

Conseil d'administration de l'Institut Polytechnique de Paris
29 septembre 2021



Pièce n°08

Research: strengths

- **AI research is developing on fast pace in IP Paris.**
 - It is one of the first European academic centers by the number of faculties (120) and PhD students (350).
 - These numbers have more than doubled in the last 5 years.
 - 1.5-2 % of papers at the largest conferences in the field (20 papers at ICML in 2021 out of 1180 accepted papers, over 100 submitted papers). On par with the largest “single site” European academic centers (TUM, EPFL)
- **In some areas, IP Paris has strong recognition**
 - Computational learning theory, computational optimal transport, distributed and federated learning, computational game theory (matched markets, competing bandits, ...)
 - Generative models and Bayesian inference (computer graphics, acoustic scene synthesis, uncertainty quantification)

Research: strengths

- **Interdisciplinary research in AI is also developing fast with applications to**
 - high-energy physics,
 - AI-related innovations for energy efficiency,
 - Finance (fintech)
 - Cybersecurity
 - Economy (matching markets), legal tech, ...
- **An eco-system from bringing AI developments from the labs to the market**
 - IP Paris has a long tradition of interacting with enterprise-driven eco-system for AI research and innovation activities
- **The largest in France population of top Maths / Computer Sciences students on a single Campus**
 - Compete with the best institution in the world at Senior BsC level
 - The largest Grad School in France (more than 300 PhDs, 500 students at MSc level)

Research: weaknesses

- **The competition for top talent has become fierce on an international scale**
 - Competition with foreign universities and academic institutions (much more attractive salary offer – not necessarily welcome packages-, difficult to attract at full professor level)
 - Competition with private AI research Labs in France and in Europe (Google, Deepmind, Facebook, but also Criteo)
- **Despite recent advances, research remains fragmented in different labs / institutions**
 - research still remain largely in “institutional silos”
 - tangle of structures, laboratories, national research institution, each with their own policies
- **Impossible to get into the AI "hype" tasks of the moment, which are gradually slipping away from academy**
 - Reinforcement Learning (largely a Google-Deepmind affair nowadays)
 - Natural Language Processing (OpenAI GPT 3 175.10⁹ parameters, trained on 500.10⁹ of textual units, 3000 MM of top engineers and researchers, illimited computational ressources, cost 50.10⁶ €)

Hi! Paris: long-term goals

- **Decompartamentalizing research between IP Paris components, creating a common identity and fostering research synergies**
 - To fight fragmentation in a pragmatic way – not creating yet another layer !-, to incentivize the dynamics of collaborative research,
 - Develop both **core AI**, **AI for sciences** and **AI for society** by becoming an international leader in certain areas - building on existing strengths in finance, energy efficiency, climate, physics, mechanics, chemistry, etc...
- **Train the AI elite, attract the best PhD students to our laboratories**
 - Supports the expansion of AI education and research (500 PhD Students, 800 MsC level students in AI in 2025, require to keep our best students – fight the US brain drain- and also to attract best students abroad)
 - But also.. provide current and future generations of all our engineers with the necessary skills and competencies in AI and to anticipate labour market trends.
- **Attract the world's top academic researchers**
 - [Long term] Build a structure as visible as MILA or the Vector Institute in Canada

Research: Chairs – Faculty positions

- **Chairs – Faculty positions**

- The center funds faculty positions for new researchers.
- Two levels: Junior (Assistant and Associate Professorships) and Senior (Full Professorships). IP Paris and HEC Paris follow their own internal recruitment procedures to identify and propose candidates.
- Total number of chairs depending on budget and additional endowment.

- **Results**

- 1 Associate Professor position (Econometrics/Applied Econometrics at ENSAE – **full endowment**)
- 2 Assistant Professors: Machine learning for computer vision at Telecom Paris – **full endowment** - ; Statistical Physics / AI at Ecole polytechnique – **partial endowment**
- 1 (invited) “top notch” Professor in AI for energy at Telecom Paris – **partial endowment**
- IP-Paris: Open offer for 1 senior professor [with competitive offers].

