Industry Careers Extravaganza

David Brind, Claire Coffey, Eloise Withnell, Haoting Zhang

Our Initial Idea

What problem does our solution address?

Through our research (see references at end of document), we found that there was a lack of an event that brings various companies with roles in health data science together. Most events of this format are usually tagged on as an afterthought in conferences or are focused solely on one company, and/or focus mostly on their research and act as a 'sales pitch'. Furthermore, we thought that since HDR UK is the national institute for health data science, they should take an active role in facilitating career searches within their industry. Currently, the careers page from the HDR UK website promotes all of their educational routes to give students the ability to apply for these roles; however it is lacking information about current job opportunities. With the introduction of the new HDR UK alumni network, we believe that this opens up an opportunity to create an annual health data science careers fair that can link closely to the alumni network and help forge new career pathways for HDR UK affiliated students.

If this proposal goes ahead, we believe this will allow students to have an honest understanding of existing health data science companies (or companies with roles in this area). It will also allow companies currently operating in this space to discover potential future employees and potentially build relationships to increase collaborations with academia. Moreover, students can leverage the large breadth of training and educational opportunities presented by HDR UK more effectively once they have an idea of their future career aspirations. Finally, we believe this will strengthen the bond between academic and industrial collaborations, as well as provide a big incentive to joining (and interacting with) the new HDR UK alumni network.

What is our idea?

We are proposing to help launch the alumni network with an annual health data science career fair involving HDR UK students (from all educational routes i.e. PhD, undergraduates, interns, apprenticeships). Companies, both partnered with HDRUK as well as guest businesses, will have the opportunity to present their work in a format that both introduces the business as well as provides an insight into the day-to-day life of working with the company as a healthcare data scientist. As an incentive to companies represented by HDR UK alumni members we propose discounted rates for their stalls.

From our research findings, we will aim to address the industry-focused pathways (biotech, pharma, HealthTech, healthcare organisations), rather than academia pathways. We will focus on the pathways that are open to those with 0 years of industry experience and that require a wide range of academic

qualifications (bachelor, master and PhD level), since we are aiming our event at current students without experience.

As we are keen to focus on students transitioning from bachelor/masters/PhD studies into industry, we believe it would be valuable to have representatives from each of these levels, either as company representatives or as speakers at the event. Industry-specific skill building, such as CV coaching and headshots, would also be a useful addition. We are keen to have keynote speakers that are industry career-focused, with a deep understanding of the skills required for our defined pathways. The Association of the British Pharmaceutical Industry (ABPI) represents over 120 pharmaceutical companies and has a team dedicated to career development. Therefore, a member of this team, such as Andrew Croydon (Skills & Education Policy and Examinations Director at ABPI), would likely be appropriate as a keynote speaker. Additionally, to gain a deeper understanding of the HealthTech industry, we would consider a speaker from Storm3, a HealthTech recruitment company.

What are the outcomes you are seeking to achieve?

We have a couple of key outputs for this proposal. Firstly, we will be running an in-person careers fair involving stands from various companies affiliated with HDR UK. We will be asking companies to produce their posters/stalls ahead of time so that these can be distributed online for attendees who cannot make it, or would like to look back in their own time. If this event cannot be run in person, then we can make use of these as well as software such as Gathertown to run a virtual fair.

Specifically, we propose to produce a brochure for attendees that will outline the itinerary for the day, the floorplan for the event, as well as highlighting the HDR UK alumni network and the benefits in becoming a member. This can include success stories from individuals who have come through the HDR UK education routes and have now reached success in their career roles and are still active within the alumni network.

