

**What I did
after my PhD**

Yet Another Math PhD

UNIVERSITÉ
PARIS DIDEROT
PARIS 7

UNIVERSITÉ PARIS VII - DENIS DIDEROT
SORBONNE PARIS CITÉ
École Doctorale de Sciences Mathématiques de Paris Centre

THÈSE DE DOCTORAT
Discipline : Mathématiques

présentée par

Léo DREYFUS-SCHMIDT

**Équivalences perverses splendides,
Conditions de stabilité et
Catégorification du complexe de Coxeter**

**Splendid Perverse Equivalences, Stability Conditions and
Categorification of the Coxeter Complex**

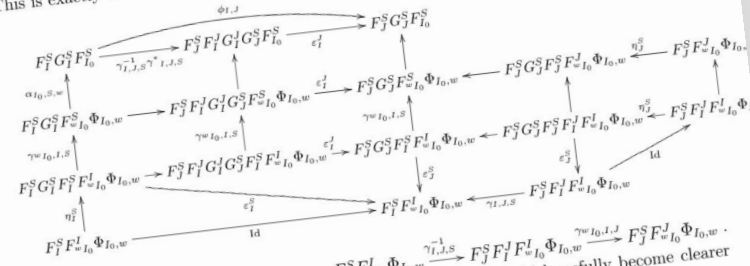
Dirigée par Michel BROUÉ

Soutenue le 26 Septembre 2014 devant le jury composé de :

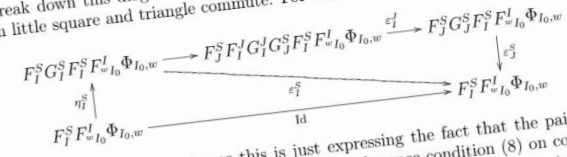
Prof. Cédric BONNAFÉ	Université Montpellier II	Rapporteur
Prof. Michel BROUÉ	Université Paris Denis Diderot	Directeur
Prof. Marc CABANES	Université Paris Denis Diderot	Examinateur
Prof. Joseph CHUANG	City University of London	Examinateur
Dr. Olivier DUDAS	Université Paris Denis Diderot	Examinateur
Prof. Radha KESSAR	City University of London	Examinateur
Prof. Raphaël ROUQUIER	University of California, Los Angeles	Co-directeur
Prof. Geordie WILLIAMSON	Bonn Universität	Rapporteur

Maths that I'll never use again ...

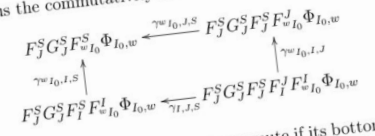
This is exactly what our next massive commutative diagram will tell us.



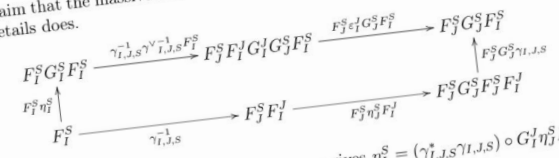
where we get from the bottom row $\psi_{I,J} : F_I^S F_{I_0}^J \Phi_{I_0,w} \xrightarrow{\gamma_{I,J,S}} F_J^S F_{I_0}^J \Phi_{I_0,w} \xrightarrow{\gamma_{I_0,I,J}} F_J^S F_{I_0}^J \Phi_{I_0,w}$.
 Let us break down this diagram into smaller ones so that it would hopefully become clearer that each little square and triangle commute. For instance the bottom left corner reads



Here the lower triangle commutes as this is just expressing the fact that the pair $(\eta_{S,I}, \epsilon_{S,I})$ forms an adjunction. As for the upper triangle, the coherence condition (8) on co-units maps gives us that it also commutes. In the same spirit, the coherence condition on transitivity of the functors F 's gives us the commutativity of the third square of the second row:



We now claim that the massive diagram above will commute if its bottom part reproduced here in more details does.



