

# Contents

2	<b>1 Summary</b>	<b>3</b>
3	<b>2 Core team</b>	<b>4</b>
4	<b>3 Application questions</b>	<b>5</b>
5	3.1 BBSRC schemes . . . . .	5
6	3.2 BBSRC remit classification . . . . .	7
7	3.3 Vision . . . . .	8
8	3.3.1 Context . . . . .	8
9	3.3.2 Specific aims . . . . .	11
10	3.4 Approach . . . . .	15
11	3.4.1 Offline analysis methods . . . . .	16
12	3.4.1.1 Initial list of methods to include in the library	16
13	3.4.1.2 Non-stationarity . . . . .	18
14	3.4.1.3 Computational efficiency . . . . .	21
15	3.4.1.4 Deliverables . . . . .	22
16	3.4.2 Real-time machine learning methods . . . . .	22
17	3.4.2.1 Real-time machine learning in neuroscience . .	22
18	3.4.2.2 Bonsai and Bonsai.ML . . . . .	24
19	3.4.2.3 Deliverables . . . . .	24
20	3.4.3 Visual exploration . . . . .	25
21	3.4.3.1 Continuous visualizations . . . . .	25
22	3.4.3.2 Epoched and interactive visual analytics . . .	26
23	3.4.3.3 Software stack for interactive visualizations .	27
24	3.4.3.4 Deliverables . . . . .	29
25	3.4.4 Offline spike sorting . . . . .	29
26	3.4.4.1 Outputs . . . . .	29
27	3.4.5 Online spike sorting . . . . .	29
28	3.4.5.1 Outputs . . . . .	29
29	3.5 Applicant and team capability to deliver . . . . .	30
30	3.6 Project partners . . . . .	32
31	3.7 Project partners: statement of support . . . . .	33
32	3.8 Trusted research and innovation (TR&I) . . . . .	34
33	3.9 Resources and cost justification . . . . .	35
34	3.10 Data management and sharing . . . . .	36

35	3.11 Facilities . . . . .	37
36	3.12 Ethics and responsible research and innovation (RRI) . . . . .	38
37	3.13 Genetic and biological risk . . . . .	39
38	3.14 Research involving the use of animals . . . . .	40
39	3.15 Conducting research with animals overseas . . . . .	41
40	3.16 Research involving human participation . . . . .	43
41	3.17 Research involving human tissues or biological samples . . . . .	44

## 1 Summary

Word limit: 550

In plain English, provide a summary we can use to identify the most suitable experts to assess your application.

We usually make this summary publicly available on external-facing websites, therefore do not include any confidential or sensitive information. Make it suitable for a variety of readers, for example:

- opinion-formers
- policymakers
- the public
- the wider research community

**Guidance for writing a summary** Clearly describe your proposed work in terms of:

- context
- the challenge the project addresses
- aims and objectives
- potential applications and benefits
- its relevance to the BBSRC long-term research and innovation priorities and, if applicable Responsive Mode Spotlight areas

## 61 **2 Core team**

62 List the key members of your team and assign them roles from the following:

- 63 • project lead (PL)
- 64 • project co-lead (UK) (PcL)
- 65 • specialist
- 66 • professional enabling staff
- 67 • research and innovation associate
- 68 • technician
- 69 • researcher co-lead (RcL)

70 Only list one individual as project lead.

71 UKRI has introduced a new addition to the ‘specialist’ role type. Public  
72 contributors such as people with lived experience can now be added to an  
73 application.

74 Find out more about [UKRI’s core team roles in funding applications and](#)  
75 [our eligibility guidance](#).

## 76 **3 Application questions**

### 77 **3.1 BBSRC schemes**

78 Word limit: 1

79 Indicate the scheme through which you are applying.

80 In the text box, copy the number corresponding to the scheme you are  
81 applying through. These are:

- 82 1. standard (no scheme)
- 83 2. Industrial Partnership Award (IPA)
- 84 3. LINK
- 85 4. Brazil (FAPESP)
- 86 5. Luxembourg (FNR)
- 87 6. NSF-Bio

88 Additional guidance

89 This is for administrative purposes to help the initial application process-  
90 ing.

91 Please follow the scheme specific guidance below and upload the addi-  
92 tional documents listed as a single PDF no larger than 8MB:

93 IPA or LINK:

- 94 • a letter from your institution's technology transfer office outlining the  
95 management of outputs from the proposed research

96 FAPESP:

- 97 • FAPESP proposal form
- 98 • FAPESP consolidated budget form
- 99 • FAPESP letter of eligibility

100 FNR:

- 101      • CVs of international collaborators
- 102      • FNR ‘INTER’ budget form
- 103      • FNR ‘INTER’ cost justification
- 104      NSF-Bio:
- 105      • US biosketches
- 106      • US budget forms

## 107 **3.2 BBSRC remit classification**

108 Word limit: 1

109 Your application will be considered by one of our four research committees  
110 made up of independent experts. Indicate which you feel would be best placed  
111 to assess your application.

112 In the text box, write only the letter (in uppercase) corresponding to the  
113 committee you feel would be best placed to assess your application. These  
114 are:

115 **A** animal disease, health and welfare

116 **B** plants, microbes, food and sustainability

117 **C** genes, development, and science, technology, engineering and maths (STEM)  
118 approaches to biology

119 **D** molecules, cells and industrial biotechnology

120 Additional guidance:

121 This is for administrative purposes to help the initial application pro-  
122 cessing. We will check your choice and make a final decision as to which  
123 committee will assess your application.

## 124 3.3 Vision

125 Word limit: 550

126 What are you hoping to achieve with your proposed work?

127 What the assessors are looking for in your response

128 Explain how your proposed work:

129 1. is of excellent quality and importance within or beyond the field(s) or  
130 area(s)

131 2. has the potential to advance current understanding, or generate new  
132 knowledge, thinking or discovery within or beyond the field or area

133 3. is timely given current trends, context, and needs

134 4. impacts world-leading research, society, the economy, or the environ-  
135 ment

136 You may demonstrate elements of your responses in visual form if rele-  
137 vant. Further details are provided in the Funding Service. References may  
138 be included within this section.

### 139 3.3.1 Context

140 Conventional systems neuroscience experiments are typically short in dura-  
141 tion and often place significant constraints on subjects behaviours to simplify  
142 data analysis. However, these restrictions may limit our ability to observe  
143 critical aspects of brain function and behaviour that only manifest in more  
144 naturalistic and extended conditions.

145 At the Sainsbury Wellcome Centre (SWC) and Gatsby Computational  
146 Neuroscience Unit (GCNU) we are pioneering **Naturalistic, Long-Duration,**  
147 **and Continual (NaLoDuCo) experiments** in mice that span weeks to  
148 months. During these experiments, we collect high-resolution behavioural  
149 and neural recordings in naturalistic settings (Figure 1).

150 To support this endeavor, we are developing the **AEON platform**, an  
151 innovative set of hardware and software tools for NaLoDuCo experimen-  
152 tal control, data store and access. We are using this platform to investi-  
153 gate the neural basis of foraging behavior in mice over prolonged periods of  
154 time (Campagner et al., 2024).