Research: Fellowships

- **Fellowships – internal call for fellowships**
 - The Center offers 3-year funding for researchers from HEC and IP-Paris (flexible budget sharing between wage premiums, hiring, organization of scientific events and funding of PhD students)
 - The aim is to **retain our top researchers and increase impact**. The call for proposals similar to ERC model.
 - Four types of grants: Pre-starting (Ph.D $\leq$ 4 years - 90k€), Starting (Ph.D $\leq$ 7 years - 180k€), Consolidator (Ph.D $\leq$ 12 years - 240k€) and Advanced (Ph.D 12 years - 300k€).
 - Original target of maximum 8 fellowships. Fixed number of fellows per call and per institution.
- **Results**
 - 28 applications for Fellowship: 2 at the Pre-starting level, 10 at the Starting level, 5 at the Consolidator level, and 11 at the Advanced level
 - 5 proposals selected for funding: 0 Fellowship at the Pre-starting level, 2 at the Starting level, 1 at the Consolidator level and 2 at the Advanced level)

Research: Fellowships - Proposals

Category	School	Title of the proposal
HEC-Starting	HEC	Compliance Technologies for the Data Supply Chain: Implementing Data Protection by Design and Legal Explainability
HEC-Starting	HEC	Examining AI and Human Evaluations
HEC-Starting	HEC	The Impacts of Algorithmic Decision-makers, Advisors, and Recommenders on Business and Society
IPP-Advanced	ENSAE	SAGMOS: Statistical Analysis of Generative Models: Sampling Guarantees and Robustness
IPP-Advanced	ENSAE	Machine learning tools to address labor market frictions
IPP-Advanced	Ecole Polytechnique	APPEAL: Aligning Performance and Principles in Ethical Learning
IPP-Advanced	ENSAE	Private Dynamic Modeling
IPP-Advanced	Ecole Polytechnique	VERNE: VERification of NEural nets
IPP-Advanced	Telecom Sud Paris	TITANS --- Trustworthy Intelligent And resilient cyber-physical Systems
IPP-Advanced	Ecole Polytechnique	CREATIVE AI: CREating and Animating Testbeds of Ideas, through Virtual Environments with Artificial Intelligence
IPP-Advanced	Ecole Polytechnique	Advancing Graph Neural Networks and Applications
IPP-Advanced	ENSAE	HD-Bayes
IPP-Advanced	Telecom Sud Paris	Increasing the energy efficiency of Transfer Learning models
IPP-Advanced	Telecom Paris	Senior Fellowship Proposal - Machine leArNing For cOmPLex Data and tasks (MANIFOLDS)
IPP-Consolidator	ENSAE	Macroeconomic nowcasting with high-dimensional Google Search data: theory and practice
IPP-Consolidator	Telecom Paris	SINNet - Natural Language Processing and Socio-Inspired Neural Networks for Conversational Systems
IPP-Consolidator	Telecom Paris	DELPIC: Deep Learning for Public-key Cryptography
IPP-Consolidator	Ecole Polytechnique	MATTER: Domain adaptation from theory to practice
IPP-Consolidator	ENSAE	Designs of Optimal Online Matching Markets
IPP-Prestarting	Ecole Polytechnique	Argumentation Mining for Social Good: Creating a Global Encyclopedia of Debates
IPP-Prestarting	Telecom Paris	Adapting deep generative networks for data-efficient training
IPP-Starting	ENSAE	Causal AI for the Evaluation of Private and Public Policies
IPP-Starting	Ecole Polytechnique	Federated Learning: new Algorithms with Guarantees - FLAG
IPP-Starting	Ecole Polytechnique	Trade-offs in Robust Distributed Learning
IPP-Starting	ENSTA	Quantifying Uncertainty of Deep Networks for autonomous driving
IPP-Starting	Telecom Paris	Large-scale data depth: computation and applications
IPP-Starting	Ecole Polytechnique	Temporal Logic to Trust Perception-based Autonomus Systems (TAPAS)
IPP-Starting	ENSAE	Interplay between Data and Geometry in Non-Linear Spaces

Research: PhD scholars

- **15 fundings for PhD Scholars**

Candidate	School	Project	Supervision of
-----------	--------	---------	----------------

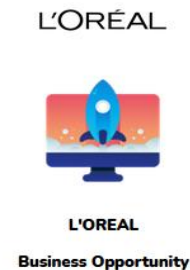
Research: Collaborative Academic Projects

- 5 funded projects

Researchers involved	Schools involved	Project
----------------------	------------------	---------