(7) on unit maps gives $\eta_I^S = (\gamma_{I,J,S}^1 \gamma_{I,J,S}) \circ G_I^J \eta_J^S \Phi_{I_0,w}^J \circ \eta_I^J$ and

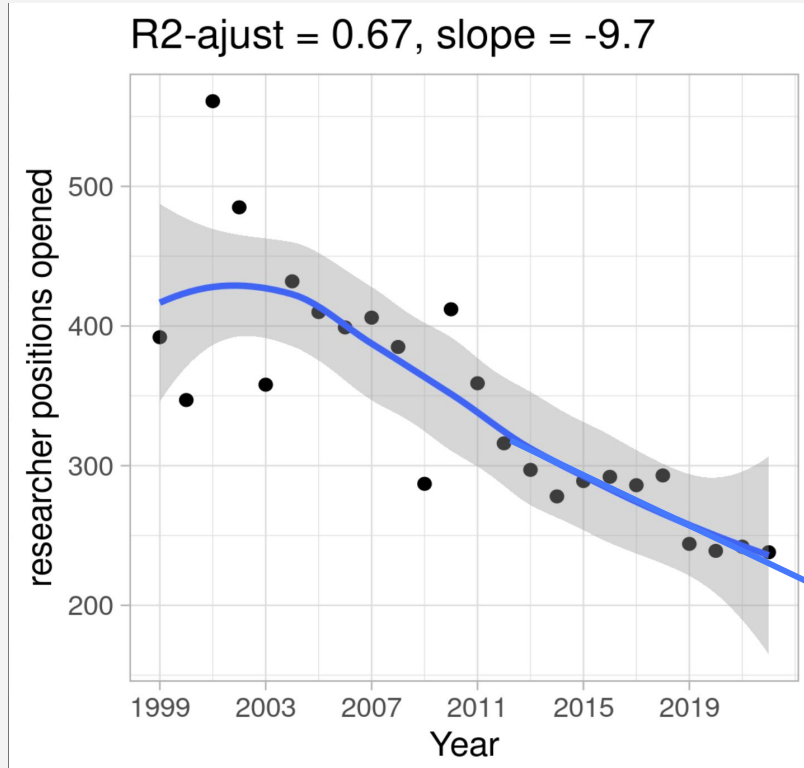
**Post-doc
in
England**



A dark, atmospheric image of a night sea. A bright, glowing light source, possibly the moon or a distant star, is visible in the upper right quadrant, casting a soft, ethereal glow over the scene. The water below is dark and textured, with some faint, distant lights visible on the horizon, suggesting a city or industrial area at night. The overall mood is mysterious and somber.

