

**Webinar | IMI2 – Call 14  
Development of a platform for  
federated and privacy-preserving  
machine learning in support of drug  
discovery**

**16 March 2018**

# Agenda

- How to use GoToWebinar – Catherine Brett, IMI
- Introduction – Colm Carroll, IMI
- The Call topic – Hugo Ceulemans, Janssen Pharmaceutica
- Opportunities for SMEs
- Questions & answers

# How to use GoToWebinar - audio

To listen via your computer, select **Computer audio**

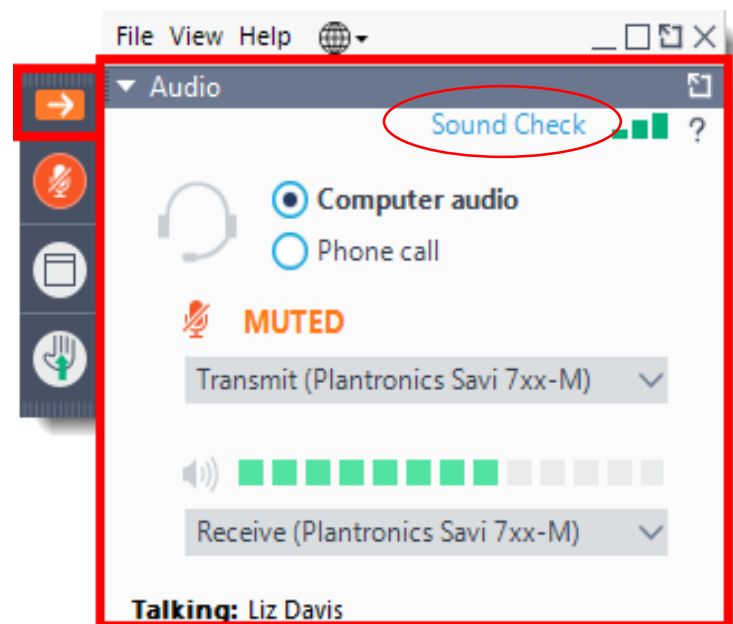
## Can't hear us?

- Check your **speakers are switched on and not muted**
- Do a **Sound Check** to make sure GoToWebinar is picking up the right speakers
- Still not working? Select **Phone call** and dial the numbers given on your phone

To listen in via your phone, select **Phone call**, pick your country, and dial the numbers given

## Can't hear us?

- Check you have selected **Phone call** in the audio panel
- Try **another country's** phone number
- Still not working? Select **Computer audio** and dial the numbers given on your phone



# How to use GoToWebinar

Expand / minimise control panel →

Microphone status →

Full screen →

Raise / lower your hand  
e.g. if you want to ask a  
question orally

Send a question in writing →

The screenshot shows the GoToWebinar interface with several key elements highlighted by a red border and green arrows:

- Expand / minimise control panel:** A green arrow points to a red circle around the expand/collapse icon (a right-pointing arrow) in the top-left corner of the control panel.
- Microphone status:** A green arrow points to the microphone icon in the control panel, which is currently muted (indicated by a red slash).
- Full screen:** A green arrow points to the full-screen icon (a square with a diagonal line) in the control panel.
- Raise / lower your hand:** A green arrow points to the hand icon in the control panel, which is currently raised (indicated by a green hand).
- Send a question in writing:** A green arrow points to the text input field in the "Questions" section, which contains the placeholder text "[Enter a question for staff]".

The interface also displays the following information:

- Audio settings: "Computer audio" is selected, "Phone call" is unselected. The status is "MUTED".
- Microphone: "Transmit (Plantronics Savi 7xx-M)" and "Receive (Plantronics Savi 7xx-M)".
- Volume: A green volume bar is shown.
- Current speaker: "Talking: Liz Davis".
- Webinar title: "Webinar Housekeeping".
- Webinar ID: "Webinar ID: 608-865-371".
- GoToWebinar logo.

# Before we start...

- This webinar is being recorded and will be published on the IMI website and / or IMI YouTube channel
- Presentation slides will be published on the webinar web page
- A participant list will be circulated and published on the website
- IMI2 – Call 14 has been launched and all Call documents & details of how to apply can be found on the IMI website



# Webinar | IMI2 - Call 14 Privacy Preserving Machine Learning

Colm Carroll  
IMI webinar • 16.03.2018

# Today's webinar

## Will cover all aspects of the Call topic

- Introduction to IMI programme
- Proposed project

## Will not cover rules and procedures

- A webinar on rules and procedures will take place on **Wednesday 11 April, 10:30 - 12:00**  
Register [here](#)

# IMI – Europe's partnership for health

## IMI mission

IMI facilitates open collaboration in research to advance the development of, and accelerate patient access to, personalised medicines for the health and wellbeing of all, especially in areas of unmet medical need.