Logistically, we can set up invites to the careers fair through Eventbrite as this will allow us to monitor attendee numbers and we can potentially leverage the current HDR UK account to send out invites. We propose that attendees are restricted to those from HDR UK education routes (internships, apprenticeships, linked health data science undergraduates courses, postgraduate researchers) however this potentially could be expanded to other students at partner universities and beyond in future years based on the success of the event. We believe that transport to the event for HDR UK students should be paid for, and employers should pay a fee to partake in the event (in return for exposure to potential future employees). This amount will be decided by speaking to individuals who have experience setting up similar events such as career fairs at universities to gauge the appropriate amount. As the event is planned to be in person, the event will begin in the afternoon to allow for people travelling to arrive at the event. Minimal catering facilities will be provided i.e. tea, coffee, biscuits for attendees, however the event will not be a full day, hence we do not expect it to warrant further catering facilities. We can leverage HDR UK channels to promote the event,

as for the first year the attendees will be HDR UK-affiliated (so email chains should be sufficient). Industry partners and employees that are collaborating with HDR UK or are members of The Alliance can be invited through appropriate liaison channels. Reaching out to more employers can be done through various social media channels, and importantly, through connections of those in the HDR UK network.

Outcomes and Outcome Indicators

We expect the students attending the event to have a clearer understanding of their desired career path (related to their progress in education i.e. undergraduate, postgraduate, apprenticeship)

- *We propose short post-event surveys for attendees to evaluate the usefulness of the event in 'demystifying' careers in health data science. This will allow us to quantitatively judge the effectiveness of the event and will include questions such as: Do you have a stronger direction for where you want to go career-wise?; Do you have a clearer idea for your preference betwee industry/academia routes?; Do you now have an understanding for the differences between working within start-ups compared to more established businesses?*

We expect industry partners and employers to gain a deeper understanding about the wants/needs for health data scientists entering employment as well as receiving an opportunity to advertise their businesses and opening roles

- *This could be followed-up in later years by looking at graduates in the alumni network and where they progress to. This will allow us to judge if these events are having an influence on their decisions or destinations*

We expect HDRUK to forge a closer interaction with affiliated partners as well as see benefits in term of alumni interaction and growth

- *Can see this interaction through involvement with education routes as well as increase in employer attendance annually for the event*
- *Through the post-event survey, we can judge the views of the alumni members on how useful the event was, which can also influence the running year-to-year. Importantly, this can give members who are not in the alumni network a better understanding of and more justification become part of it*