Les Sirènes de l'Industrie

Le CNRS



2044

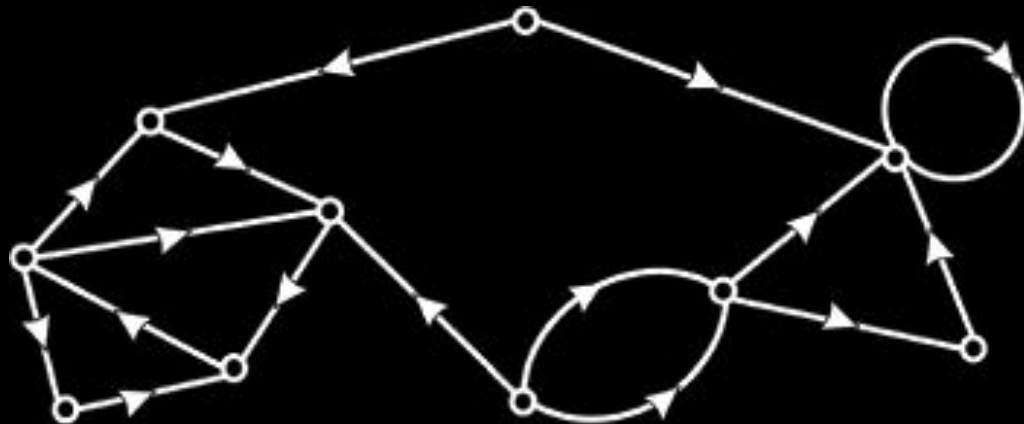
The Big Data



2012

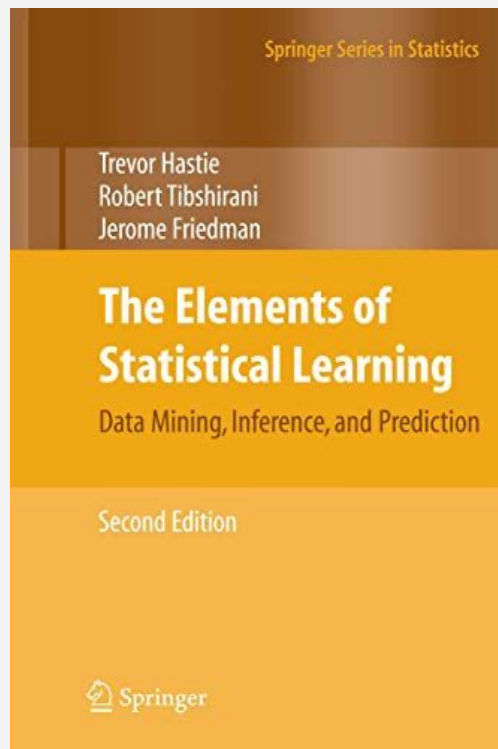


2017



The Path from Algebra to Data

Elements of Statistical Learning



Dimension Reduction?

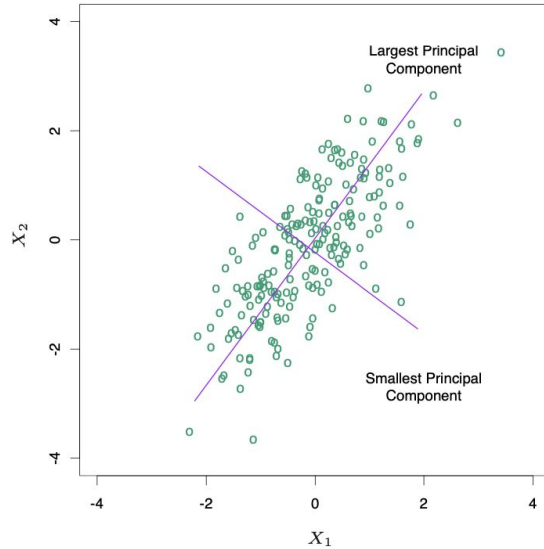
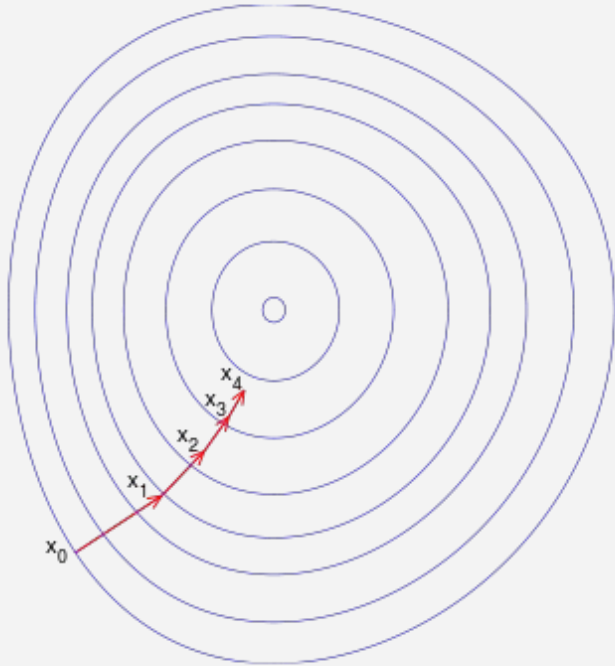


FIGURE 3.9. *Principal components of some input data points. The largest principal component is the direction that maximizes the variance of the projected data, and the smallest principal component minimizes that variance. Ridge regression projects \mathbf{y} onto these components, and then shrinks the coefficients of the low-variance components more than the high-variance components.*

SVD!