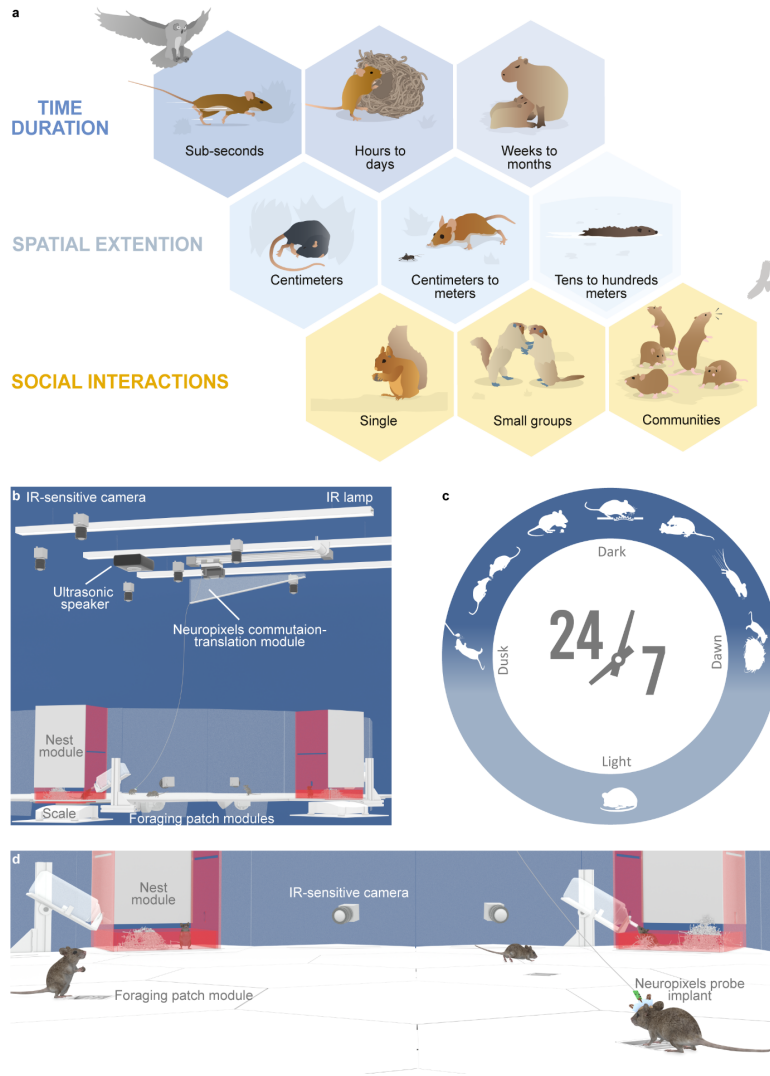


Figure 1: **a**: Example of natural behaviours in rodents that take place over different timescale, spatial extensions and involving different numbers of individuals.

**b-d**: Close-up views of one possible configuration of the Aeon environment in which naïve mice and mice chronically implanted with Neuropixels probe can live while expressing a variety of natural behaviours including exploring, drinking, escaping, foraging, nesting, sleeping, eating and interacting socially.

155 Our US partner, the **Allen Institute for Neural Dynamics (AIND)**  
156 is also performing NaLoDuCo experimentation, using the AEON platform,  
157 studying naturalistic olfactory learning over weeks to month outside conven-  
158 tional task structures (Fink et al., 2024).

159 **NeuroGEARS Ltd**, our industrial partner, is a UK-based company sup-  
160 porting academic institutions implementing innovative technology for scien-  
161 tific investigation. It is the main developer of the **Bonsai** software ecosystem  
162 for experimental control (Lopes et al., 2015), used by thousands of scientists  
163 around the world, and powering the AEON platform. NeuroGEARS has  
164 played a central role in the development of the AEON platform, and pro-  
165 vides services to both the SWC and the AIND.

166 NaLoDuCo experimentation will enable researchers to explore neural mech-  
167 anisms underlying ethological behaviours in naturalistic environments over  
168 months, for the first time. The experiments will shed new light on a wide  
169 range of poorly understood neural mechanisms, including how the brain  
170 structures complex behavioural sequences as a function of the animal needs,  
171 learning, adaptation, sleep-dependent memory consolidation and social dy-  
172 namics. **The data generated from NaLoDuCo experiments represent**  
173 **an entirely new resource in neuroscience**, with the potential to drive  
174 breakthroughs and discoveries that are beyond the reach of traditional ex-  
175 periments.

176 While **naturalistic, long-duration, or continuous** neuroscience ex-  
177 periments have been conducted in the past (Nagy et al., 2023; Ho et al.,  
178 2023; Ray et al., 2025; Weissbrod et al., 2013; Dhawale et al., 2017; Newman  
179 et al., 2024), to the best of our knowledge, **we are the first ones to inte-**  
180 **grate all three of these features in a single experimental paradigm.**

181 This emerging paradigm of long-duration experimentation is poised to  
182 become mainstream in the coming years. However, experiments spanning  
183 weeks to months generate massive datasets—often reaching hundreds of ter-  
184 abytes—posing significant challenges in data acquisition, management, distri-  
185 bution, visualization, and analysis. To address these challenges, we (GCNU,  
186 SWC, AIND, and NeuroGEARS Ltd) will collaboratively extend the AEON  
187 platform with functionality to **visualise and statistically analyze pre-**  
188 **viously collected NaLoDuCo experimental data on the cloud**, and  
189 **to perform real-time machine to enable the intelligent control of**  
190 **NaLoDuCo experiments.**

### 3.3.2 Specific aims

Data generated by NaLoDuCo experiments will be of general interest to the neuroscience community. **We want to share our NaLoDuCo foraging and odor learning recordings and allow other groups collecting this type of data to share their own.** However, this dissemination is not trivial, as datasets are of the order of hundreds of terabytes, and it will take users several days to download them over standard Internet connections.

Instead of bringing data to users, we will bring users to data, by storing datasets in the cloud (or in institutional clusters), and providing **cloud software to allow users to visually explore and statistically analyse behavioural and neural NaLoDuCo datasets where they live** (1 and 2 in Figure 2).

Our statistical analysis of neural time series will require knowledge of the spiking activity of single units; i.e., spike sorting. In long-duration experiments with freely moving animals spike sorting is a challenging problem, because movements of recording probes change the shape of spike waveforms over time and complicate the assignment of spikes to units based on their waveforms. We will address this problem by developing **spike sorting methods for long-duration, continual and high-channel-count recordings** (3 in Figure 2).

Funded by a BBSRC award we are adding machine learning functionality to Bonsai in order to enable a new type of experimentation controlled by advanced machine learning inference on behavioral and neural recordings (Bonsai.ML, [Guilbeault et al., 2025](#)). We have developed this functionality for conventional short duration experiments. We will add to Bonsai.ML **real-time machine learning functionality for processing nonstationary data**, such as that generated in NaLoDuCo experiments.

Most of the online neural data analysis methods that we will add to AEON require sorted spikes. We will adapt the previous offline **spike sorting methods for long-duration experiment to operate in real-time** (5 in Figure 2).

## References

Campagner, D., Bhagat, J., Lopes, G., Calcaterra, L., Ahn, J., Almeida, A., Carvalho, F. J., Cruz, B., Erskine, A., Lo, C., Nguyen, T. T.,