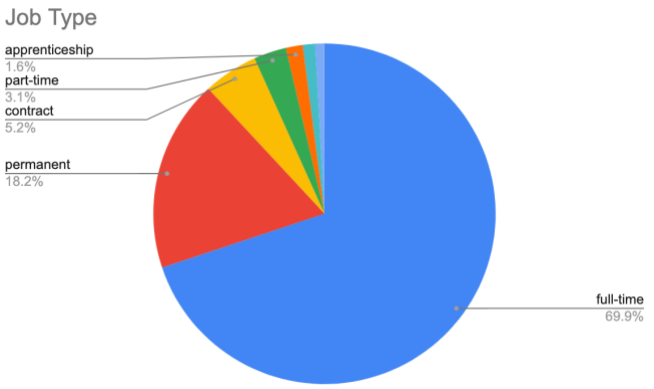
Evidence to Support our Idea

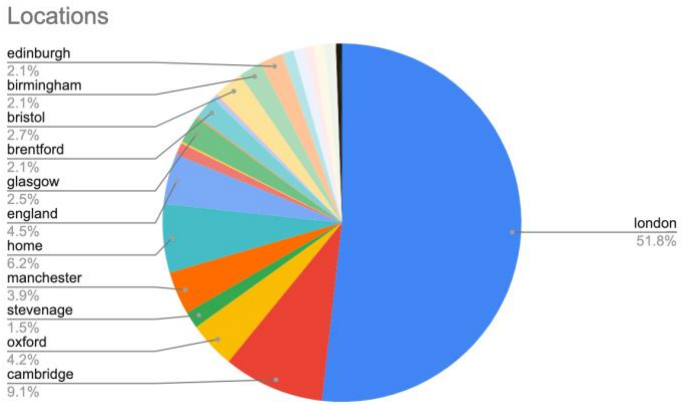
Literature reviews

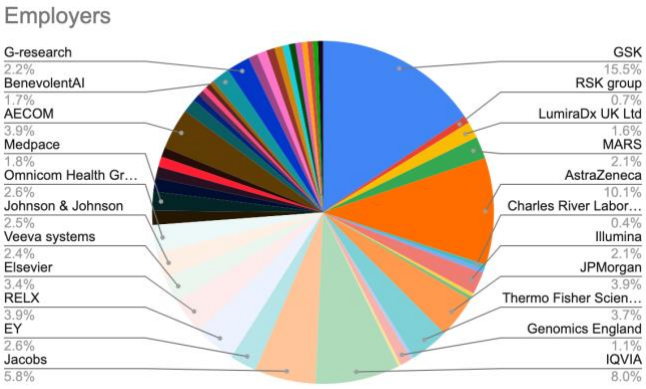
The health data science careers pathways landscape

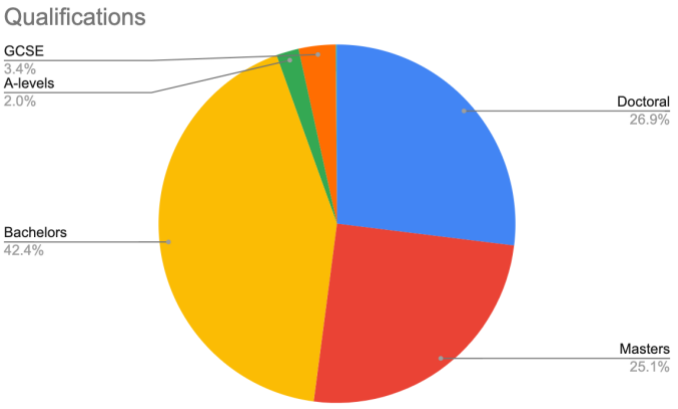
To first gain an overview of the jobs available in the health data science (HDS) space, we searched on the 'indeed' careers website, as this is one of the most popular job listings websites in the UK. We understood that relevant roles may be titled as something other than 'health data scientist', and so in order to ensure a

thorough overview, we performed searches using the following: health data scientist; health data science; medical data science; medical data scientist; machine learning medicine; machine learning biology; and machine learning healthcare. We hoped that this range of searches would encompass a variety of roles. In total, this returned **16054** jobs. It is likely that not all of the results are relevant, and many are overlapping, but this was not so important at this stage, as we wanted a general overview to guide our decisions about the pathways on which to focus. The results are presented below.









Even if these do not accurately represent typical HDS career paths as there are potential biases (are the search terms accurate? Do some employers not advertise on these sites? Are start-ups represented?), we can see some general trends: the majority of jobs are full time; most are based in London; there are a large range of employers, with pharmaceutical giants dominating the landscape; the majority of jobs require university-level qualifications but there is a large representation for both undergraduate and postgraduate training required. We used this as a guide for our further investigations into HDS pathways, to see if these core elements and trends are visible in: previous research; specific career pathway example; and as advertised by employers.

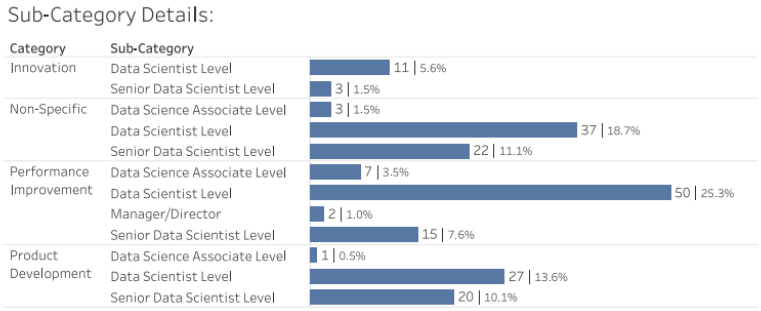
Health data science careers literature analysis

To further understand HDS career pathways, we searched the literature for publications containing the terms "health data scientist" or "health data science" and "careers" or "jobs". There was a lack of relevant research published, likely due to the recent increase in HDS careers. The majority of publications focused on academic HDS pathways (Munevar, 2017) or HDS academic course structure (Aldridge, 2019) with a few statements about HDS careers in industry. The most relevant publication characterised the US HDS job market (Meyer, 2019). The report focused on analysing recent job postings which we believed could be useful for our research. Of particular interest were the author's results to the question "What are the types of data scientist positions for which U.S. healthcare organisations are hiring and what is the focus of the work?" Figure 1 summarises the key skills required for HDS jobs in the US. These skills are broken down into junior and senior data scientist levels. Statistics, R, machine learning, storytelling, and Python were the most common keywords used in job postings.