# IMI 2 budget (2014 – 2024)

## EU funding goes to:

Universities  
Hospitals  
SMEs  
Mid-sized companies  
Patient groups



€1.638 bn



€1.425 bn

Other  
€213 m

IMI 2 total budget  
€3.276 billion

## EFPIA companies

receive no funding  
contribute to projects 'in kind'

Associated Partners  
e.g. charities, non-EFPIA companies

# IMI – Ecosystem for innovative collaborations

- Allow engagement in a cross-sector, multi-disciplinary consortium at the forefront of cutting-edge research
- Provide the necessary scale by combining funding, expertise, knowledge, skills and resources
- Build a collaboration based on trust, creativity and innovative and critical thinking
- Learn from each other - new knowledge, skills, ways of working
- Take part in transformative research that will make a difference in drug development and ultimately patients' lives

IMI is a **neutral platform** where **all involved** in drug development can engage in **open collaboration** on **shared challenges**.

# How a topic is generated

Industrial partners align themselves around a real challenge for industry and agree to work together **and commit resources**

New ideas from public sector, universities, SMEs etc. are needed to address the challenge

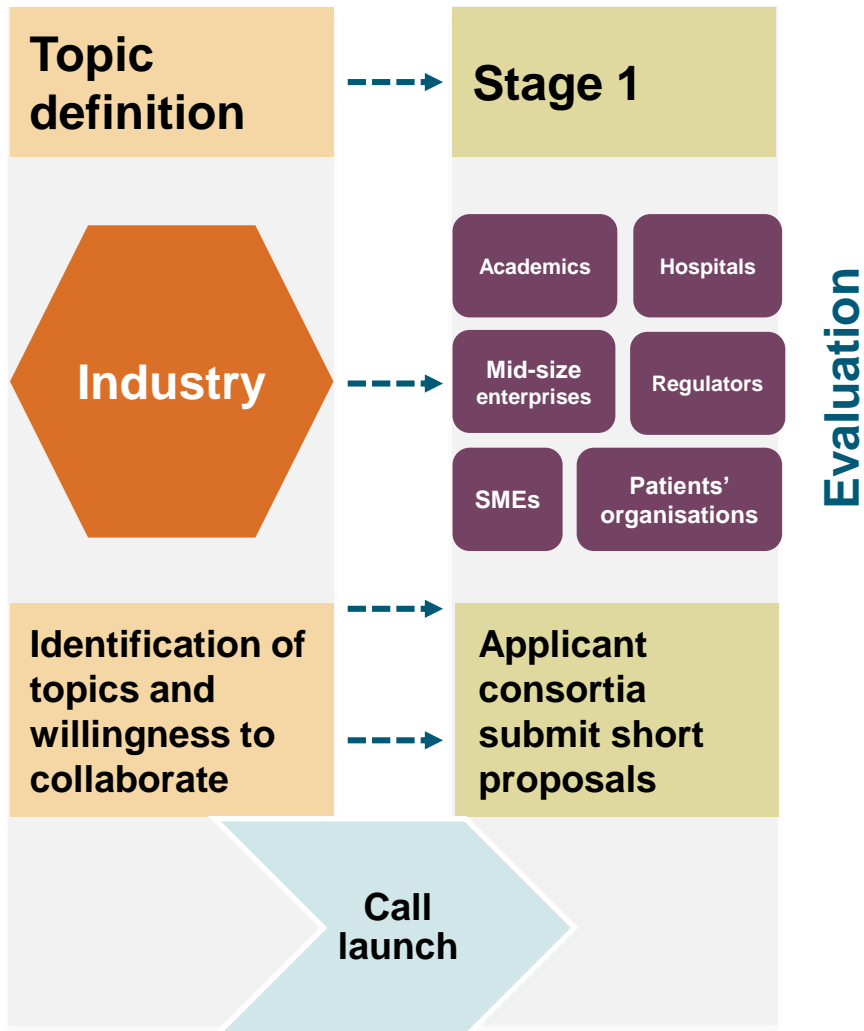
Scale is a key to success and is provided through IMI funding