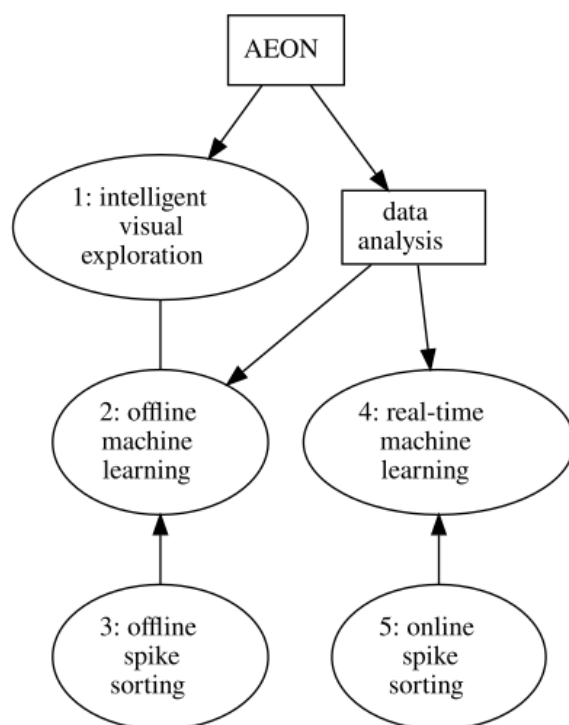


Figure 2: Specific aims

225 Pouget, A., Rapela, J., Ryan, T., Reggiani, J., and Group, S. S.  
226 F. B. W. (2024). Aeon: an open-source platform to study the  
227 neural basis of ethological behaviours over naturalistic timescales.  
228 [https://www.sfn.org/-/media/SfN/Documents/NEW-SfN/Meetings/](https://www.sfn.org/-/media/SfN/Documents/NEW-SfN/Meetings/Neuroscience-2024/Abstracts-and-Sessions/Abstract-PDFs/SFN24_Abstracts-PDF-Posters_SAT_PM.pdf)  
229 [Neuroscience-2024/Abstracts-and-Sessions/Abstract-PDFs/](https://www.sfn.org/-/media/SfN/Documents/NEW-SfN/Meetings/Neuroscience-2024/Abstracts-and-Sessions/Abstract-PDFs/SFN24_Abstracts-PDF-Posters_SAT_PM.pdf)  
230 [SFN24\\_Abstracts-PDF-Posters\\_SAT\\_PM.pdf](https://www.sfn.org/-/media/SfN/Documents/NEW-SfN/Meetings/Neuroscience-2024/Abstracts-and-Sessions/Abstract-PDFs/SFN24_Abstracts-PDF-Posters_SAT_PM.pdf).

231 Dhawale, A. K., Poddar, R., Wolff, S. B., Normand, V. A., Kopelowitz, E.,  
232 and Ölveczky, B. P. (2017). Automated long-term recording and analysis  
233 of neural activity in behaving animals. *Elife*, 6:e27702.

234 Fink, A. J., Hogan, M., and Schoonover, C. E. (2024). Olfactory investigation  
235 in the home cage. *Neurobiology of Learning and Memory*, 213:107951.

236 Guilbeault, N., Lopes, G., and Rapela, J. (2025). Bonsai.ml: machine  
237 learning for intelligent experimental control. [https://bonsai-rx.org/](https://bonsai-rx.org/machinelearning/)  
238 [machinelearning/](https://bonsai-rx.org/machinelearning/).

239 Ho, H., Kejzar, N., Sasaguri, H., Saito, T., Saido, T. C., De Strooper, B.,  
240 Bauza, M., and Krupic, J. (2023). A fully automated home cage for long-  
241 term continuous phenotyping of mouse cognition and behavior. *Cell Re-*  
242 *ports Methods*, 3(7).

243 Lopes, G., Bonacchi, N., Frazão, J., Neto, J. P., Atallah, B. V., Soares, S.,  
244 Moreira, L., Matias, S., Itskov, P. M., Correia, P. A., et al. (2015). Bonsai:  
245 an event-based framework for processing and controlling data streams.  
246 *Frontiers in neuroinformatics*, 9:7.

247 Nagy, M., Naik, H., Kano, F., Carlson, N. V., Koblitz, J. C., Wikelski, M.,  
248 and Couzin, I. D. (2023). Smart-barn: Scalable multimodal arena for real-  
249 time tracking behavior of animals in large numbers. *Science Advances*,  
250 9(35):eadf8068.

251 Newman, J. P., Zhang, J., Cuevas-López, A., Miller, N. J., Honda, T., van der  
252 Goes, M.-S. H., Leighton, A. H., Carvalho, F., Lopes, G., Lakunina, A.,  
253 et al. (2024). Onix: a unified open-source platform for multimodal neural  
254 recording and perturbation during naturalistic behavior. *Nature Methods*,  
255 pages 1–6.

256 Ray, S., Yona, I., Elami, N., Palgi, S., Latimer, K. W., Jacobsen, B., Witter,  
257 M. P., Las, L., and Ulanovsky, N. (2025). Hippocampal coding of identity,

258 sex, hierarchy, and affiliation in a social group of wild fruit bats. *Science*,  
259 387(6733):eadk9385.

260 Weissbrod, A., Shapiro, A., Vasserman, G., Edry, L., Dayan, M., Yitzhaky,  
261 A., Hertzberg, L., Feinerman, O., and Kimchi, T. (2013). Automated  
262 long-term tracking and social behavioural phenotyping of animal colonies  
263 within a semi-natural environment. *Nature communications*, 4(1):2018.

## 264 **3.4 Approach**

265 Word limit: 3,300

266 How are you going to deliver your proposed work?

267 What the assessors are looking for in your response

268 Explain how you have designed your approach so that it:

- 269 1. is effective and appropriate to achieve your objectives
- 270 2. is feasible, and comprehensively identifies any risks to delivery and how  
271 they will be managed
- 272 3. uses a clearly written and transparent methodology (if applicable)
- 273 4. summarises the previous work and describes how this will be built upon  
274 and progressed (if applicable)
- 275 5. will maximise translation of outputs into outcomes and impacts
- 276 6. describes how your, and if applicable your team's, research environment  
277 (in terms of the place and relevance to the project) will contribute to  
278 the success of the work

279 You may demonstrate elements of your responses in visual form if rele-  
280 vant.

281 Please make sure to check sizing and readability of the image using 'read  
282 view' prior to submission. Further details are provided in the Funding Ser-  
283 vice.

284 References may be included within this section.

285 Within the 'Approach' section we also expect you to:

- 286 • provide a detailed and comprehensive project plan including milestones  
287 and timelines in the form of an embedded Gantt chart or similar (please  
288 make sure to check sizing and readability of the image using 'read view'  
289 prior to submission)

290 BBSRC's [action plan for EDI](#) outlines our commitment to removing bar-  
291 riers to participation in our programmes, ensuring investments do not inad-  
292 vertently prevent access or usage by individuals from minority groups, for  
293 example disabled researchers.

294 To this end, applications should identify how accessibility and inclusive-  
295 ness in the widest sense have been incorporated into the design of the project.  
296 For example, you may wish to reference relevant institutional strategies and  
297 policies which support equality, diversity, and inclusion as they relate to ac-  
298 cess to equipment and facilities and indicate how the proposed project has  
299 been designed and will be delivered with broad access in mind.

300 We have collected unprecedented NaLoDuCo datasets at the SWC and  
301 AIND, comprising continuous, multimodal recordings over weeks to months.  
302 However, the scientific value of these massive datasets cannot be fully realized  
303 without robust tools for visual exploration and analysis. To address this need,  
304 we will develop and disseminate an open-source library of methods for the  
305 visualization and analysis of NaLoDuCo experimental data.

306 This library will include tools for both offline and online analysis (Sec-  
307 tions 3.4.1 and ??), visual exploration (Section 3.4.3), and scalable offline  
308 and real-time spike sorting (Sections 3.4.4 and 3.4.5). Together, these meth-  
309 ods will make NaLoDuCo data more accessible, interpretable, and actionable  
310 for the neuroscience community.

### 311 3.4.1 Offline analysis methods

312 **Modern neuroscience lacks robust methods to characterize long-**  
313 **duration and continual time series**, especially in settings where the sta-  
314 tistical properties of the data evolve over time. This limitation present a  
315 methodological gap that must be addressed in order to unlock the scientific  
316 potential of NaLoDuCo experiments.

317 To bridge this gap, we will develop and disseminate a software library  
318 containing new implementations of machine learning methods specifically  
319 tailored to: (1) operate effectively under **non-stationary** conditions, and  
320 (2) scale to **very long time series**.