Data Scientist Skills	Overall (All Job Postings) n = 3218	Overall Percentage Distribution	Data Scientist Level n = 2006	Senior Data Scientist Level n = 1094
Statistics (eg, general linear model, analysis of variance)	138	4%	94	40
R	136	4%	87	44
Applying machine learning techniques	133	4%	85	44
Storytelling; delivering actionable results	132	4%	91	38
Python	125	4%	79	41
Communicating findings	117	4%	79	36
Developing products	117	4%	75	39
Data-driven problem solving	112	3%	71	39
Data manipulation	108	3%	70	34
Developing algorithms	106	3%	66	34
Setting up/maintaining data platforms	97	3%	65	28
SQL	95	3%	63	27
Implementing models into production	89	3%	54	33
SAS	84	3%	54	27
Work in multidisciplinary teams	84	3%	55	28
Creating visualizations	73	2%	45	25
Identifying business problems to address	64	2%	38	23
Big and Distributed Data	61	2%	34	24
Hadoop	61	2%	25	34
Unstructured Data (eg, noSQL, text mining)	56	2%	34	20
Other	1230	38%	742	436
TOTAL	3218	100%	2006	1094

Figure 1: The range of data science skills mentioned in US HDS job postings (Meyer, 2019).

Additionally, the author defined sub-categories of roles within HDS job postings: performance improvers; product developers; modellers; and innovators (Figure 2). Performance improvers worked on areas such as quality measures, financial performance, and patient outcomes. Product developers focused on areas including population health, digital health, decision support, speech/language solutions, behavioural health, and claims analytics. Modeller job postings were nonspecific and required core data science skills, often focusing on machine learning. Innovators addressed areas such as health standards, personalised or precision medicine, genomics, and biology. These could offer us a useful way to further break down the HDS careers within the UK when defining our career pathways.



Degree Required		
Advanced Degree	8	4.0%
Bachelor's and Above	82	41.4%
Master's and Above	76	38.4%
PhD and MD or Equivalent	18	9.1%
Not Listed	14	7.1%
Total	198	100.0%

Figure 2: Sub-categories and degree required found within HDS job postings. Each sub-category has different degree requirements which are described in more detail in the report. (Meyer, 2019).

Different skills were required within each category. 'Product developer' and 'modeller' careers require more data science domain expertise and can apply those skills to a wider range of areas. On the other end of the spectrum were the innovators, who had more training in research and science, particularly precision medicine. 5–7 years of experience were required in more than half of these positions, however in every category there were roles open to those with 0 years of experience. The top hiring organisations included biotech, vendors, and recruiters. Due to the lack of a similar analysis for the UK, we believe a similar publication based on an analysis of UK HDS job postings would be useful.

Linkedin individual career pathways

To have a better understanding of pathways for health data scientists, we performed case studies using Linkedin user profiles. We searched for keywords including "data scientist", "healthcare", "statistician" and "bioinformatics" in job titles, and searched for healthcare institutions and companies that are well-known in the UK. After searching for more than 50 profiles, we picked two examples as listed below. They are not comprehensive due to the diversity of HDS pathways, but they represent two typical pathways. Other pathways include staying in academia throughout one's career, or creating a start-up, but these are less common.

Example 1: Academic route (PhD + Postdoc + Research Associate, diverting to industry.

Job title	Education	Previous career	Listed Skills
Senior Data Scientist in Bioinformatics at a pharmaceutical company	- Master's degree in Chemistry - PhD in Bioinformatics and Molecular Biology	- Postdoctoral Fellow at a bioinformatics institute (1 year) - Research Associate at a cancer research institute (4 years)	- Healthcare related: Bioinformatics, Genomics, Computational Biology, Molecular Biology, Biochemistry, etc - Programming/Data Science related: R, Python, Unix, Data Management, etc

Example 2: Industrial route, with applied HDS jobs. A PhD is not necessarily required.