Outcomes should be transformative for the industry as well as having a clear “public” value

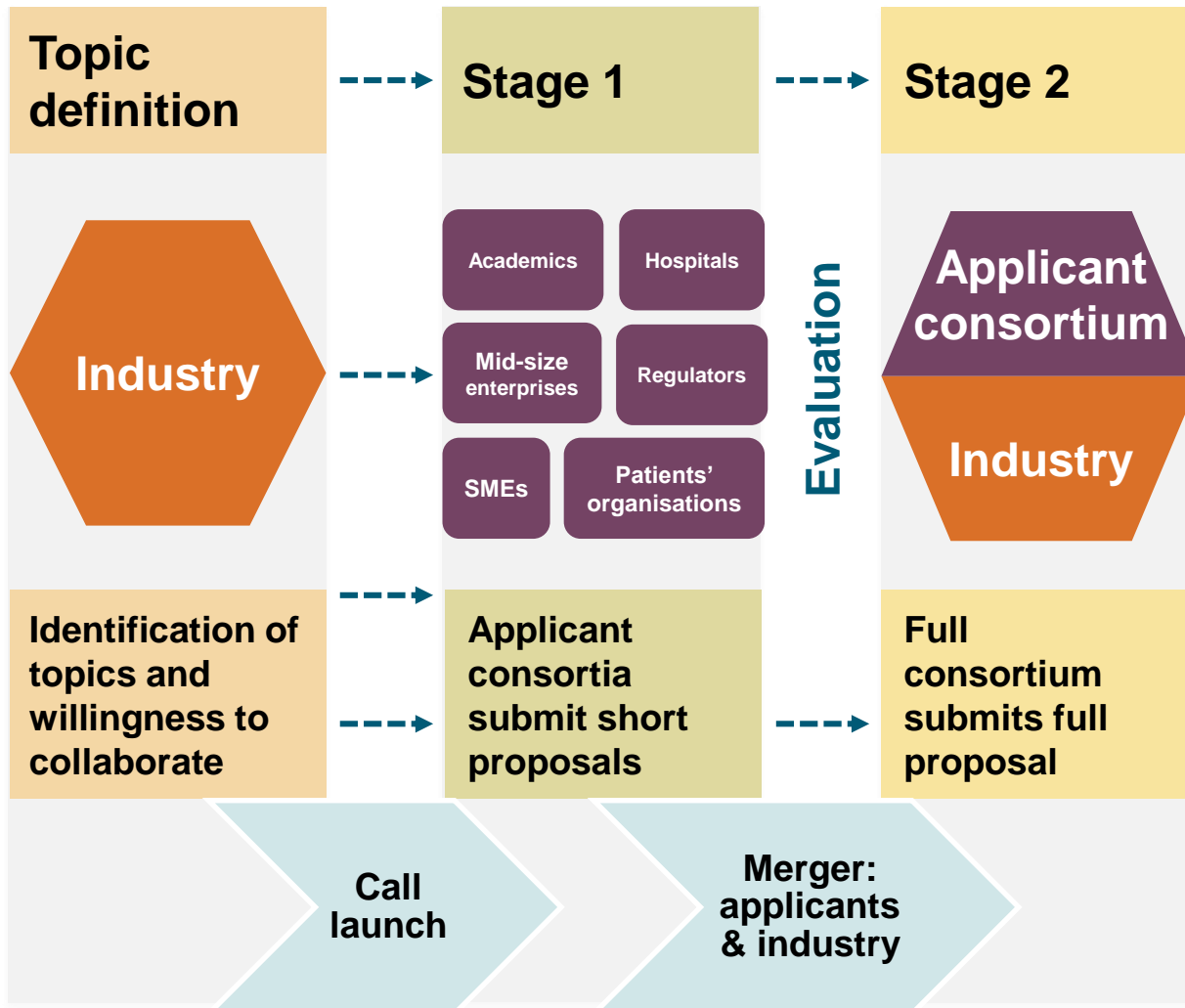
# Typical IMI project life cycle



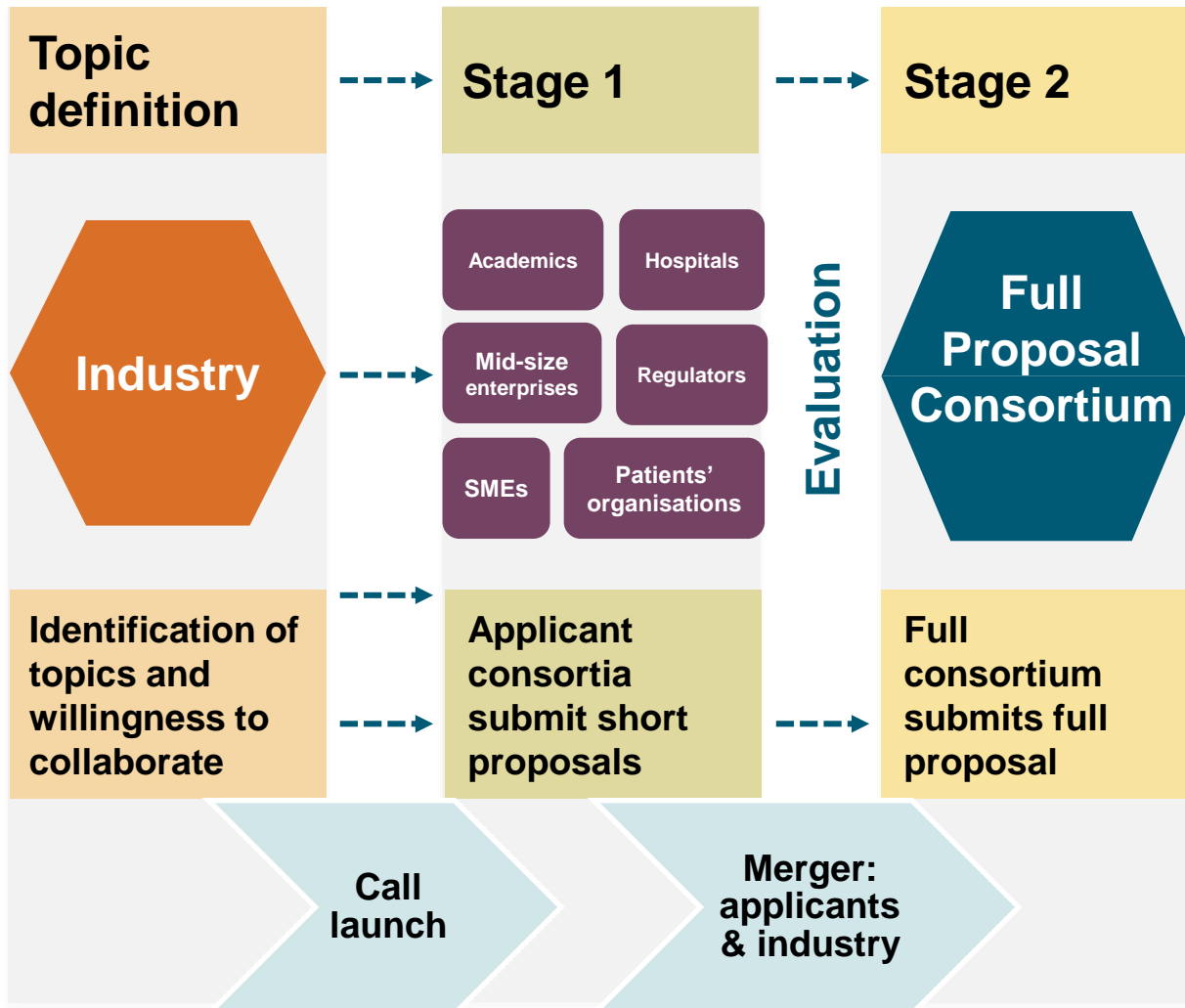
# Typical IMI project life cycle



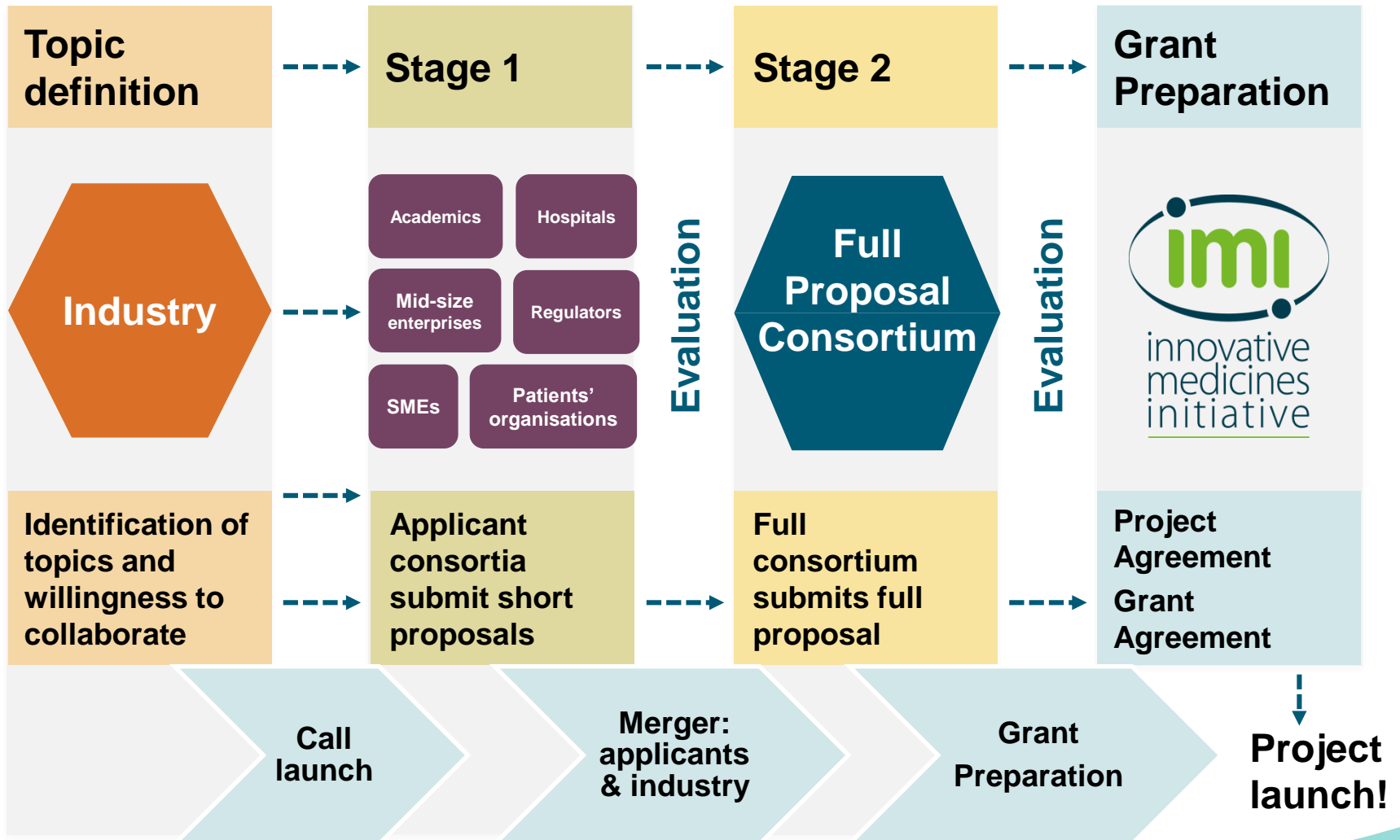
# Typical IMI project life cycle



# Typical IMI project life cycle



# Typical IMI project life cycle





# Submitting a proposal

- <http://europa.eu/!Mg84kq>

The screenshot displays the European Commission Research & Innovation Participant Portal. The header includes the European Commission logo and the text 'RESEARCH & INNOVATION Participant Portal'. A navigation bar contains links for 'HOME', 'FUNDING OPPORTUNITIES', 'HOW TO PARTICIPATE', 'EXPERTS', and 'SUPPORT', along with a search bar and 'LOGIN' and 'REGISTER' buttons. The main content area is titled 'Calls for Proposals' and features a 'Horizon 2020' section with a globe icon. Below this, there are two main categories: 'Excellent Science' and 'Industrial Leadership', each with several sub-topics. A 'Status' filter is set to 'Calls with forthcoming topics' and 'Calls with open topics'. The 'Sort by' dropdown is set to 'Publication date', and a search filter 'IMI2' is entered in the search bar. A sidebar on the left lists 'EU Programmes 2014-2020' with various program categories, including 'H2020' which is highlighted.

