
Participant pack

The Alan Turing Institute

Data Study Group

9-13 Decemebr 2019

Welcome

Welcome to the Data Study Group, a week-long event bringing together some of the country's top talent from data science, artificial intelligence, and wider fields, to analyse real-world data science challenges.

This participant pack provides you with essential information for the week, as well as preparatory material before arrival. We look forward to welcoming and working with you.

The Alan Turing Institute is the national centre for data science and artificial intelligence, founded in 2015, with the mission to make great leaps in data science and artificial intelligence research to change the world for the better.

We break down disciplinary boundaries; at the Turing, computer scientists, engineers, statisticians, mathematicians, and social scientists under one shared goal, with no departmental boundaries. We are a collaborative hub, with roots in universities and centres of research excellence across the UK, and strong links to a growing network of industry, public sector, and third sector partners.

Our researchers are motivated by driving impact, both through theoretical development and application to real-world problems.

Table of contents

Welcome	2
Table of contents.....	3
Introduction to the Data Study Group	4
Schedule	5
Challenges	6
Outputs and outcomes	6
Facilitators	7
The week.....	8
Code of conduct	9
Data policy & confidentiality	10
Accessing data and IT resources.....	10
Photography/video	11
Social media.....	11
Housekeeping, health & safety	12
Accommodation	12
Accessing The Alan Turing Institute	13
Food/drink	14
Travel expense claims	14
Contacts	15
Emergencies	15
First-aiders and fire marshals.....	16
Executive team.....	17

Introduction to the Data Study Group

This week-long event brings together researchers and industry to work on real data science and AI challenges posed by collaborating organisations. These Challenge Owners (CO) present a real-world challenge to the participants ranging from postgraduate students to faculty members. Participants then choose which challenge they would like to work on over the week and split into groups to start designing solutions, co-ordinated by a facilitator. During the week, each group collaboratively writes a report summarising the results obtained, and each group presents the results on Friday morning to all participants and CO's.

The Turing Data Study Groups are organised by a team of Turing researchers, students, and business team members who will be present throughout the event to ensure everything runs smoothly.

Schedule

This is an approximate schedule for the week (final schedule available on arrival)

Monday	Tuesday	Wednesday	Thursday	Friday
9:30 Registration	9:00 Start	9:00 Start	9:00 Start	8.30 – 9:30 Coffee and pastries
10:00 Presentations				10:00 Presentations
12:30 Lunch	13:00 Lunch	13:00 Lunch	13:00 Lunch	14:00 Lunch and close
17:30 Social offsite and dinner *	19:00 Dinner	18:00 Dinner	18:00 Dinner	Optional socialising / networking
	Finish 21:00	Finish 20:00	Finish 22:45	

Registration will begin at 9:30 on Monday 9th December with presentations starting at 10:00.
It is vital you arrive on time so may be fully briefed on all challenges.

Please remember to bring some photo ID with you.

On Friday 13th December, there is an optional 8:30 start for those who would like to rehearse their presentations. Please checkout of the accommodation before coming to the Turing.

Presentations will begin at 10:00 and the Data Study Group week will close with lunch, ending by 14:00. We anticipate participants might want to spend Friday afternoon socialising, networking, and celebrating the completion of their projects.

The institute will also be taking part in National Festive Jumper day on Friday 13th December, if you would like to take part please wear your finest festive jumper and bring along a £2 charity donation.

* Please note, the off-site social will require travelling on the London transport network. You will need to cover the immediate cost of your travel either using an Oyster card, contactless card or a paper ticket. To be reimbursed for this please write 'social travel' on your travel expenses form at the end of the Data Study Group (this is the only expense that will not require proof of receipt).

Challenges

- **WWF**- Smart monitoring for conservation areas.
- **Dstl**- Anthrax and nerve agent detector: Identification of hazardous chemical and biological contamination on surfaces using spectral signatures.
- **Dstl**- Bright-field image segmentation.
- **The National Archives**- Discovering topics and trends in the UK Government Web Archives.
- **Agile Datum**- Automating the evaluation of local government planning applications.
- **SenSat**- Semantic and Instance Segmentation of 3D Point Clouds.