Numerical Optimization?



SGD!

Learn Python

The screenshot shows a learning interface with three main sections: a task panel on the left, a code editor in the center, and an output panel on the right.

Task Panel (Left):

- Header: **Tasks** 14/14 Complete
- Text: "otherwise our data will not be correctly sorted anymore!"
- Section 13: "Three mice walk into the store. They don't have much money (they're mice), but they do each want different pizzas." The instruction is to slice the `pizza_and_prices` list and store the 3 lowest cost pizzas in a list called `three_cheapest`.
- Section 14: "Great job! The mice are very pleased and will be leaving you a 5-star review." The instruction is to print the `three_cheapest` list.

Code Editor (Center):

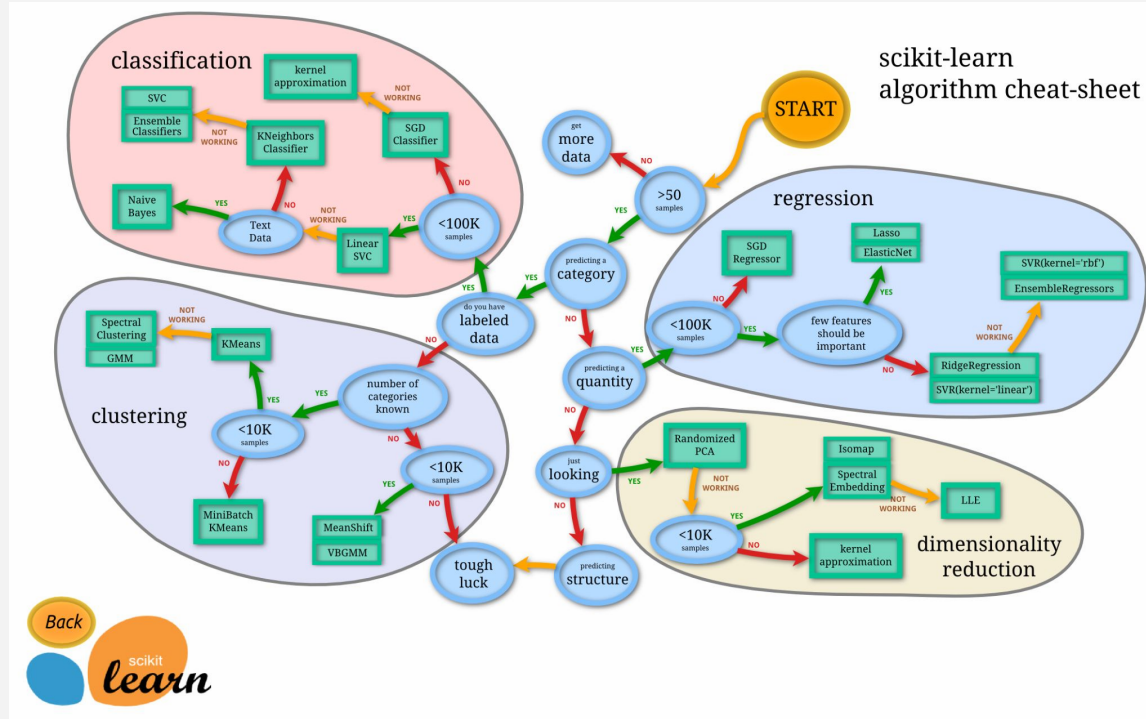
```
script.py
1 # Your code below:
2 # A list of toppings for pizzas
3 toppings = ["pepperoni", "pineapple", "cheese",
4             "sausage", "olives", "anchovies", "mushrooms"]
5 print(toppings)
6 prices = [2, 6, 1, 3, 2, 7, 2]
7 num_two_dollar_slices = prices.count(2)
8 print(num_two_dollar_slices)
9 num_pizzas = len(toppings)
10 print("We sell" + str(toppings) + "kinds of pizzas")
11 pizza_and_prices = [[2,"pepperoni"], [6,"pineapple"], [1,
12                    "cheese"], [3,"sausage"], [2,"olives"], [7,"anchovies"],
13                    [2,"mushrooms"]]
14 print(pizza_and_prices)
15 pizza_and_prices.sort()
16 print(pizza_and_prices)
17 cheapest_pizza = pizza_and_prices[0]
18 pricest_pizza = (pizza_and_prices[-1])
19 pizza_and_prices.pop()
20 pizza_and_prices.insert(0, [2.5, "peppers"])
21 print(pizza_and_prices)
```