European Commission  
RESEARCH & INNOVATION  
Participant Portal

European Commission > Research & Innovation > Participant Portal > Calls

HOME FUNDING OPPORTUNITIES HOW TO PARTICIPATE EXPERTS SUPPORT Search LOGIN REGISTER

EU Programmes 2014-2020

Search Topics

Updates

Calls

H2020

3rd Health Programme

Asylum, Migration and Integration Fund

Consumer Programme

COSME

Internal Security Fund - Borders

Internal Security Fund - Police

Justice Programme

## Calls for Proposals

Horizon 2020 [Advanced search for topics](#)  
[Calls for tenders on TED](#)

- Excellent Science
  - European Research Council (ERC)
  - Future and Emerging Technologies (FET)
  - Marie-Sklodowska-Curie Actions
  - Research Infrastructures
- Industrial Leadership
  - Leadership in enabling and industrial technologies (LEIT)
    - Information and Communication Technologies

Status  Calls with forthcoming topics  Calls with open topics  Calls with only closed topics

Sort by  Call title  Call identifier  Publication date

# Proposal Template

- Available on IMI website & H2020 submission tool
- For first stage proposals, the page limit is **30 pages**.

Title of Proposal

List of participants

Table of Contents

|                                      |   |
|--------------------------------------|---|
| <b>1. EXCELLENCE</b>                 | <b>3. IMPLEMENTATION</b>  |
| 1.1 Objectives                       | 3.1 Outline of project plan — Work packages, and major deliverables |
| 1.2 Relation to the call topic text. | 3.2 Management structure and procedures                             |
| 1.3 Concept and approach             | 3.3 Consortium as a whole   |
| 1.4 Ambition                         | 3.4 Table 3.1a: List of work packages                               |
| <b>2. IMPACT</b>                     | <b>4. PARTICIPANTS</b>  |
| 1 Expected impacts                   | 4.1. Participants (applicants)                                      |

# Evaluation Criteria (1/2)

## ■ Excellence

- Clarity and pertinence of the proposal to meet all key objectives of the topic;
- Credibility of the proposed approach;
- Soundness of the concept, including trans-disciplinary considerations, where relevant;
- Extent that proposed work is ambitious, has innovation potential, and is beyond the state of the art;
- Mobilisation of the necessary expertise to achieve the objectives of the topic, ensure engagement of all relevant key stakeholders.

## ■ Impact

- The expected impacts of the proposed approach as mentioned in the Call for proposals;
- Added value from the public private partnership approach on R&D, regulatory, clinical and healthcare practice as relevant;
- Strengthening the competitiveness and industrial leadership and/or addressing specific societal challenges;
- Improving European citizens' health and wellbeing and contribute to the IMI2 objectives.

# Evaluation Criteria (2/2)

- **Quality and efficiency of the implementation**
  - Coherence and effectiveness of the outline of the project work plan, including appropriateness of the roles and allocation of tasks, resources, timelines and approximate budget;
  - Complementarity of the participants within the consortium (where relevant) and strategy to create a successful partnership with the industry consortium as mentioned in the topic description in the Call for proposal;
  - Appropriateness of the proposed management structures and procedures, including manageability of the consortium.

# Tips for writing a successful proposal

- Read **all the call-relevant material**:  
[www.imi.europa.eu](http://www.imi.europa.eu)
- Begin forming your consortium **early**  
Partner search tools & networking events
- Provide **reviewers** with all the information requested to allow them to evaluate your proposal
- **Finalise and submit your proposal early**
- Contact the **Programme Office** (**NOT** topic writers):  
[infodesk@imi.europa.eu](mailto:infodesk@imi.europa.eu)

# Common mistakes

- Admissibility/Eligibility criteria not met:
  - submission **deadline** missed
  - minimum of **3 legal entities** from **3 H2020 member & associated states** not met
- The proposal does **not address all the objectives** of the topic
- A proposal is **scientifically excellent** but will have **limited impact**
- **Complementarity** with Industry consortium not well described.

# Find project partners

- Network with **your contacts**
- **Network** with fellow webinar participants
- Use **Partner Search Tools**:
  - H2020 portal: <http://europa.eu/!Mg84kq>
  - German NCP version: <http://www.imi-partnering.eu>
- Get in touch with your **local IMI contact point**:  
[www.imi.europa.eu/content/states-representatives-groups](http://www.imi.europa.eu/content/states-representatives-groups)
- Talk to your **Health National Contact Point (NCP)**
- Network on **social media** (e.g. IMI LinkedIn group)

# SME Participation

IMI encourages the participation of SMEs in applicant consortia as they can offer a complementary perspective to other organisations.