Note: You can only work on one challenge during the Data Study Group week.

Facilitator	Challenge	Principal Researcher
Dan Nielson	WWF	Kasra Hosseini
Laura Merritt	Dstl (Anthrax and nerve agent detector)	Andrew Dowsey
Sam Blakeman	Dstl (Bright-field image segmentation)	Jeremy Pike
Fazl Barez	The National Archives	David Beaven
Adriaan Hilbers	Agile Datum	Jack Roberts
Ondrej Bohdal	SenSat	Wen Xiao

Outputs and outcomes

On Friday, each group will:

- Submit a report of the work undertaken
- Give a 20-minute presentation (followed by questions)

Your group facilitator will be able to provide advice and guidance on the format.

Data Study Group teams have achieved incredible results in previous events. For inspiration, take a look at the [case studies on our website](#).

Facilitators

Facilitate (verb) via [wiktionary](#)

- To make easy or easier
- To help bring about
- To preside over (a meeting, a seminar)

Collaborative work does not come for free. It's hard to effectively manage the communications between individual researchers. This is particularly true when they are from different fields of study, at different points in their research training (from first year PhD students to faculty to professional data scientists and every stage in between) and when you are only together for five days.

The role of each challenge facilitator is to bring out the best in the team members and oversee the interactions within the team. We want delegates to have a good time during their visit to the Turing, to feel that your expertise is appreciated and harnessed and to feel proud of the group's contribution as a whole. Your facilitator will be responsible for ensuring that interpersonal communications are clear, the work conducted is documented and included in the report (even if it is incomplete or doesn't lead to a conclusive result) and that the project is managed in a transparent way.

The best facilitators make it seem like they did nothing at all. Don't be fooled. They are key to the success of a Data Study challenge.

Meet your Facilitators



Fazl Barez - The National Archives

Fazl is a first year PhD student at Edinburgh Centre for Robotics. His research is in unification of symbolic methods and deep learning for explainable AI. Outside studies, Fazl enjoys playing chess and listening to jazz.

Website: http://www.inf.ed.ac.uk/people/students/Fazl_Barez.html



Laura Merritt - Dstl (Anthrax and nerve agent detector: Identification of hazardous chemical and biological contamination on surfaces using spectral signatures

Laura is a 3rd year PhD student with the University of Reading, based at the UK Centre for Ecology and Hydrology. Her research focuses on improving the fit of dispersal kernels to data for predicting species movements under climate change using Bayesian methods and alternative cost metrics. She has also undertaken an internship at the National Physical Laboratory, working on detecting ischemia in heart tissue using deep learning on PET scan time series. Outside of academia Laura enjoys dancing and walking her dog.



Ondrej Bohdal- SenSat

Ondrej is a first-year PhD student in Data Science at the University of Edinburgh. His research focuses on developing new meta-learning methods to learn how to solve few-shot learning problems across different domains. Ondrej studied Artificial Intelligence and Mathematics at Edinburgh for his undergraduate degree and then joined CDT Data Science – first finishing a research Master's degree and currently continuing as a PhD student. He has also done tech internships at Amazon, JPMorgan Chase & Co. and Metaswitch. In his free time, he often participates in hackathons and datathons, plays board games, does Jiu Jitsu or swims.



Adriaan Hilbers – Agile Datum

Adriaan is a second-year PhD student in statistics and machine learning applied to renewable energy planning at Imperial College London. Originally from the Netherlands, he has a bachelor's degree in Mathematics and a master's degree in Mathematical and Theoretical Physics, both from the University of Oxford. Between his master's degree and start of PhD, he spent around a year working in strategy consulting in Amsterdam. Next to his studies, he is the Director of Content Creation for the Energy Journal, a biannual energy magazine released in collaboration between Imperial College, UCL, LSE and ESCP.



Dan Nielsen – WWF

Dan is a 4th year PhD student in Mathematical Logic at the University of Bristol. He is currently doing an internship at the tech start-up Barbal, where he is applying deep learning methods to solve NLP problems related to technical documents. Dan also facilitated the data science group at a recent Bristol-Bath hackathon, which ended up winning a prize! In his spare time you'll often find him in a yoga class.