Output Panel (Right):

```
['pepperoni', 'pineapple', 'cheese', 'sausage',
'olives', 'anchovies', 'mushrooms']
3
We sell['pepperoni', 'pineapple', 'cheese',
'sausage', 'olives', 'anchovies', 'mushrooms']kinds
of pizzas
[[2, 'pepperoni'], [6, 'pineapple'], [1, 'cheese'],
[3, 'sausage'], [2, 'olives'], [7, 'anchovies'],
[2, 'mushrooms']]
[[1, 'cheese'], [2, 'mushrooms'], [2, 'olives'],
[2, 'pepperoni'], [3, 'sausage'], [6, 'pineapple'],
[7, 'anchovies']]
[[1, 'cheese'], [2, 'mushrooms'], [2, 'olives'],
[2, 'pepperoni'], [3, 'sausage'], [6, 'pineapple'],
[2.5, 'peppers']]
The three cheapest pizzas are [[1, 'cheese'], [2,
'mushrooms'], [2, 'olives'], [2, 'pepperoni'], [3,
'sausage'], [6, 'pineapple']]
```

Footer: 14/14 Complete, Back, Next

Learn ML Python



Job Hunt



Landing a Job

4-5 months transition

Some Rejections


Dataiku ?

Candidature Spontanée

CEO *normalien*

VP Product Math PhD

Dataiku: A Unified Platform to Systemize AI



Centralized workbench for everyone

Streamlined path to production

Integrated in your stack, governed at scale

IT/Data Ops **Data Expert** **Business Expert** **Enterprise Leader**

Operate **Code** **Click** **Consume**

Cataloging and Exploration **Data Prep and Pipelining** **Visualization and Stat Analysis** **Machine Learning** **MLOps** **AI Governance** **AI Apps**

Connect to Existing Data & Compute **Elastic Architecture Built for the Cloud** **Enterprise Identity & Access Controls** **Monitoring, Documentation & Compliance**

©2022 dataiku, Inc. | dataiku.com | contact@dataiku.com | @dataiku

The diagram illustrates the Dataiku platform's unified approach to systemizing AI. It features a central horizontal flow of four user roles: IT/Data Ops (Operate), Data Expert (Code), Business Expert (Click), and Enterprise Leader (Consume). Below this, a series of six icons represent key capabilities: Cataloging and Exploration, Data Prep and Pipelining, Visualization and Stat Analysis, Machine Learning, MLOps, and AI Governance. At the bottom, four foundational pillars are listed: Connect to Existing Data & Compute, Elastic Architecture Built for the Cloud, Enterprise Identity & Access Controls, and Monitoring, Documentation & Compliance. The entire content is framed by three key benefits: Centralized workbench for everyone, Streamlined path to production, and Integrated in your stack, governed at scale.