#### 321 3.4.1.1 Initial list of methods to include in the .library

322 We will initially populate this library with new implementations of methods  
323 already in use at the GCNU, SWC, and AIND to analyze neural and behav-  
324 ioral time series from NaLoDuCo foraging and olfactory learning experiments  
325 in mice. These methods span multiple domains—kinematics, neural dynam-  
326 ics, behavioral state segmentation, forecasting, and joint modeling—and are



327 designed to work together within an integrated analysis pipeline. We de-  
328 scribed these methods below and summarise them in Table 1.

329 **Behavioral Analysis:** The first step in behavioral analysis involves multi-  
330 body-part tracking. For this, we will use **deep learning-based pose esti-**  
331 **mation** methods such as **SLEAP**, which enable accurate and efficient track-  
332 ing of multiple unmarked mice across long recording sessions.

333 From the tracked poses, we will infer continuous kinematic variables using  
334 **linear dynamical systems (LDS)**, including particle filter-based variants  
335 to handle uncertainty and measurement noise. These kinematic features will  
336 be used to infer discrete behavioral states with **Hidden Markov Models**  
337 **(HMMs)**, as implemented in tools such as **MoSeq**.

338 We will relate these inferred states and kinematic variables to foraging-  
339 related outcomes—such as patch-leaving decisions—using both **generalized**  
340 **linear models (GLMs)** and **deep neural networks**. These models will  
341 allow us to capture both interpretable and high-capacity representations of  
342 behavioral decision-making processes.

343 To recover the latent strategies guiding animal behavior, we will apply  
344 **inverse reinforcement learning** methods such as **HIQL**, which estimate  
345 the underlying reward functions and policies based on observed actions.

346 NaLoDuCo recordings uniquely support behavioral forecasting over ex-  
347 tended horizons—ranging from hours to days—far beyond what is feasible  
348 in conventional short-duration experiments. To capitalize on this, we will  
349 apply long-horizon forecasting models using **recurrent neural networks**  
350 **(RNNs)** and **transformer architectures**, which are well-suited to model-  
351 ing long-range temporal dependencies.

352 **Neural Data Analysis:** Analysis of high-density electrophysiology will be-  
353 gin with **latent variable modeling** to reduce the dimensionality of popula-  
354 tion neural recordings. We will use both linear and nonlinear latent dynamics  
355 models, including **svGPFA**, which uses Gaussian processes, and **LFADS**, a  
356 deep generative model based on recurrent neural networks.

357 The resulting low-dimensional trajectories will be used to infer discrete  
358 neural states via **HMMs**, using methods such as **SSM**. For neural activity  
359 forecasting across long durations, we will again employ **RNNs** and **trans-**  
360 **formers**, which can model complex temporal structure in spiking activity.

361 We will also decode the animal’s position from hippocampal spike trains  
362 using **point-process decoders**, enabling the analysis of spatial coding and

363 replay phenomena during naturalistic foraging behavior. We will build on  
364 existing implementations such as [replay\\_trajectory\\_classification](#).

365 **Joint Neural-Behavioral Modeling:** To understand the interactions be-  
366 tween neural dynamics and behavior, we will use models that extract **shared**  
367 **latent representations** from both domains. These models will help reveal  
368 how cognitive and behavioral states are jointly encoded in neural activity.

369 We will adapt [Recognition-Parametrized Models \(RPM\)](#), a Bayesian ap-  
370 proach developed at the GCNU, which infers latent variables that explain  
371 multiple observation streams through highly nonlinear relationships. We  
372 will also use [CEBRA](#), a state-of-the-art contrastive learning framework de-  
373 signed for multimodal representation learning, to discover temporally and  
374 semantically aligned neural-behavioral structure.

### 375 3.4.1.2 Non-stationarity

376 Many conventional methods for analyzing neural and behavioral time series  
377 assume that the underlying data-generating processes are stationary—that  
378 is, their statistical properties remain constant over time. While this assump-  
379 tion may be acceptable in short-duration experiments, it breaks down in  
380 long-duration and continual recordings. In such settings, animals learn and  
381 adapt, their internal states and motivations fluctuate, and their behavior and  
382 physiology are influenced by biological rhythms such as circadian, ultradian,  
383 and infradian cycles. These changes induce non-stationarity in the data,  
384 making models that assume stationarity progressively less reliable or even  
385 obsolete.

386 To address this challenge, we will adapt and develop methods that are  
387 explicitly designed to operate in non-stationary environments. Our approach  
388 draws on techniques from multiple domains, including adaptive signal pro-  
389 cessing, machine learning, and Bayesian inference.

390 **Adaptive Signal Processing.** The field of adaptive signal processing has  
391 produced robust methods for modeling linear systems with time-varying dy-  
392 namics ([Haykin, 2002](#)). The recursive least-squares (RLS) algorithm, for  
393 example, is a well-known adaptation of the ordinary least squares algorithm  
394 that continuously updates model parameters to perform linear regression  
395 under non-stationary conditions. We will use RLS to study time-varying re-  
396 lations between behavioral states, as inferred by hidden Markov models, and

Table 1: Initial data analysis methods to disseminate

<b>Domain</b>	<b>Functionality</b>	<b>Method</b>	<b>Model Type</b>
behaviour	multi-body-part tracking	SLEAP	deep neural network
behaviour	kinematics inference	LDS	linear dynamical system
behaviour	kinematics inference	LDS	particle filter
behaviour	state inference	SSM	hidden Markov model
behaviour	regression		generalized linear model
behaviour	regression		deep neural network
behaviour	policy inference	L(M)V-IQL	reinforcement learning
behaviour	long-duration forecasting		RNN
behaviour	long-duration forecasting		transformers
brain	latents inference	svGPFA	Gaussian processes
brain	latents inference	LFADS	RNN
brain	state inference	SSM	hidden Markov model
brain	long-duration forecasting		RNN
brain	long-duration forecasting		transformers
brain	decoding	NA	point-process decoder
brain & behaviour	latents inference	RPM	Bayesian inference + deep neural network
brain & behaviour	latents inference	CEBRA	contrastive learning

397 foraging visit durations.

398 **Continual Learning.** The field of continual learning develops adaptive  
399 methods for artificial neural networks. In classic continual learning, a model  
400 learns a sequence of discrete, well-defined tasks. But in NaLoDuCo experi-  
401 mentation, as in many real-world settings there are not specific task bound-  
402 aries. So methods that do not require task boundaries are needed. They are  
403 studied by the subfield of task-free continual learning and include online regu-  
404 larization (which constrain the update of relevant weights), experience replay  
405 (which maintain a small, representative buffer of past samples) and ensemble  
406 methods (which combine the predictions of multiple individual models  
407 with, for example, different learning rates). We will use these techniques, for  
408 example, to train pose tracking models on month-long continuous recordings.

409 **Adaptive State-Space Models.** In state-space modeling, the Kalman  
410 filter provides a principled way to handle non-stationary Gaussian processes  
411 with drifting mean and covariance. More flexible approaches are needed when  
412 data exhibit abrupt regime shifts or complex latent dynamics. Switching  
413 state-space models, such as Switching Linear Dynamical Systems (SLDS)  
414 and Switching Hidden Markov Models (sHMMs), model discrete changes in  
415 underlying system dynamics. For nonlinear, non-Gaussian signals, particle  
416 filters approximate the posterior distribution through sequential sampling.  
417 Bayesian online learning techniques offer a general framework for continually  
418 updating model parameters as new data arrive. Using these techniques we  
419 will build models that robustly infer kinematics over months.

420 **Concept Drift in Machine Learning.** In the machine learning litera-  
421 ture, non-stationarity is often framed under the concept of *concept drift*,  
422 which refers to changes in the joint distribution of inputs and outputs over  
423 time. Such drift can take various forms—sudden, gradual, or cyclical (e.g., re-  
424 emergence of behavioral patterns linked to circadian or ultradian rhythms).