Sam Blakeman – Dstl (Bright-field image segmentation)

Sam is a final year PhD student studying Computational Neuroscience at Birkbeck, University of London. His PhD project focuses on improving the transfer ability of reinforcement learning methods using inspiration from different memory systems in the brain. Before his PhD Sam completed a bachelors degree in Natural Sciences and a masters degree in Neuroscience. He also recently completed the Data Science for Social Good program at Imperial College London, which has inspired me to do more applied machine learning research particularly in the fields of bioscience and environmental science. Outside of work he enjoys triathlons and is always keen to show people new places in London to swim, bike or run!

The facilitators are supported by **the meta-facilitators** – Catherine, Jules, Alvaro and Marya from the Data Study Group executive. If you have any issues with your facilitator, or just want to talk to someone different, do come over and see them.



The week

Code of conduct

The Alan Turing Institute Data Study Group (DSG) is an inclusive event where people should feel comfortable sharing their work, opinions, and perspectives. All of us commit to engaging with each other mindfully to ensure an environment that promotes shared learning and collaboration.

We are dedicated to a harassment-free experience for everyone, regardless of gender, gender identity and expression, sexual orientation, disability, physical appearance, body size, race, age, religion, choice of text editor, or the number of tabs open in a browser window.

We are confident you will come together to build a supportive and collaborative atmosphere during the week. The following bullet points set out explicitly what we hope you will consider to be appropriate community guidelines:

- **Be respectful to others.** Do not engage in homophobic, racist, transphobic, ageist, ableist, sexist, or otherwise exclusionary behaviour.
- **Use welcoming and inclusive language.** Exclusionary comments or jokes, threats or violent language are not acceptable. Do not address others in an angry, intimidating, or demeaning manner. Be considerate of the ways the words you choose may impact others. Be patient and respectful of the fact that English is a second (or third or fourth) language for some participants.
- **Do not harass people.** Harassment includes unwanted physical contact, sexual attention, or repeated social contact. Know that consent is explicit, conscious and continuous—not implied. If you are unsure whether your behaviour towards another person is welcome, ask them. If someone tells you to stop, do so.
- **Respect the privacy and safety of others.** Do not take photographs of others without their permission. Note that posting (or threatening to post) personally identifying information of others without their consent ("doxing") is a form of harassment.
- **Be considerate of others' participation.** Everyone should have an opportunity to be heard. In update sessions, please keep comments succinct in order to allow maximum engagement by all participants. Do not interrupt others on the basis of disagreement; hold such comments until they have finished speaking.
- **Don't be a bystander.** If you see something inappropriate happening, speak up. If you don't feel comfortable intervening but feel someone should, feel free to ask a member of the executive team for support.

As an overriding rule, please be intentional in your actions and humble in your mistakes.

This Code of Conduct governs participation for all interactions related to the Data Study Group week. **It applies to all participants during every session, including after-hours working and social events.**

The full Code of conduct and further information on how to report an issue can be found online: https://docs.google.com/document/d/1c8cWCFd2LM27jEdDYbS7YL0-zqX_7AE5-1Vf7z8nEOA/edit?usp=sharing

This Code of conduct is not intended as a static set of rules by which everyone must abide. Rather, you are invited to make suggestions for updates or clarifications within the google document.

The Code of conduct is adapted (lightly) from [Citizen Lab Summer Institute 2017 Code of Conduct](#) and used under a [CC BY 2.5 CA license](#).

Data policy & confidentiality

Some of the data available to groups contains sensitive or personal information. Adhering to the signed Terms and Conditions is therefore of utmost importance.

To reiterate: it is acknowledged you may have access to confidential information over the course of the Data Study Group. You undertake that you will not, either during the term of this agreement or at any time thereafter, use or disclose to any person, firm or company any confidential information.

Accessing data and IT resources

Bring your:

- Laptop
- Mobile phone

Projects will be hosted on the Microsoft Azure platform. We will be using two step factor authentication. **Therefore, you will need access to your mobile phone.**

Whenever you log in, you will be asked to confirm access. This will either be in the form of a telephone call or you can use the Microsoft Authenticator app (highly recommended to speed up your work) available from most app stores.