Job title	Education	Previous career	Listed Skills
Data Scientist at a pharmaceutical company	- Bachelor's degree in Chemistry - Master's (research) in Drug Discovery and Development - Master's (taught) in Health Data Science	- Data Science Graduate Scientist at current company (2 years 1 month) - a 4-month internship in HDS company	- Healthcare related: pharmaceutical industry, UV/Vis spectroscopy, HPLC, NMR, medicinal, chemistry, etc - Programming/data science related: machine learning, R, Python, SQL, STATA, SAS, MongoDB, PyTorch, etc

From these 50+ profiles, the following can be summarised:

- Most of them have acquired at least a Master's degree. A PhD degree is common. These degrees often have an interdisciplinary nature, where healthcare and data science are both important components.
- HDS professionals are usually proficient in multiple programming languages and data science skills.
- Knowledge in at least one healthcare domain is usually present in their education history.
- Internship experience at relevant healthcare companies or institutions is common when a more industrial or applied pathway is taken.

Employers' perspectives

We then searched for companies that facilitated jobs within HDS or offered transitional routes from academia, as well as recruitment agencies specific to HDS.

In terms of **career transitions from academia into industry**, there were a few schemes set up to allow for this.

- *Faculty AI <link>* provides a fellowship programme for new graduates/post-doctoral researchers within STEM based fields to fast track careers within data science. They have a specific health & life sciences branch that has connections with companies such as: IQVIA, NHS England, BHF,

Breast cancer now. These allow for 8 week placements within their customer companies with the main goal being employment post placement.

- *AstraZeneca* <*link*> provides a data sciences & AI graduate programme that allows for three 8 month placement periods giving new graduates a chance to be involved in a multitude of projects.
- *NHS* <*link*> provides a NHS Graduate Digital, Data & Technology scheme that offers 3 distinct pathways. The most relevant to our event would be the Data arm in which they employ data analysts and information analysts. The scheme lasts for 2 years with the end goal being employment within the NHS in your desired position
- *Storm3* <*link*> is a recruitment agency specifically for the HealthTec arena. They have 2 specific consultants for data & analytics as well as other areas within HealthTec i.e. engineering & devops, sales & marketing, product management. They recruit for companies all over the world.
- *WeEngage* <*link*> is another recruitment agency specifically for HealthTec with a specific specialism within data & analytics.

In terms of pursuing **academia fellowships & research** as a career role, there are a few national institutes that try and support this.

- *NIHR Fellowships* <*link*> offer a range of fellowship opportunities for people in different stages of their career. Most relevant to our talks are HDS opportunities for post-doctoral students. This falls under the advanced fellowship options.
- *HDRUK Fellowships* <*link*> offer specific post-doctoral fellowship opportunities exclusively to HDS.
- *BHF Fellowships* <*link*> offer post-doctoral fellowship opportunities for an array of audiences e.g. clinical/non-clinical. For our event we would be focusing on data science related awards such as: BHF-Turing Cardiovascular Data Science Awards, Cardiovascular Catalyst Awards.
- *Wellcome Fellowships* <*link*> offer a vast array of awards and fellowships for different career stages. The one at most interest to us would be the Early-Career Awards that fund people for health related disciplines up to 3 years after their PhD.
- Various other charitable foundations offer fellowships e.g. Cancer Research UK etc - Plenty of fellowship opportunities for people who want to pursue research as a career.

Interviews

We had an interview with Sarah Cadman (Director of Talent and Training, HDRUK) to test the feasibility of our proposal. We received positive feedback which motivated us to continue with our idea. We received advice on how to implement such an event, such as details about different venues, locations to consider and funding options available.

The Evolution of our Extravaganza

How did ideas change after mentoring? What assumptions were challenged?