Education: Hi!ckathon, Hackathon data challenge

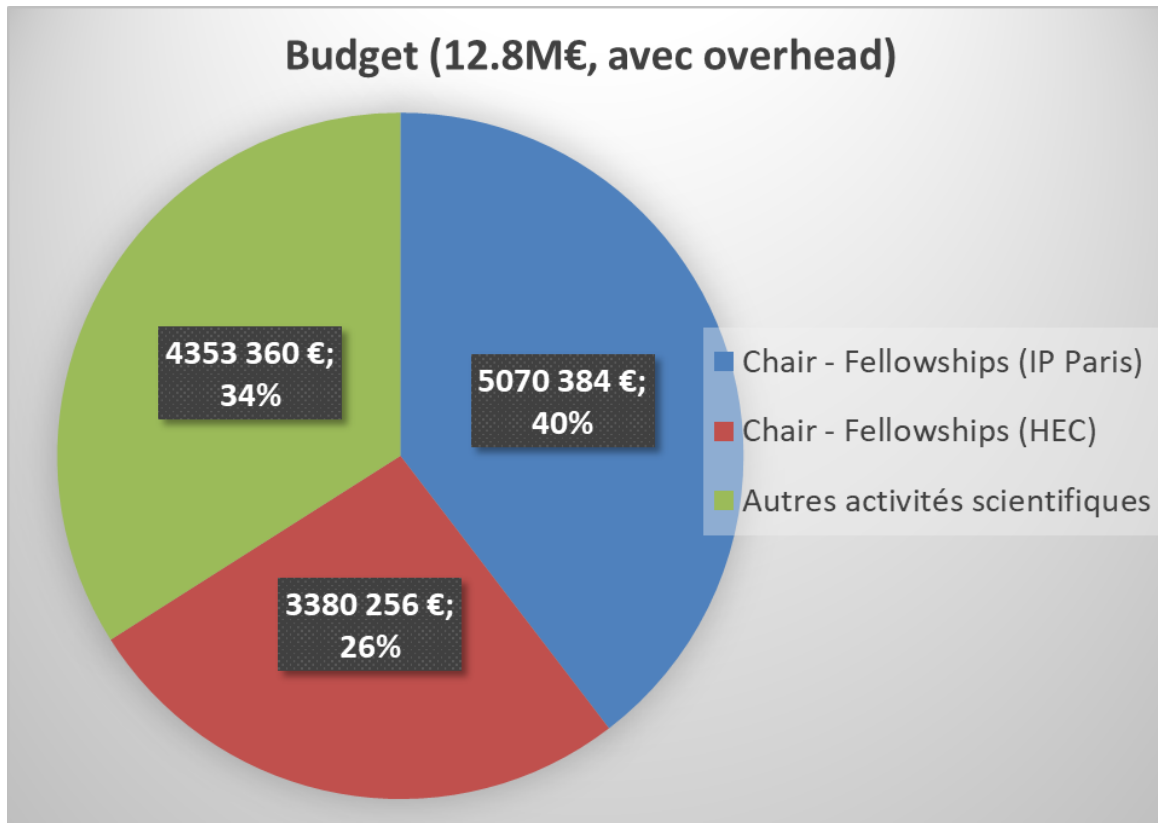
- **Event : March 2021**
- **Theme “AI for Energy Efficiency”**
 - Optimising the energy costs of three buildings for frugal solutions. 3 deliverables: scientific approach, data collection and algorithmic design, business pitch.
 - Participation
 - 187 students from HEC Paris and IP-Paris
 - 35 interdisciplinary teams and 69% of the teams were mixed HEC and IP-PARIS students
- **Five awards sponsored by the Center’s founding corporate donors**