425 Techniques for handling concept drift generally fall into three categories:  
426 (1) *detection methods*, which monitor for significant changes in data distri-  
427 bution; (2) *adaptation methods*, which incrementally update models using  
428 strategies such as sliding windows, online learning, or ensemble-based ap-  
429 proaches; and (3) *forgetting mechanisms*, which allow models to discard out-  
430 dated information while retaining relevant past knowledge.

431 We will apply techniques from the concept drift literature to models that  
432 fall outside the previous categories of focus (e.g., linear models, artificial  
433 neural networks, and state-space models). In particular, we will explore their  
434 use in building **Recognition-Parametrized Models (RPMs)** to estimate joint  
435 behavioral and neural latent variables over timescales of weeks to months.

436 In summary, robust analysis of NaLoDuCo datasets requires models that  
437 continuously adapt to evolving data distributions. Our offline analysis frame-  
438 work will integrate both established adaptive algorithms and cutting-edge  
439 methods from continual learning and concept drift to meet this challenge.

### 440 3.4.1.3 Computational efficiency

441 Neural and behavioral data analysis is most effective when computations  
442 are performed quickly, ideally in real time. Slow computations discourage  
443 data exploration and hinder scientific discovery. The large dataset sizes gen-  
444 erated by NaLoDuCo experimentation pose a significant challenge for fast  
445 data analysis.

446 To overcome this limitation, we will combine distributed and GPU com-  
447 puting. Distributed computing is a paradigm in which tasks and data are  
448 divided across multiple computers. Instead of relying on a single powerful  
449 machine, distributed computing accelerates processing by executing multiple  
450 parts of a computation in parallel. GPU computing is a parallel computing  
451 approach that uses Graphics Processing Units (GPUs) to accelerate compu-  
452 tational tasks. Unlike traditional Central Processing Units (CPUs), which  
453 execute a few complex operations sequentially, GPUs consist of thousands  
454 of smaller cores optimized for executing many operations simultaneously.  
455 Distributed computing allows to split workload across multiple machines,  
456 overcoming memory and computational limitations. It is particularly useful  
457 for scaling to massive datasets.

458 For distributed computing to deliver substantial speed improvements,  
459 computations must be decomposable into independent parallel tasks. Due to  
460 their serial dependencies, time series models are difficult to decomposed in  
461 this manner. Still, time series models can benefit from distributed comput-  
462 ing infrastructures, as many parts of time series pipelines are parallelizable,  
463 like preprocessing steps (e.g., filtering, feature extraction, normalization) or  
464 parallel model evaluation across hyperparameter sweeps. In addition, when  
465 datasets are too large to fit in memory, distributed computing (e.g., with  
466 Ray, Dask, or Spark) can Distribute I/O and preprocessing, train models in

467 parallel on different subsets (e.g., one model per animal or time window) and  
468 run hyperparameter sweeps or model variants in parallel. Furthermore, even  
469 with serial dependencies GPU acceleration significantly speeds up the pro-  
470 cessing of each item in the time series, specially when large matrix operation  
471 are involved.

472 We will develop accelerated implementations of all methods in the library.  
473 These implementations will use **JAX** for model learning, inference, and nu-  
474 merical computation, **Apache Spark** or **Dask** to distribute pre-processing and  
475 feature extraction, and **Ray** to distribute machine learning and deep learning  
476 functionality.

477 Related to this item is the library **Thunder**, which accelerate the anal-  
478 ysis of large scale neural data. It was pioneering by introducing the use  
479 of distributed computing in neural data analysis. Our library is different  
480 from **Thunder** in that, besides analyzing large scale neural data, it processes  
481 continual recordings, and needs to overcome non-stationarity problems. In  
482 addition, **Thunder** implements simpler methods assuming independent and  
483 identically distributed data, while our library contains more sophisticated  
484 time series ones.

#### 485 **3.4.1.4 Deliverables**

- 486 1. repository containing implementations of machine learning algorithms  
487 for offline processing NaLoDuCo experimental data, adapted to op-  
488 erate in non-stationary environments, and optimized to perform at  
489 scale when running on public clouds or institutional high-performance-  
490 computing clusters.
- 491 2. SWC NaLoDuCo foraging dataset stored in DANDI.
- 492 3. deployment of the methods in 1 in Amazon EC2 instances, to allow  
493 users to analyze on the cloud the datasets in 2.

#### 494 **3.4.2 Real-time machine learning methods**

##### 495 **3.4.2.1 Real-time machine learning in neuroscience**

496 Real-time machine learning (RTML) is widely used across sectors such as  
497 finance, logistics, and environmental monitoring. For instance, in climate  
498 science, RTML enables real-time wildfire and flood detection from satellite

499 data, as well as the forecasting of extreme weather events using streaming  
500 radar and sensor signals. In food delivery systems, RTML is used to esti-  
501 mate delivery times based on traffic, kitchen queue lengths, and historical  
502 performance, and to dynamically optimize dispatching routes.

503 Surprisingly, RTML is still underutilized in neuroscience. This represents  
504 a missed opportunity—particularly in the context of NaLoDuCo experimen-  
505 tation—where adaptive, low-latency computation could significantly enhance  
506 both experimental control.

507 **Real-Time Experimental Design Verification.** In traditional neuro-  
508 science workflows, analysis is done offline, often days or weeks after data  
509 collection. Errors or design flaws are only discovered post hoc, sometimes  
510 necessitating a costly repetition of experiments. This problem is exacerbated  
511 in NaLoDuCo settings, where experiments may last weeks or months. RTML  
512 can address this by providing online assessments of experiment progress and  
513 data quality, allowing early detection of issues and in-situ protocol adjust-  
514 ments.

515 **Intelligent Neuromodulation.** Neuromodulation can be performed op-  
516 tically, chemically, or electrically. Typically, stimulation is delivered at pre-  
517 defined times or based on simple thresholds in neural or behavioral signals.  
518 With RTML, these interventions can be driven by more sophisticated models  
519 that infer high-level internal states from ongoing data.

520 For example, a scientist may hypothesize that a peak in a latent neural  
521 variable—estimated in real time from a prefrontal cortex population—signals  
522 the onset of a foraging decision. To test this, she uses an online latent variable  
523 model to forecast the peak’s occurrence and triggers optogenetic inactivation  
524 just before the predicted moment. If the intervention disrupts foraging onset,  
525 this provides causal support for the hypothesis.

526 **Intelligent Data Storage.** As the richness and duration of NaLoDuCo  
527 experiments increase, storing all raw data becomes infeasible. We will need  
528 RTML algorithms to make real-time decisions about what data to retain and  
529 what to discard.

530 For example, consider a setup with ten high-resolution cameras moni-  
531 toring a mouse in a large arena. Storing all video streams continuously is  
532 inefficient. Instead, a tracking model can estimate the animal’s location in

533 real time. When the confidence of the tracker is high, only the streams from  
534 relevant cameras are saved. When uncertainty is high, more data can be  
535 preserved for later inspection.

### 536 3.4.2.2 Bonsai and Bonsai.ML

537 Bonsai is a widely adopted open-source software ecosystem for experimental  
538 control in neuroscience (Lopes et al., 2015). With support from the BBSRC,  
539 we are developing software infrastructure to enable intelligent experimenta-  
540 tion through the Bonsai.ML package.

541 We have already integrated several real-time ML models into Bonsai.ML,  
542 including linear regression, linear dynamical systems, hidden Markov mod-  
543 els, and Bayesian point-process decoders. In collaboration with researchers  
544 at SWC and UCL, we have applied these tools to real-time inference of vi-  
545 sual receptive fields, foraging kinematics, behavioral state classification, and  
546 spatial decoding from hippocampal spiking activity.