For example, being closer to the market, SMEs can drive the tangible outputs of the project, and help ensure these outputs are sustained beyond the project lifetime and therefore help lead to faster impact on healthcare.

Therefore, where possible, include SMEs in your Short Proposal.



# Development of a platform for Federated & Privacy-preserving Machine Learning in support of Drug Discovery

**Ceulemans Hugo**  
16 March 2018 • IMI webinar

# Need for public-private collaboration

- To efficiently meet and exceed patient and regulatory expectations, return on data investment in drug discovery must be improved
  - leverage more than project-specific data, using multi-task learning
  - unlock untapped data by deep learning of more complex features
  - learn jointly, over federated & privacy-preserving machine learning
- To address fragmented data ownership that mars big data potential
  - many more big data spaces with multiple owners than with one
  - naïve mutual data and model sharing most often unviable
  - learning with differentiated data and model ownership required
- Complementary expertise and interests of academia, SMEs and industry needed to realize this new concept and related opportunities
  - SaaS, clouded HPC, data and computation privacy & security, use cases with potential in other industries, ...

# Objectives of the full project

- Delivery of a safe and scalable platform for federated and privacy-preserving machine learning
  - **privacy preservation**: strict protection of each partner's data assets. Activities of compounds in assays, assay annotations and resulting models remain under respective owner control.
  - **federated machine learning**
    - learning itself is distributed over physically separated partners (beyond consolidated learning over federated data, which does not preserve model privacy)
    - no ensembling of data for specific learning tasks across partners (← assay nature is not disclosed). Predictive performance improvements stem from multi-task effect, across partners
  - combines 'federated' (shared) and 'private' (like the data, under owner control) model components

# Objectives of the full project

- Evaluation of security, scalability, operational and predictive performance on extensive pre-clinical and private data from industry (dwarfing public data – and generated out of scope of the data)
  - 5M+ cpds with 100M+ dose-response quality activity points
  - 10M+ cpds with 1B+ single conc. points (as from HT screening)
  - several sets of high-content activities (100s – 1000s of readouts per well, e.g. from microscopy or transcriptional screens) acquired at high throughput (100k+ compounds)
- This evaluation requires the generated of aggregated predictive performance metrics (to be shared and published). Persistent individual predictive models and their performance are not essential to the evaluation – however, the ability to generate them is needed for future exploitation of the platform.

# Pre-competitive nature

- The project aims to realize and evaluate the new and generic concept of federated and privacy-preserving learning. Post project, availability of such platform is aimed for as an affordable but sustainable service to any constellation of partners interested.
- Federated and privacy-preserving learning (and hence the formation of big data spaces) implies standardization of data and preparation. The EFPIA partners will openly release software they develop for affordable standardization and encourage applicants to accommodate this.
- Publication and dissemination of
  - (sufficiently detailed) requirements, guidelines, processes and workflows to document established best practices
  - evaluation of the (aggregated) predictive performance gains.
  - identification of any barriers to uptake, and solutions to reduce

# Expected impact

- Currently, many potential big data spaces blocked by need to fully share data and models. Federated & privacy-preserving machine learning will stimulate the formation of many new big (public-private) data spaces with fragmented data ownership, also beyond pharma.
- For drug discovery: cross-industry federated learning will help virtualization of costly and time-consuming in vitro & in vivo tests.
- For health: future big data spaces of privacy-sensitive health and clinical data will help inform drug development, public and private pharmaco-economics, data-driven best practice formulation.
- Opens new lines of research and applications that complement data ownership with model ownership concepts.
  - opportunities in SaaS, clouded HPC, data security and privacy
  - opportunities in new business models for data owners and generators too

# Suggested architecture of the project

- WP1: sufficiently standardized pre-processing of data
  - method selection, implementation and execution
- WP2: industrial IT technical scoping and deployment
  - formulation of industrial IT requirements, auditing against those, industry side deployment (ensure platform respects intentions)
- WP3: federated machine learning algorithms and analysis
  - prototyping and analysis of predictive and privacy performance (on public data)
- WP4: industry evaluation of privacy/predictivity analysis and results
  - evaluation of privacy/predictivity balance in analysis and implementation and evaluation of aggregated results (on industry data)