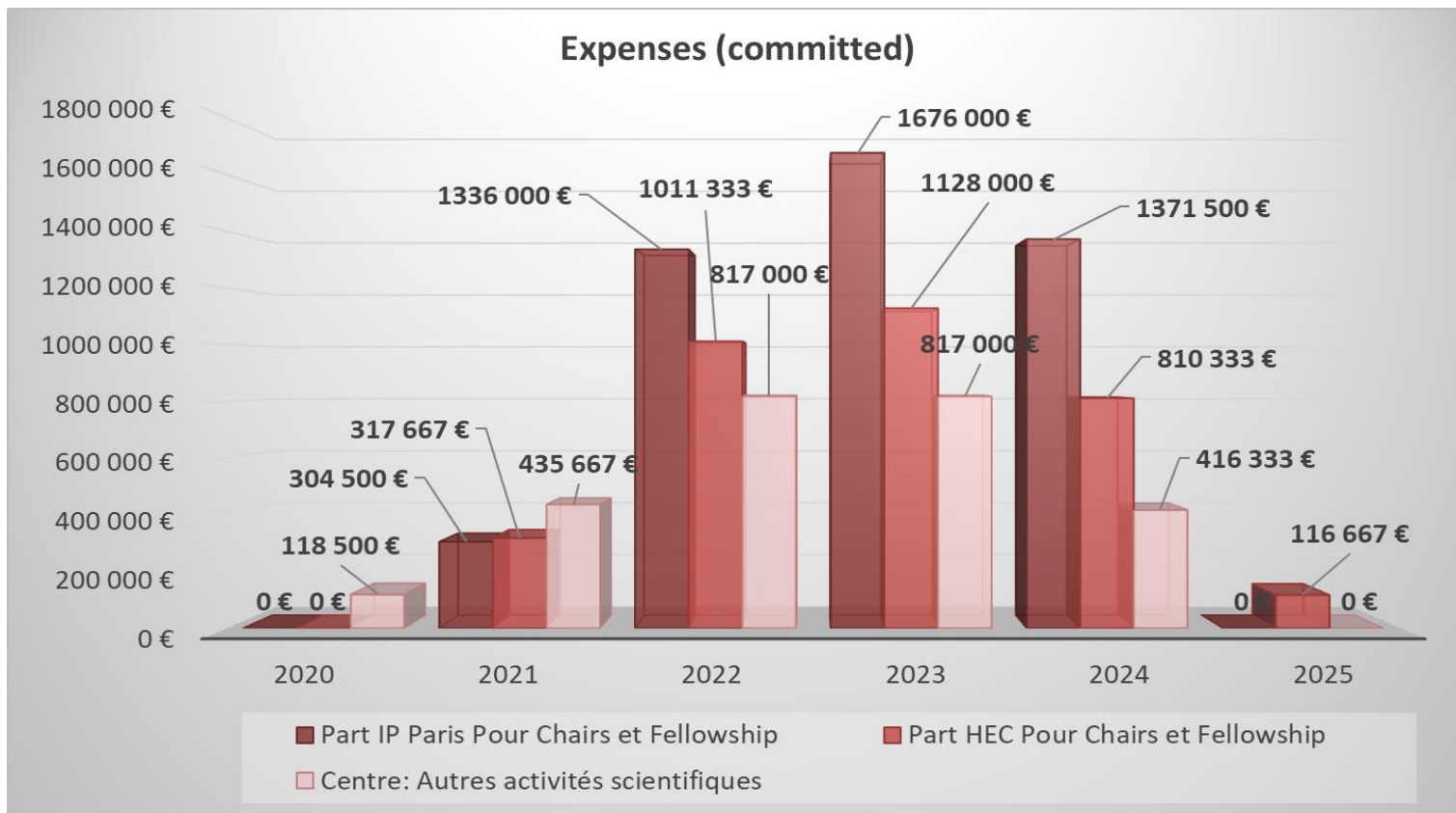
Research: 1st Summer School Hi! PARIS

- **Program** <https://www.summerschool.hi-paris.fr>
 - Themes: Machine Learning in Asset Management, Data Visualization, Stochastic Optimization and Federated Learning.
 - Target audience: PhD students, academics who want to expand their knowledge in these areas, and research engineers / teams from corporate donors.
 - 3 days, there will be
 - 2 panels: Industry panel with corporate donors / Academic panel
 - 2 keynote speakers: Claire Monteleoni, University of Colorado Boulder - Chair for Inclusive Excellence and Michel Wedel, University of Maryland - Chair in Consumer Science
 - 8 tutorials in two parallel sessions: Track A: Data Science for Business and Society ; Track B: Theory and methods of AI.
- **100 active participants**
 - despite the hybrid mode of participation and the weariness of distance learning workshops / summer schools that "really" don't work

Budget overview (1/3)