547 However, existing Bonsai.ML methods assume stationarity, which—as  
548 discussed in Section 3.4.1—is inappropriate for NaLoDuCo data. We will  
549 adapt these methods to operate under non-stationary conditions using tech-  
550 niques outlined in Section 3.4.1.2.

551 All new RTML methods for non-stationary experimental control will be  
552 released as open-source extensions to the Bonsai.ML package.

553 At both the SWC and AIND, Bonsai is used for experimental control. In  
554 collaboration with scientists at these institutes, we will apply our new RTML  
555 methods to process non-stationary data and address cutting-edge scientific  
556 questions in state-of-the-art NaLoDuCo experiments.

### 557 3.4.2.3 Deliverables

- 558 1. New methods for processing non-stationary data, integrated into the  
559 Bonsai.ML package and made available to the neuroscience community.
- 560 2. Peer-reviewed publications co-authored with researchers at SWC and  
561 AIND, demonstrating scientific discoveries enabled by the new RTML  
562 capabilities.



### 563 3.4.3 Visual exploration

564 Visualizations are essential for extracting insight from any dataset. Given  
565 the scale of NaLoDuCo datasets, downloading them locally is impractical.  
566 Therefore, visualization methods must operate where the data resides—either  
567 in the cloud or on institutional compute clusters.

568 We will develop visualization functionality for both continuous datasets  
569 (Section 3.4.3.1) and epoched datasets, where epochs are anchored around  
570 events identified by advanced machine learning methods (Section ??).

#### 571 3.4.3.1 Continuous visualizations

572 Continuous visualizations will enable users to seamlessly explore large-scale  
573 behavioral and neural datasets spanning weeks to months. Users should  
574 be able to fluidly zoom out to gain a high-level overview (e.g., across an  
575 entire month) and zoom in to inspect millisecond-level detail. Our goal is  
576 to provide an interactive experience analogous to Google Maps—where one  
577 can zoom from a global perspective down to individual buildings—with time  
578 series data.

579 To achieve this, we will employ a combination of tiling, hierarchical stor-  
580 age, and streaming techniques:

#### 581 Multi-Resolution Tiling.

- 582 • Large volumetric and time series datasets will be preprocessed into tiles  
583 at multiple spatial and temporal resolutions.
- 584 • When the user zooms into a specific time or spatial window, only the  
585 relevant tiles at the appropriate resolution will be rendered, minimizing  
586 latency and resource use.

#### 587 Hierarchical Storage.

- 588 • Data will be organized using hierarchical file formats (e.g., Zarr, HDF5)  
589 that support chunked access and multi-resolution storage.
- 590 • These formats allow efficient random access to specific subsets of data  
591 and integration with modern data infrastructure.

592 **On-Demand Streaming.**

- 593 • Visualizations will stream data dynamically based on the user’s current  
594 view, leveraging cloud infrastructure to deliver data at the required  
595 resolution and scale.
- 596 • We will develop custom APIs for real-time access and transformation  
597 of neural and behavioral data streams.

598 **3.4.3.2 Epoched and interactive visual analytics**

599 A key strength of our platform is its support for **epoched visualization**  
600 **and interactive, closed-loop visual analytics**, which together enable the  
601 discovery and refinement of neural and behavioral patterns in long-duration  
602 datasets.

603 Epoched visualizations are essential for analyzing data around events of  
604 interest—such as decision points, sensory cues, or machine learning-inferred  
605 transitions. These visualizations will support:

- 606 • Grouping trials or epochs by event type, time of day, or machine  
607 learning-inferred state
- 608 • Overlaying neural, behavioral, and environmental variables aligned to  
609 key event markers
- 610 • Flexible sorting and filtering of epochs to uncover context-dependent  
611 patterns

612 We will implement interactive interfaces that allow researchers to define,  
613 explore, and compare arbitrary epoch-based segments. These will support  
614 exploratory data analysis as well as hypothesis-driven comparisons across  
615 conditions, individuals, and time periods.

616 **Machine Learning-Defined Events.** A core feature of our system will be  
617 the ability to align epochs not just to experimenter-defined events, but also  
618 to latent state transitions inferred via unsupervised methods (e.g., hidden  
619 Markov models, behavioral clustering, inverse reinforcement learning). This  
620 will support deeper investigation into emergent patterns in long-duration,  
621 naturalistic behavior.

622 **Closed-Loop Analytics.** There will be a *closed-loop interaction* between  
623 visualizations and machine learning algorithms: algorithmic outputs will gen-  
624 erate new visualizations, and visual insights will guide further machine learn-  
625 ing analysis, forming an iterative discovery cycle. This process allows the  
626 visualization platform to function not just as a display tool, but as a central  
627 component in data-driven scientific inquiry.

628 In this loop:

- 629 • **Machine learning algorithms** extract latent states, classify behav-  
630 iors, infer structure, or forecast dynamics from NaLoDuCo data.
- 631 • These outputs feed into the visualization engine to generate novel views  
632 (e.g., state-aligned rasters, dynamic embeddings, attention maps).
- 633 • **Users explore these visualizations interactively**, discovering un-  
634 expected, task-agnostic, or contextual patterns.
- 635 • New queries and insights drive further rounds of machine learning anal-  
636 ysis—closing the loop.

637 This design enables researchers to co-develop computational models and  
638 scientific hypotheses iteratively, with human insight and machine inference  
639 deeply intertwined.

### 640 3.4.3.3 Software stack for interactive visualizations

641 To support scalable, cloud-based, and interactive visualization of NaLoDuCo  
642 datasets, we will develop our system using a modern and modular software  
643 stack optimized for high performance, extensibility, and ease of integration  
644 with existing neuroscience infrastructure.

#### 645 Frontend (User Interface).

- 646 • **React.js** will serve as the primary framework for building a dynamic,  
647 modular, and responsive web-based interface.
- 648 • Visualization components will leverage libraries such as **D3.js**, **Plotly**,  
649 and **Deck.gl** to render interactive time series, raster plots, and behav-  
650 ioral trajectories at scale.
- 651 • For GPU-accelerated rendering of large datasets, we will use **WebGL**  
652 and related technologies such as **regl** or **Three.js**.

653 **Backend (Computation and Data Services).**

- 654 • The backend will be written in **Python**, using **FastAPI** or **Flask** to  
655 serve data and model outputs to the frontend.
- 656 • Time series pre-processing, tiling, and downsampling will be handled  
657 via **NumPy**, **Xarray**, and **Dask** to enable scalable, distributed pro-  
658 cessing.
- 659 • For storage, we will use chunked, cloud-native formats such as **Zarr**  
660 and **HDF5**, which allow efficient retrieval and hierarchical access to  
661 long-duration recordings.
- 662 • Machine learning integration will rely on **PyTorch**, **scikit-learn**, and  
663 model serving frameworks such as **TorchServe** or **ONNX Runtime**.

664 **Cloud Infrastructure.**

- 665 • The system will be deployable on commercial or academic cloud plat-  
666 forms (e.g., AWS, GCP, or institutional clusters).
- 667 • For orchestration of services, we will use **Kubernetes**, enabling auto-  
668 scaling and distributed deployment of visualization and ML services.
- 669 • **Docker** containers will ensure reproducibility and portability across  
670 environments.
- 671 • The visualization system will integrate directly with the **DANDI Archive**  
672 for cloud-native access to neurophysiology data.

673 **Data Interoperability.**

- 674 • All tools will be compatible with **Neurodata Without Borders**  
675 (**NWB**) and follow FAIR data principles.
- 676 • The system will expose APIs for programmatic access to raw and  
677 derived data, enabling integration with existing tools like **Bonsai**,  
678 **CaImAn**, or **napari**.