Our ideas for the event did change after mentoring. However, most of the ideas that changed were surrounding the delivery of our event instead of the actual content/substance of our event. Our mentors gave us positive feedback on the content which we were planning to deliver, they were more inquisitive with our plans for areas such as funding, logistics and inclusivity.

The first suggestion to change in our plan was the naming of the event, while this seems like a minor adjustment, it is a worthwhile change to make. The name of the event gives overall impressions to the attendees as well as the companies applying to participate. The key unique element of our idea is that companies will be giving an honest, open account of a health data scientist's role within those companies as well as the companies gaining relevant exposure and advertising their roles. As such, framing this as a careers event will perhaps give the wrong impression to people as they will picture a more typical careers event with networking etc. We want ours to be distinctly different and a change in name will reflect that.

Issues surrounding inclusivity are ones that we overlooked initially and are very important to ensure a wide demographic of attendees. Not only this, inclusivity issues also link to new ideas surrounding the logistics of our event. We were challenged on issues such as physical disabilities and how these may impact venue choices, as well as issues surrounding neurodivergence and making the event as inclusive as possible for these situations. Furthermore, some logistic choices may also influence inclusivity for example, taking our careers event to be a hybrid model would limit students who cannot attend in person but also do not have access to IT equipment to attend virtually.

Finally, there were many questions surrounding our funding model. Different models were suggested such as: sponsoring, individual business contributions, attendee contributions as well as combinations. All of these have positives and negatives but ultimately attendee contributions would be ruled out as it would just further amplify issues surrounding inclusivity. Equal business contributions would negatively impact small start-ups opposed to big established companies, and is also an infeasible model for non-traditional companies such as charities. It could be adapted to be weighted by company revenue for example to remove this issue but it may provide a barrier for bigger companies to attend.

What follow-up investigations did we conduct?

We were given useful advice into companies that had successfully run career events, such as FacultyAI. Whilst we had conducted research into some of these companies previously, we further investigated the career events of these companies.

What had we not considered?

Our overview of our proposal had focused on explaining our research to support our idea for the event, and therefore we had not thoroughly presented our concrete plan as to "how" the event could take place with our mentors.. This is something our mentors suggested that we more thoroughly communicate in further project proposals.

Our Final Solution

What changed? Is the idea viable? What outcomes would or would not be met? What assumptions were correct or not correct?

The key components of our plan that have changed - some of these being under review - are areas surrounding inclusivity, funding and event logistics. We believe our idea is viable, and is something that could be seen as a natural extension to HDRUK's first career event on 21st June, provided we review potential funding models for our event. We need a clear plan in place for how enough money will be generated for the hosting of the event, as well as potential travel supplements for attendees.

We believe all of our outcomes are still achievable as none of these have been majorly impacted by potential project changes.

References

1. UCL AI In Medicine - https://uclmed.tech/ai-in-medicine/
2. Responsible AI Series From Institute of Engineering and Technology - https://ai-systems.theiet.org/?utm_source=twitter&utm_medium=social&utm_campaign=ai&utm_content=twitter
3. Artificial Intelligence in Genomic Medicine From The Royal Society of Medicine - https://www.rsm.ac.uk/events/medical-genetics/2021-22/mgq50/
4. Translating Artificial Intelligence into Clinical Practise From The Christabel Pankhurst Institute - https://www.eventbrite.co.uk/e/translating-artificial-intelligence-into-clinical-practice-tickets-157236303015?aff=estw&utm-campaign=social&utm-content=attendeeshare&utm-medium=discovery&utm-source=tw&utm-term=listing%20@EventbriteUK
5. AI and Data Science in the Age of Covid-19 - https://www.turing.ac.uk/events/ai-and-data-science-age-covid-19
6. Cambridge Centre for AI in Medicine Inaugural Event - https://ccaim.cam.ac.uk/ccaims-inaugural-online-event-join-us/
7. University of Cambridge Career Fairs (Specifically STEM) - https://unicamcareers.edublogs.org/2020/08/20/virtual-careers-fairs/
8. Imperial College London Data and Analytics Career Fair - https://www.imperial.ac.uk/careers/whats-on/careers-fairs/data/