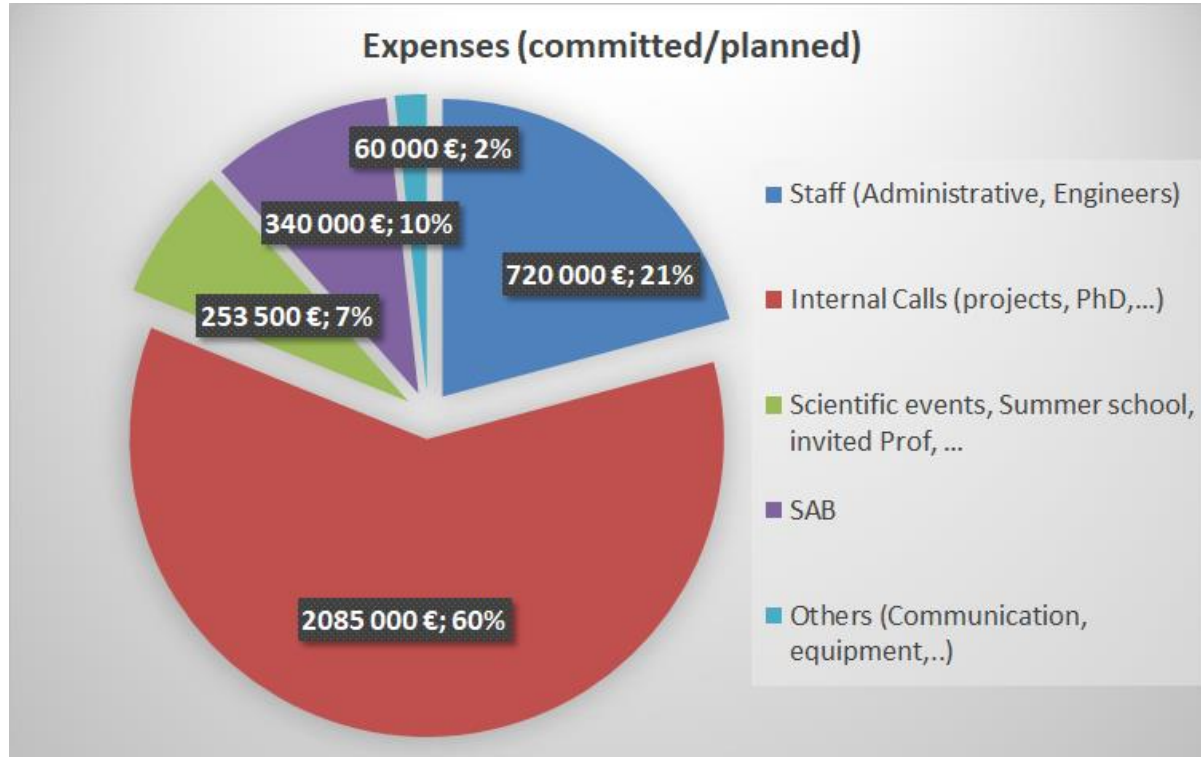


Budget overview (2/3)



Budget overview (3/3)

- Other (center) scientific activities



Hi! PARIS new actions:

- **Research**

- Reading groups with corporate donors
- « research in pair » program [small groups of researchers / engineers on topics given by corporate donors]
- Communication, including scientific communication

- **Teaching**

- "Data Bootcamp" project
- 2nd Hi!ckathon project
- Mapping (i)Existing lecture set and (ii)Student projects : timing and organization

Research: Communication

- **Recruitment of the Hi! PARIS project manager (April 2021) in charge of communication**
- **Hi! PARIS website:** www.hi-paris.fr/
 - Open calls and recrutement / News page (in preparation)
- **Meetings with IP PARIS and HEC communication teams**
 - draft communication plan in progress
- **Future agenda**
 - Hi! PARIS **Social networks**: Launch at last in september 2021
 - Hi! PARIS **Welcome workshop** for fellows and recruits (main topic, research plan): October 2021
 - Welcome workshop in October 2021 to present the fellows and the recruits (main topic, research planWorkshop/webinar)
 - Hi! PARIS **Newsletter** (4/year): Launch at last in October. 2021
 - Portrait of Hi! PARIS new recruited fellows/PhD scholars/collaborative academic projects
 - Hi! PARIS Activities summary (open call for proposals, results, webinars, student events, etc.)
 - Publication visibility of HEC and IP PARIS researchers within the scope of Hi! PARIS → repository/library of the research meant in the scope of Hi! PARIS
 - Signature of the publication of the researchers funded by Hi! PARIS

Teaching: Data Events

- **Organisation by Capgemini**
- **Lessons learned from the 1st hackathon:**
 - better timing
 - embedding in the syllabus of two data events (Hi!ckathon and Data Bootcamp)
- **Two data events are proposed**
 - **2nd Hi!ckathon**, a competitive event that brings together teams of students to create an MVP to solve a Data Science problem
 - Long format: a 5-week educational adventure culminating in an intense hackathon on one week-end. Jan. 17 – Feb. 16 + Feb. 19-20
 - **Data Bootcamp**, a before "back-to-school" seminar to raise passion and bring out talents
 - Format: 2x 1 week of courses (integrating theory sessions and group coaching). Aug. 23 – Sept. 03, 2021.