Internal communications - Slack

For internal communications, participants will be invited to the Data Study Group [Slack](#) channel. Due to the confidentiality of some of the challenges, **it is imperative that no discussions on the actual data itself is conducted on the Slack channel.** This will be policed, and we will ask you to delete any posts that refer to the challenge data.

Please use the Slack channel to organise in your groups, ask the executive team for additional help, IT support, sharing resources, and to ask any big data science questions that may stump you in your challenge quest. In the spirit of collaboration everyone involved is technical support for data science.

One of our goals for the Slack channel is to build an ongoing Turing DSG participant community which continues after the event. As such, be aware that previous participants can be found in the Slack channel. We hope you would like to continue to keep in contact with us and each other. You can leave the channel at any time.

Photography/video

Please note photos will be taken and video will be recorded during the week for Turing and partner communications and marketing purposes. We find that many participants are proud and excited to taking part in the Data Study Group and the Turing Communications team enjoy capturing and sharing moments from the week as it evolves. The team - and potentially partners - will be actively gathering testimonials from participants during and after the event and you are encouraged to share your experiences. Please let us know if you are interested in being interviewed!

If you do not wish to be photographed, please make staff aware.

Presentations taking place on Friday will be recorded and may be live streamed for industry representatives unable to attend in person. Recordings will be kept for internal use by Turing and the Challenge Owners and not published following the event.

Social media

It is natural to want to tell your friends, family, and colleagues what you are up to, and we welcome you joining us on social media to share your experiences over the week. However, due to the sensitive nature of some of the challenges and of the data you will be working with, we ask for caution on what you share on social media in relation to the challenges or the datasets you are working with.

Specifically, under no circumstances post pictures or screenshots, in any form, of any of the data provided. In some circumstances, this could be considered a criminal offence, and you would be breaking the terms & conditions you have digitally accepted.

If in doubt about what is appropriate to post, please speak to a Data Study Group facilitator or a member of the Turing's communications team (on-site throughout the week).

The Turing can be found on Twitter @turinginst and the hashtag for the Data Study Group is #TuringDSG.

Follow us on Turing Social Media

[Youtube](#)

[Facebook](#)

[LinkedIn](#)

Send us your photos, videos, comments, quotes, updates on how projects are progressing etc, we would love to share them!

Email: communications@turing.ac.uk

Housekeeping, health & safety

Accommodation

Accommodation has been booked for those who required it, at the Hub by Premier Inn Kings Cross:

50 Wharf Road,
Kings Cross,
London,
N1 9FA

Check in time is from 2pm and check out is no later than 12pm. The Hotel may reserve the right to charge an additional night for late check outs.

If you arrive earlier than 2pm, you may store your luggage free of charge at the hotel or bring it with you to the Institute. On departure date you can also store luggage if leaving later than 12pm, however, luggage cannot be left overnight.

Upon arrival you may be asked for a payment card, this will not be charged unless you are late to check out or incur any other expenses during your stay. Breakfast is included and is free of charge.