Data Science Projects



Hiring PhDs (for ~~ML~~ AI)

Upskilling

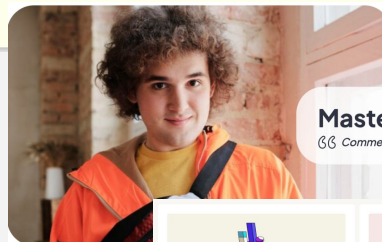


Pet project

Les 15 meilleurs masters en Data Science

Métier Data

Les entreprises recrutent des Data analysts, Data scientists, Data engineers...
! Découvrez les meilleurs masters qui ouvrent les portes de la Data Science.



Master Data science

📖 Comme Jules, 40% des bacheliers utilisent Diplomeo pour trouver leur école



Data Scientist

Bootcamp ou Temps partiel

- ✓ Programmation Python
- ✓ Machine Learning avancé
- ✓ Deep Learning
- ✓ Big Data / Database
- ✓ Système complexe et IA
- ✓ Projet Data

[↓ Voir programme complet](#)



Data Analyst

Bootcamp ou Temps partiel

- ✓ Programmation Python
- ✓ Business Intelligence
- ✓ Machine Learning
- ✓ Dataviz
- ✓ Text Mining

[↓ Voir programme complet](#)



Data Engineer

Bootcamp ou Temps partiel

- ✓ Python avancé
- ✓ Base de données (SQL, NoSQL)
- ✓ Big Data (Hadoop, Spark)
- ✓ Git, GitHub, CI/CD (Jenkins)
- ✓ APIs (Flask, Fast APIs)
- ✓ Airflow, Docker, Kubernetes

[↓ Voir programme complet](#)



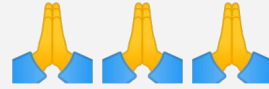
Data Product Manager

Temps partiel

- ✓ Acculturation Data et RGPD
- ✓ Business Intelligence
- ✓ Datavisualisation - Power BI
- ✓ Chefferie de projet
- ✓ Méthode agile
- ✓ Fondamentaux Python
- ✓ Projet Data professionnel

[↓ Voir programme complet](#)

Academic Resume ≠ Industry CV



WORK EXPERIENCE

VP Research Dataiku	2020 - present
<ul style="list-style-type: none"> Responsible for the definition and execution of Dataiku Research roadmap. Managing researchers on broad variety of topics (Causal ML, Data Shift, Data Discovery). 	
Research Director - Head of Dataiku AI Lab Dataiku	2018 - 2020
<ul style="list-style-type: none"> Creation and management of Dataiku's research department that focuses on new ways to bring artificial intelligence to organizations. Construction and research supervision of various Machine Learning projects (Deep Learning, Active Learning, AutoML). 	
Senior Data Scientist Dataiku	2015 - 2018
<ul style="list-style-type: none"> Conception and deployment of data science products (recommender systems, churn prediction, predictive maintenance, pricing,...) for European clients. 	
Graduate Student & Teaching Assistant in Mathematics University of Paris VII	2013 - 2014
University of California, Los Angeles	2011 - 2012
University of Oxford	2010 - 2011

EDUCATION

Ph.D. Pure Mathematics, University of Oxford & University of Paris VII Thesis: Splendid and Perverse Equivalences. tl;dr: Conducted research at the intersection of representation theory, category theory and group theory, leading to three research papers.	2014
M.S. Pure Mathematics, Imperial College, London & University of Paris VII	2010
B.S. Pure Mathematics, Ecole Normale Supérieure de Lyon	2008

SKILLS

Programming Languages: Python, SQL (in many flavour), NoSQL, Spark.

Languages: French (native), English (bilingual).

Main Publications:

- Ensembling Shift Detectors: an Extensive Empirical Evaluation* in ECML-PKDD 2021
- Rebuilding trust in active learning with actionable metrics* in ICDM 2020.
- Splendid and Perverse Equivalences* in Journal of Algebra 2016.