# Suggested architecture of the project

- WP5: implementation to enterprise-ready software
  - implementation of methods from WP3, to meet IT requirements from WP2 and privacy/predictivity balance from WP4
- WP6: secure stand-alone and federated infrastructure
  - set-up of secure infrastructure for industry data owners and of infrastructure connecting owner infrastructures
- WP7: operations and deployment
  - sustainable detailed operating model and monitoring run execution
- WP8: project governance & management, dissemination & sustainability
  - strategic, operational, IP and financial management & administration
  - detailed sustainability plan to make results accessibly beyond project



# Expected contributions of applicants

- Delivery of a platform for federated and privacy-preserving machine learning
  - scientific software prototypes and analysis on public data (WP3\*)
  - enterprise-ready software (ready to be audited against industry defined requirements) (WP5\*)
  - enterprise-ready infrastructure solutions (ready to be audited against industry defined requirements) (WP6\*)
  - operating model and deployment (WP7\*)
  - project governance, dissemination and sustainability (WP8)

Applicants and industry to be represented in each WP, but applicants are the main drivers of WPs marked with (\*). Applicants and industry contribute equally to WP8.

# Expected contributions of industry

- Definition of industrial use case and requirements and enablement of evaluation of the platform on industrial scale data
  - sufficiently standardized preprocessing of data (including public data preprocessed to the same standards) (WP1\*)
  - industrial IT technical scoping and deployment (WP2\*)
  - industry evaluation of privacy/predictivity analysis and results (WP4\*)
  - project governance, dissemination and sustainability (WP8)
  - funding for independent audit against requirement and project administration support

Applicants and industry to be represented in each WP, but industry are the main drivers of WPs marked with (\*). Applicants and industry contribute equally to WP8.

# What's in it for you?

- Academic researchers:
  - novel lines of research in machine learning, combining aspects of performance, novel data sources, privacy and security and scale in the societally and industrially relevant context of health.
- SMEs:
  - novel types of machine learning applications in the context of SaaS, clouded HPC, infrastructure, ... . Application in pharma, life sciences and health and beyond. Exposure to knowledge partners and industry side IT and user community.
- Industry:
  - virtualization of drug discovery, best practices, data sharing solutions, ...

# Key deliverables of the full project

- coherent federated and privacy-preserving machine learning platform by month 12 and updated annually.
  - prototype scientific software version available 10 months before enterprise ready release, to enable documentation and analysis
  - analysis of predictivity and privacy preservation using public data
  - enterprise ready (i.e. audit-ready) software and infrastructure
  - from 2<sup>nd</sup> year onwards ability to benefit from high-content datasets
- evaluation on industrial data
  - appropriately prepared datasets; open standardization software
  - definition of and audit against security and privacy requirements
  - enablement and evaluation of runs
- publication & dissemination of best practices, sustainability plan



**Thank you**

[www.imi.europa.eu](http://www.imi.europa.eu)



@IMI\_JU



# SME participation

Colm Carroll  
IMI webinar • 16.03.2018

# SME Participation

IMI encourages the participation of SMEs in applicant consortia as they can offer a complementary perspective to other organisations.

In particular, in this topic, SMEs can participate by bringing expertise in:

- Hands-on expertise in solutions for big data handling at industrial scale, ICT security and information leakage aspects, high performance computing infrastructures, software engineering
- Machine learning technologies in the context of federated learning
- Deploying computational approaches in drug discovery and development
- General project management in the context of EU-funded projects



# Questions



# Questions?

Raise your hand  
if you want to ask a  
question orally



The screenshot shows a webinar interface with two main panels. The top panel is titled 'Audio' and includes a 'Sound Check' indicator. It has radio buttons for 'Computer audio' (selected) and 'Phone call'. A microphone icon is labeled 'MUTED'. Below this are dropdown menus for 'Transmit (Plantronics Savi 7xx-M)' and 'Receive (Plantronics Savi 7xx-M)'. A volume slider is also present. The bottom panel is titled 'Questions' and contains a text input field with the placeholder '[Enter a question for staff]' and a 'Send' button. At the bottom of the interface, it says 'Talking: Liz Davis', 'Webinar Housekeeping', 'Webinar ID: 608-865-371', and the 'GoToWebinar' logo.

Send a question in writing



After the webinar, send any questions  
to the **IMI Programme Office**

[applicants@imi.europa.eu](mailto:applicants@imi.europa.eu)