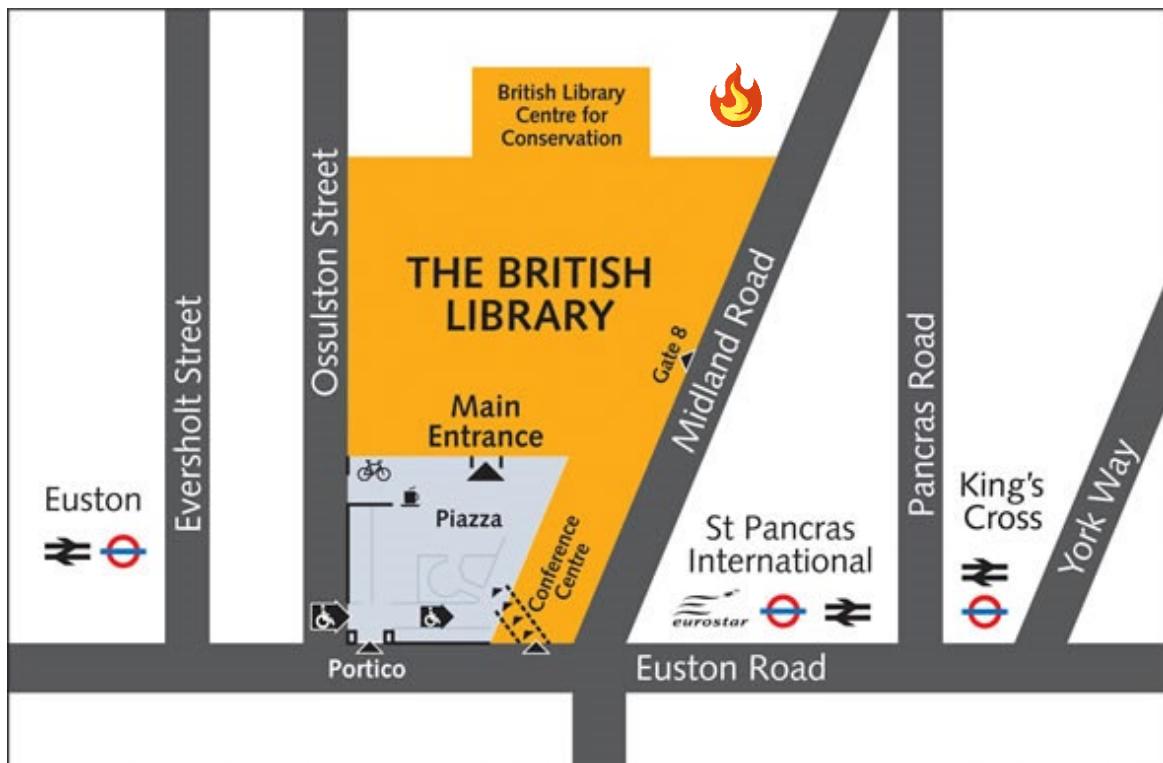
Email: **datastudygroup@turing.ac.uk** with questions regarding accommodation.

Accessing The Alan Turing Institute

The Data Study Group will be held at The Alan Turing Institute, which is a bespoke space on the first floor of the British Library.

Due to the early starts/late finishes you will need to enter and exit through the staff entrance – Gate 8 on Midland Road. You will need to sign in when you arrive at the Turing office itself each day. If you wish to leave the building during the day and return, you will need to sign in and out at the Turing reception due to fire safety monitoring and use the main library entrance.

On arrival on Monday, you will be issued with an event badge. **You must wear this at all times, and when entering and leaving the building.**



During the day, enter and exit through the main British Library entrance (9:30-20:00). You may need to queue to enter the building and be subject to a baggage check as part of the Library's standard security checks.

You can find the Turing offices on the first floor of the British Library, next to the Business & IP Centre. We have our own entrance.

Do check out the Enigma machine just outside the entrance too!



The Alan Turing Institute

British Library
96 Euston Road
London NW1 2DB

Food/drink

Lunch and dinner will be provided throughout the week. Dietary requirements communicated in advance will be catered for wherever possible. All meals will come with a vegan options.

The caterers cannot provide guaranteed allergen free food; whilst every effort is made to ensure allergens are catered for, traces may be found as ingredients may appear in other recipes. Therefore, we recommend that participants with serious allergies speak to a member of the Data Study Group team so that alternative arrangements can be made.

Hot/cold drinks and fruit are available in the kitchen.

Please bring a bottle for water, the Institute will not have enough glasses for everyone.

Travel expense claims

The most economic and low-carbon means of travel should be chosen wherever practical and claimants should use existing travel cards and discounts wherever possible to reduce costs.

To claim travel expenses, submit the Turing expenses and compliance forms which has been emailed to you along with this pack.

Complete **both** forms with copies of receipts and/or tickets for each item. Payments will not be processed without receipts. If your bank account is not located in the UK, include IBAN and swift code.

If using a Transport for London Oyster Card, ensure you have registered it online at www.tfl.gov.uk before you travel. Once registered, you can export your travel record, and submit as evidence for travel expense claims. A receipt for an Oyster card 'top-up' will not be accepted as evidence for your expense claim.

Contacts

A facilitator will be assigned to each group and will be your main point of contact for any issues. In addition, you can speak to the DSG team at any time.