Music: 12 years of classical piano and 5 years of jazz guitar.

What You'll Need To Fight Against

She's going to take 3 years to complete any project.

He can't do quick and dirty. He'll overcomplicate things.

She is not used to *actual* supervision, she'll go rogue.

He'll question everything. He'll spend his time on arXiv.

What You'll Need To Show

You embrace change, you've moved on from your PhD.

You can adapt your communication to your audience.

You are precise and concise.

You care about the business.

Job Targeting

Large companies or startups ?

Prep for the interviews. Know the company, do your homework.

Reach out to people in the company with similar profile.

A Little Thing Called Crédit d'Impôt Recherche

Crédit Impôt Recherche : Calculez le coût réel d'un jeune docteur

Okay
Doc 
OKAYDOC.FR



Voici la formule appliquée :

Salaire brut + cotisations sociales

X temps passé sur des opérations de R&D x 2

X 2 (dépenses de fonctionnement à 200%)

X taux de CIR applicable.

Exemple d'un salaire de 45K€ avec
80% du temps consacré à la R&D

$45\text{K€} \times 1,41 \times 2 \times 0,8 \times 2 \times 0,3 = 60,9\text{K€}$ de CIR

Une fois le CIR déduit, le jeune docteur ne coûte donc que 3000€ à l'entreprise (car son salaire chargé est de 45K€ x 1,41 de cotisations patronales soit 63,45€)



Q&A

Général :

- Combien de temps avez-vous mis à trouver un poste fixe (MCF/CR dans le public et CDI dans le privé) après votre doctorat ?
- Peut-on donner des cours sans devenir MCF ?
- Quelles difficultés avez-vous rencontrées dans votre recherche de poste académique ?
- Avez-vous renoncé à chercher un poste académique du fait de la trop forte compétitivité ?
- Combien de temps à l'avance par rapport à la défense de la thèse est-il suggéré de candidater à des post-docs ?

Sortie du monde académique :

- Comment avez-vous effectué la jonction entre votre thèse académique et le privé ? Êtes-vous passé par des cabinets de recrutement ? Comment avez-vous choisi le poste en entreprise correspondant à votre profil ?
- Avez-vous ressenti le besoin de vous "vendre" ? Comment se "vendre" à une entreprise lorsqu'on a un profil académique comme le nôtre ?
- À quel point votre domaine est-il compétitif ? Y-avait-il beaucoup de pression, d'opportunités professionnelles ?

Travail en entreprise et/ou hors de l'académique :

- Quel aspect de votre thèse s'est révélé le plus important dans votre carrière jusqu'à présent ? Par exemple : compétences techniques (logiciels, codage), mathématiques théoriques, spécialisations dans certains domaines...
- Est-ce que la R&D en entreprise vous a déçu pendant votre parcours professionnel car peu intéressante, peu poussée, ou par manque de temps/moyens ? Vous sentez-vous épanoui intellectuellement ?
- Aviez-vous initialement prévu de "rester dans l'académique" ? Qu'est-ce qui vous a fait bifurquer ? Regrettez-vous d'avoir quitté la recherche académique ?
- Dans l'industrie, les compétences acquises en doctorat sont-elles facilement transférables notamment pour des postes à responsabilités ?
- En France, le doctorat est-il toujours vu comme une surqualification ou commence-t-il à être valorisé notamment pour des postes de direction technique ?
- Un doctorat en mathématiques fondamentales ou appliquées peut-il être disqualifiant pour des postes type ingénieur ou au contraire un atout de par la capacité d'abstraction nécessaire à la réalisation mathématique ?
- Existe il des thèses en entreprise en suisse ou Australie dans le domaine de l'IA ?
- Quel est l'intérêt de faire une thèse selon vous, plutôt que de commencer directement en CDI si on veut être data scientist ?
- Meilleure reconnaissance d'une thèse en IA si on la fait en France ou à l'étranger ?