679 This software stack ensures that our visualization tools will be perfor-  
680 mant, scalable, and usable across a wide range of environments, from local  
681 lab systems to cloud-based scientific platforms.

#### 682 **3.4.3.4 Deliverables**

- 683 1. visualisations for continuous behavioural and neural recording
- 684 2. visualisations for epoched behavioural and neural recording
- 685 3. visualisations for model outputs
- 686 4. indexing system to support intelligent visualisations
- 687 5. deployment of the above items to allow users to visualise NaLoDuCo
- 688 DANDI datasets on the cloud

#### 689 **3.4.4 Offline spike sorting**

##### 690 **3.4.4.1 Outputs**

- 691 1. Repository with implementations and benchmarking of offline spike
- 692 sorting algorithms for long-duration recordings

#### 693 **3.4.5 Online spike sorting**

##### 694 **3.4.5.1 Outputs**

- 695 1. Repository with implementations and benchmarking of online spike
- 696 sorting algorithms

## 697 **References**

- 698 Haykin, S. S. (2002). *Adaptive filter theory*. Pearson Education India.
- 699 Lopes, G., Bonacchi, N., Frazão, J., Neto, J. P., Atallah, B. V., Soares, S.,  
700 Moreira, L., Matias, S., Itskov, P. M., Correia, P. A., et al. (2015). Bonsai:  
701 an event-based framework for processing and controlling data streams.  
702 *Frontiers in neuroinformatics*, 9:7.

### 703 3.5 Applicant and team capability to deliver

704 Word limit: 1,650

705 Why are you the right individual or team to successfully deliver the pro-  
706 posed work?

707 What the assessors are looking for in your response

708 Please ensure the current job titles of the core team members are included  
709 here to ensure eligibility can be established for the core team roles assigned.  
710 Find out more about [UKRI's core team roles in funding applications](#) and our  
711 [eligibility guidance](#).

712 Evidence of how you, and if relevant your team, have:

- 713 • the relevant experience (appropriate to career stage) to deliver the pro-  
714 posed work
- 715 • the right balance of skills and expertise to cover the proposed work
- 716 • the appropriate leadership and management skills to deliver the work  
717 and your approach to develop others
- 718 • contributed to developing a positive research environment and wider  
719 community

720 You may demonstrate elements of your responses in visual form if rele-  
721 vant.

722 Further details are provided in the Funding Service.

723 The word limit for this section is 1,650 words: 1,150 words to be used for  
724 R4RI modules (including references) and, if necessary, a further 500 words  
725 for Additions.

726 Use the Résumé for Research and Innovation (R4RI) format to showcase  
727 the range of relevant skills you and, if relevant, your team (project and project  
728 co-leads, researchers, technicians, specialists, partners and so on) have and  
729 how this will help deliver the proposed work. You can include individuals'  
730 specific achievements but only choose past contributions that best evidence  
731 their ability to deliver this work.

732 Complete this section using the R4RI module headings listed. Use each  
733 heading once and include a response for the whole team, see the UKRI guid-  
734 ance on R4RI. You should consider how to balance your answer, and empha-  
735 sise where appropriate the key skills each team member brings:

- 736     • contributions to the generation of new ideas, tools, methodologies, or  
737       knowledge
- 738     • the development of others and maintenance of effective working rela-  
739       tionships
- 740     • contributions to the wider research and innovation community
- 741     • contributions to broader research or innovation users and audiences  
742       and towards wider societal benefit

#### 743     Additions

744     Provide any further details relevant to your application. This section is  
745     optional and can be up to 500 words. You should not use it to describe  
746     additional skills, experiences, or outputs, but you can use it to describe any  
747     factors that provide context for the rest of your R4RI (for example, details  
748     of career breaks if you wish to disclose them).

749     Complete this as a narrative. Do not format it like a CV.

750     References may be included within this section.

751     The roles in funding applications policy has descriptions of the different  
752     project roles.

### 753 **3.6 Project partners**

754 Add details about any project partners' contributions. If there are no project  
755 partners, you can indicate this on the Funding Service.

756 A project partner is a collaborating organisation who will have an integral  
757 role in the proposed research. This may include direct (cash) or indirect (in-  
758 kind) contributions such as expertise, staff time or use of facilities. Project  
759 partners may be in industry, academia, third sector or government organisa-  
760 tions in the UK or overseas, including partners based in the EU.

761 If you are applying via the IPA or LINK scheme, please include details  
762 of industry partners here.

763 If applying under the BBSRC-NSF lead agency scheme, please include  
764 details of your US partner here.

765 Add the following project partner details:

- 766 • the organisation name and address (searchable via a drop-down list or  
767 enter the organisation's details manually, as applicable)
- 768 • the project partner contact name and email address
- 769 • the type of contribution (direct or in-direct) and its monetary value

770 If a detail is entered incorrectly and you have saved the entry, remove the  
771 specific project partner record and re-add it with the correct information.

772 For audit purposes, UKRI requires formal collaboration agreements to be  
773 put in place if an award is made.



### 774 **3.7 Project partners: statement of support**

775 Word limit: 3,000

776 Only complete a statement of support if you have named project partners  
777 in the project partner section above. A statement is required to be provided  
778 from each partner you named in the ‘Project partners’ section.

779 If you are applying via the IPA or LINK scheme, please include details  
780 of industry partner support here.

781 What the assessors are looking for in your response

782 A project partner is a collaborating organisation who will have an integral  
783 role in the proposed research. This may include direct (cash) or indirect (in-  
784 kind) contributions such as expertise, staff time or use of facilities.

785 Each statement should:

- 786 • confirm the partner’s commitment to the project
- 787 • clearly explain the value, relevance, and possible benefits of the work  
788 to them
- 789 • describe any additional value that they bring to the project

790 Ensure you have prior agreement from project partners so that, if you are  
791 offered funding, they will support your project as indicated in the ‘Project  
792 partners’ section.

793 For audit purposes, UKRI requires formal collaboration agreements to be  
794 put in place if an award is made.

795 Do not provide a statement of support from host and project co-leads’  
796 research organisations.

797 Do not provide a statement of support from collaborators. Contributions  
798 from collaborators not listed as project partners can be outlined in ‘Applicant  
799 and team capability to deliver’.

### 800 **3.8 Trusted research and innovation (TR&I)**

801 Word limit: 100

802 Does the proposed work involve international collaboration in a sensitive  
803 research or technology area?

804 What the assessors are looking for in your response

805 Demonstrate how your proposed international collaboration relates to  
806 TR&I, including:

- 807 • list the countries your international project co-leads, project partners  
808 and visiting researchers, or other collaborators are based in
- 809 • if international collaboration is involved, explain whether this project  
810 is relevant to one or more of the 17 areas of the UK National Security  
811 and Investment (NSI) Act
- 812 • if one or more of the 17 areas of the UK National Security and Invest-  
813 ment (NSI) Act are involved list the areas

814 If your proposed work does not involve international collaboration, you  
815 will be able to indicate this in the Funding Service. If your proposed work  
816 does not involve international collaboration, please indicate this by stat-  
817 ing ‘N/A’. We may contact you following submission of your application to  
818 provide additional information about how your proposed project will comply  
819 with our approach and expectation towards TR&I, identifying potential risks  
820 and the relevant controls you will put in place to help manage these risks.

### 821 3.9 Resources and cost justification

822 Word limit: 1,000

823 What will you need to deliver your proposed work and how much will it  
824 cost?

825 What the assessors are looking for in your response

826 The FEC of your project can be up to a maximum of £2 million. We will  
827 fund 80million, we will fund £1.6 million and your research organisation will  
828 be expected to fund £400,000.