Teaching: "Data Bootcamp" project

Principe :

Afin de rythmer la semaine et de ludifier l'apprentissage, il conviendra d'alterner des séquences présentant des objectifs pédagogiques différents :

Présentations en plénière

Cours théoriques

Mises en pratique

Restitutions en plénière

Exemple d'une session d'une semaine :

A affiner lors de la phase de préparation

Lundi	Mardi	Mercredi	Jeudi	Vendredi
Kick-off de l'évènement	Data Processing / Feature engineering	Machine learning : intuitions et applications	Machine learning : fonctionnalités avancées	Sprint final : finalisation des projets
Modules de remise à niveau	Challenge : Analyse exploratoire du dataset	Challenge : Conception de l'algorithme niveau 1	Challenge : Conception de l'algorithme niveau 2	Présentation des projets
Présentation de la problématique business	Livrable intermédiaire : dataset nettoyé et enrichi	Echange entre groupes, nouveaux indices	Data visualisation	REX et clôture
Découverte du dataset	Use case mécènes 1	Use case mécènes 2		

CAB du 11 juin 2021 – Les deux événements data prévus en 2021/2022

Organisation en cours



Evénements

POC

Année 2020/2021

Hackathon "Flash" 48h

Evénement compétitif réunissant, sur une période courte, des individus ou des équipes autour de la création d'un MVP pour résoudre une problématique Data Science.

Phase pilote

Année 2021/2022

Data Bootcamp 2 x 1 semaine

Un séminaire de pré-rentree autour de la création d'un programme de Machine Learning pour une acculturation alternant théorie et pratique

Hackathon long 5 semaines

Un hackathon de deux jours, sur le modèle du hackathon flash, précédé d'une phase de montée en compétence des étudiants de 5 semaines

Déploiement

Année 2022/2023

Data Bootcamp 2 x 1 semaine

Même principe que les événements 2021/2023 (potentiels ajustements à décider à l'issue de la phase pilote)

Hackathon long 5 semaines

Data Challenge ?

Ou autre événement à définir



Objectifs

1. **Rayonnement du centre :** interne (écoles) et externe (attirer de nouveaux mécènes)
2. **Rôle fédérateur :** rassembler les élèves des 6 écoles et engager une collaboration mécènes / enseignants / élèves

1. **Pilote pédagogique :** démontrer aux responsables de formation la valeur pédagogique des événements
2. **Test de l'octroi d'ECTS :** Impliquer un ensemble de formations « pilotes » dans une première initiative commune délivrant des ECTS
3. **Consultation large :** Directeurs des études, responsables de formations, directeurs écoles
4. **Conception des maquettes définitives :** au plus tard en janvier 2022, pour partage aux responsables pédagogiques avant la finalisation par les écoles des programmes de l'année 2022/2023

1. **Intégration complète dans les cursus :** Inscrire pleinement les événements dans les calendriers
2. **Octroi d'ECTS :** Si la phase pilote est concluante, délivrer des ECTS dans le cadre des formations sélectionnées
3. **Convergence :** Assurer une trajectoire commune des écoles du Centre sur l'IA et la Data, ainsi qu'une convergence enseignement/recherche

Teaching: Mapping (i)Existing lecture set and (ii)Student projects : timing and organization

- **Construction of the catalogue with two objectives:**
 - 1. Identify opportunities for courses/lectures given by corporate donors:**
 - Description of the courses, with potential identification of the specific courses where it is thought that an intervention/collaboration of a corporate donor would be interesting (*e.g. presentation of specific applications, platforms, usecases,..*)
 - 2. Identify projects type that could be proposed by corporate donors:**
 - Description of all curricula where projects are carried out, with identification of projects characteristics (*duration, size in number of students and number of hours/days, type of programs such as specialized masters, first year or final engineering programs, ...*)

Hi! PARIS Activities – Summary Roadmap

	sept-20	oct-20	nov-20	déc-20	janv-21	févr-21	mars-21	avr-21	mai-21	juin-21	juil-21	août-21	sept-21	oct-21	nov-21	déc-21
Chair Programme (target for hiring: September 1st)										★			★			
Fellowship Programme (target for beginning: September 1st)						★				★			★			
Hi! PARIS Programme calls																
Collaborative academic projects (target for beginning: September 1st)										★			★		★	
Phd theses (target for hiring: September 1st)										★			★			
Hi! PARIS events																
Summer School (preparation, invitation, website, etc...)											★					
Hackathon, Data challenge							★									
Data Bootcamp													★			
Scientific events (webinar, Meetup, workshop on a specific topic)					★			★								★
Yearly Hi! PARIS Scientific event													★			