Catherine Lawrence (Programme Manager) – 07508 071 851

Jules Manser (DSG coordinator) - 0203 862 3312

Daisy Parry (DSG Administrator) - 07948850167

You can also speak to the events, communications or facilities team at any time.

Emergencies

The weekly fire drill is on Monday morning around 09:15 and is preceded with an announcement. At all other times, if the alarm goes off, you will need to evacuate the British Library. In the event of an evacuation, please exit the building via the Turing main entrance, go down the stairs and exit the British Library also by its main entrance, and proceed to the Assembly Point at the rear courtyard of the British Library, on Midland Road. Your facilitator will be able to instruct you.

The Assembly Point is marked with  on the building access map on page 16.

Requests for emergency support (ambulances, paramedics etc.) must be made via Security Control Room 0207 412 7007. Please notify your facilitator or a member of the DSG team for assistance.

First-aiders and fire marshals

			
Sharmila 	Kate 	Daisy 	Jade 
			
Amerik 	Anneca 	Jessie 	Jules 

Executive team



Sebastian Vollmer – Data Study Groups Director

Sebastian created the Data Study Groups with the aim of fostering challenge-driven research and translational activities as an inclusive community effort. Sebastian is a Turing Fellow, an associate Director on the Health programme at the Turing, and an Associate Professor for Mathematics and Statistics at University of Warwick. His work has two pillars: foundations of Bayesian inference and Monte Carlo methods and applications of data science. He enjoys his free time with his family and likes board games and discussing politics.



Marya Bazzi – Principle Researcher

Marya is a Turing Research Fellow. Before joining the Turing, Marya was a postdoctoral scholar at Oxford University and Head of Analytics at a London-based FinTech. Marya is particularly interested in developing research projects at the interface between academia and industry.



Alex Bird – Participant recruitment and DSG advisor

Alex is a PhD student at the Turing and has acted as a Facilitator at previous Data Study Groups. Alex leads on the participant recruitment and selection process for the DSG team. Alex has been a member of the DSG team since its inception and provides advice and guidance in running the DSG, as well as the week itself. His research interests lie primarily in time series models and approximate inference, currently working with applications to medical monitoring and drug dosing.

“Data Study Groups function as a democratic algorithm for boosting people. And they’re also good fun”.



Alvaro Cabrejas Egea – Science Liaison

Alvaro is an Enrichment student at The Alan Turing Institute. His research areas are reinforcement learning and traffic prediction and control. He is interested in economic and agent games in networks, machine learning and time series analysis.

“The Turing DSG is an extraordinary opportunity to apply your skills while experiencing the current challenges that industry faces, all in a fun interdisciplinary and collaborative environment.”



Helen Hu – Science Liaison

Helen is PhD student in Statistics at UCL and Turing. Her research is on handling not at random missing data and misclassification.

“The DSG is a great opportunity to work on interesting small projects and learn about how other people approach problems differently.”



Catherine Lawrence – Senior Programme manager

Catherine is a member of the Partnerships team at the Turing, and works with Jules and the rest of the DSG team to deliver the DSGs. Catherine has just joined the Health programme at Turing, moving over from the Defence & Security and Data Science at Scale programme.

“The Data Study Groups are all about open collaboration; it’s brilliant to see participants sharing skills and learning from each other too”



Jules Manser - Data Study Group operations management

Jules is a member of the Partnerships team at the Turing and has played a lead role in delivering Data Study Groups working closely with Sebastian since their inception in 2016.

“The DSG is a great way to test your skills against real problems. But you should also have fun!”



Bilal Mateen – Clinical and Bio-sciences liaison

Bilal is a doctor at Kings College Hospital, and a data science oriented academic-clinician who works on medical applications of machine learning at Warwick Medical School, The Alan Turing Institute, as well as several other institutions.

“Every time I’ve participated in a DSG Challenge, it’s those who know least about the background subject matter that help make the amazing intuitive leaps that underlie the best solutions”



Daisy Parry- Data Study Group administrator

Daisy has recently joined The Alan Turing Institute and works closely with Jules to deliver the Data Study Group events.