829 Please note, equipment over £10,000 is funded by BBSRC at 50%. The  
830 Funding Service does not currently have the ability to record this. For this  
831 round we ask that you include equipment over £10,000 in ‘Exceptions’ at  
832 100% of cost. We will cut this to 50% at award. You must ensure you have  
833 prior agreement from your research organisation to fund the remaining 50%.

834 Justify the application’s more costly resources, in particular:

- 835 • project staffsignificant travel for field work or collaboration (but not  
836 regular travel between collaborating organisations or to conferences)
- 837 • any equipment that will cost more than £10,000
- 838 • any consumables beyond typical requirements, or that are required in  
839 exceptional quantities
- 840 • all facilities and infrastructure costs
- 841 • all resources that have been costed as ‘Exceptions’

842 Assessors are not looking for detailed costs or a line-by-line breakdown  
843 of all project resources. Overall, they want you to demonstrate how the  
844 resources you anticipate needing for your proposed work:

- 845 • are comprehensive, appropriate, and justified
- 846 • represent the optimal use of resources to achieve the intended outcomes
- 847 • maximise potential outcomes and impacts

### 848 **3.10 Data management and sharing**

849 Word limit: 500

850 How will you manage and share data collected or acquired through the  
851 proposed research?

852 What the assessors are looking for in your response Provide a data man-  
853 agement plan that clearly details how you will comply with UKRI's published  
854 data sharing policy, which includes detailed guidance notes.

### 855 **3.11 Facilities**

856 Word limit: 250 Does your proposed research require the support and use of  
857 a facility?

858 What the assessors are looking for in your response

859 If you will need to use a facility, follow your proposed facility's normal  
860 access request procedures. Ensure you have prior agreement so that if you  
861 are offered funding, they will support the use of their facility on your project.

862 For each requested facility you will need to provide the:

- 863 • name of facility, copied and pasted from the facility information list  
864 (DOCX, 42KB)
- 865 • proposed usage or costs, or costs per unit where indicated on the facility  
866 information list
- 867 • confirmation you have their agreement where required

868 Facilities should only be named if they are on the facility information list  
869 above. If you will not need to use a facility, you will be able to indicate this  
870 in the Funding Service.

871 **3.12 Ethics and responsible research and innovation**  
872 **(RRI)**

873 Word limit: 500

874 What are the ethical or RRI implications and issues relating to the pro-  
875 posed work? If you do not think that the proposed work raises any ethical  
876 or RRI issues, explain why.

877 What the assessors are looking for in your response

878 Demonstrate that you have identified and evaluated:

- 879 • the relevant ethical or responsible research and innovation considera-  
880 tions
- 881 • how you will manage these considerations

882 If you are collecting or using data you should identify:

- 883 • any legal and ethical considerations of collecting, releasing or storing  
884 the data (including consent, confidentiality, anonymisation, security  
885 and other ethical considerations and, in particular, strategies to not  
886 preclude further reuse of data)
- 887 • formal information standards that your proposed work will comply with

888 You may demonstrate elements of your responses in visual form if rele-  
889 vant. Further details are provided in the Funding Service.

### 890 **3.13 Genetic and biological risk**

891 Word limit: 700

892 Does your proposed research involve any genetic or biological risk?

893 What the assessors are looking for in your response

894 In respect of animals, plants or microbes, are you proposing to:

- 895 • use genetic modification as an experimental tool, like studying gene  
896 function in a genetically modified organism
- 897 • release genetically modified organisms
- 898 • ultimately develop commercial and industrial genetically modified out-  
899 comes

900 If yes, provide the name of any required approving body and state if  
901 approval is already in place. If it is not, provide an indicative timeframe  
902 for obtaining the required approval. Identify the organism or organisms as  
903 a plant, animal or microbe and specify the species and which of the three  
904 categories the research relates to.

905 Identify the genetic and biological risks resulting from the proposed re-  
906 search, their implications, and any mitigation you plan on taking. Assessors  
907 will want to know you have considered the risks and their implications to  
908 justify that any identified risks do not outweigh any benefits of the proposed  
909 research.

910 If this does not apply to your proposed work, you will be able to indicate  
911 this in the Funding Service.

### 912 **3.14 Research involving the use of animals**

913 Does your proposed research involve the use of vertebrate animals or other  
914 organisms covered by the Animals Scientific Procedures Act?

915 What the assessors are looking for in your response

916 If you are proposing research that requires using animals, download and  
917 complete the Animals Scientific Procedures Act template (DOCX, 74KB),  
918 which contains all the questions relating to research using vertebrate animals  
919 or other Animals (Scientific Procedures) Act 1986 regulated organisms.

920 Save it as a PDF, ensuring it is no larger than 8MB. The Funding Service  
921 will provide document upload details when you apply. If this does not apply  
922 to your proposed work, you will be able to indicate this in the Funding  
923 Service.



### 924 3.15 Conducting research with animals overseas

925 Word limit: 700

926 Will any of the proposed animal research be conducted overseas?

927 What the assessors are looking for in your response

928 If you are proposing to conduct overseas research, it must be conducted  
929 in accordance with welfare standards consistent with those in the UK, as in  
930 Responsibility in the use of animals in bioscience research. Ensure all named  
931 applicants in the UK and overseas are aware of this requirement. Ensure all  
932 named applicants in the UK and overseas are aware of this requirement.

933 If your application proposes animal research to be conducted overseas,  
934 you must provide a statement in the text box. Depending on the species  
935 involved, you may also need to upload a completed template for each species  
936 listed.

937 Statement

938 Provide a statement to confirm that:

- 939 • all named applicants are aware of the requirements and have agreed to  
940 abide by them
- 941 • this overseas research will be conducted in accordance with welfare  
942 standards consistent with the principles of UK legislation
- 943 • the expectation set out in Responsibility in the use of animals in bio-  
944 science research will be applied and maintained
- 945 • appropriate national and institutional approvals are in place

946 Templates

947 Overseas studies proposing to use non-human primates, cats, dogs, equines  
948 or pigs will be assessed during NC3Rs review of research applications. Pro-  
949 vide the required information by completing the template from the question  
950 ‘Research involving the use of animals’.

951 For studies involving other species, such as:

- 952 • rodents
- 953 • rabbits
- 954 • sheep

- 955       • goats
- 956       • pigs
- 957       • cattle
- 958       • *xenopus laevis* and *xenopus tropicalis*
- 959       • zebrafish

960       Select, download, and complete the relevant Word checklist or checklists  
961 by exploring NC3Rs checklist for the use of animals overseas.

962       Save your completed template as a PDF and upload to the Funding Ser-  
963 vice. If you use more than one checklist template, save it as a single PDF.  
964 The Funding Service will provide document upload details when you apply.

965       If conducting research with animals overseas does not apply to your pro-  
966 posed work, you will be able to indicate this in the Funding Service.

### 967 **3.16 Research involving human participation**

968 Word limit: 700

969 Will the project involve the use of human subjects or their personal in-  
970 formation? What the assessors are looking for in your response

971 If you are proposing research that requires the involvement of human sub-  
972 jects, provide the name of any required approving body and whether approval  
973 is already in place. Justify the number and the diversity of the participants  
974 involved, as well as any procedures.

975 Provide details of any areas of substantial or moderate severity of impact.

976 If this does not apply to your proposed work, you will be able to indicate  
977 this in the Funding Service.

978 **3.17 Research involving human tissues or biological**  
979 **samples**

980 Word limit: 700

981 Does your proposed research involve the use of human tissues, or biolog-  
982 ical samples?

983 What the assessors are looking for in your response

984 If you are proposing work that involves human tissues or biological sam-  
985 ples, provide the name of any required approving body and whether approval  
986 is already in place.

987 Justify the use of human tissue or biological samples specifying the nature  
988 and quantity of the material to be used and its source.

989 If this does not apply to your proposed work, you will be able to indicate  
990 this in the Funding Service.