: bioinformatics and systems biology (development of an integrated metabolic and gene signalling network).
Key Outputs

- Integration of the components for learning, simulation, revision and experimental design into an adaptive laboratory system Adalab.
- Application of AdaLab to improve diauxic shift models: high reduction of prediction error (30% - 75%)
Summary
Model revisions: M0-M1

- **M0:** the best existing model - Zimmer model (from hundreds of articles).

- **M1:** is built using an in-house ensemble network inference method, using M0 as prior model, and focusing on subset of genes for extension as the top ones in CoRegNet rankings; + bioinformatics database.

- **Improvements:** The network inference produces many new links that we filter using cross validation on Brauer dataset. **M1 outperforms M0.**
Model revisions: M1 - M smart

- **M1** is considered as a Hypothesis to devise experiments to improve the model. Experiment selection taking into account inconsistencies between various kinds of simulated gene expression (forward /backward) has been implemented.

- Revision from 80 carefully selected experiments

- **Improvements:** M1-smart outperforms M1
Model revisions: M1 - M smart
Model revisions: M1 - M random

- **M1-rand**: a revision of M1 using high-throughput experiments. A random set of 80 genes was selected to identify strains to experiment with.

- Selected 40 experiments for which M1 random and M1 smart disagree most, on these:

- **Improvements**: M1-rand outperforms M1; M1-smart outperforms M1-rand
Next Steps

**WP2:** Produce a final version of the knowledge representation of the application domain.

**WP3:** Finalise works on the integrated AdaLab framework and the application.

**WP4:** Improve the revision and experiment design strategies and to evaluate and validate the learning results with real experiments

**Overall:** to make all the Adalab components work better together
Dissemination

- Public Engagement
  - London Science Museum Event
  - Interview for Belgian television channel
- Dissemination to non-specialist audiences
  - Project Website and promotional literature
  - Newsletters
  - Public lectures and seminars (e.g. Pint of Science; Pendle Science café; Summer School of Science)
- Dissemination via specialist media
  - Presentations (44 given to date)
  - Journal publications (7 generated to date)
- **International AdaLab Workshop in June 12, 2018 in the Royal Society, London.**
Thanks