BRAINWORKS Interview Notes

Mohammad M. Ghassemi, Ph.D.

1. Interview Purpose and Approach

1.1. Structure

<u>Section 1:</u> Overview of the interview purpose and approach.

<u>Section 2:</u> Executive summary of the key interview takeaways.

<u>Section 3:</u> Contact information of the PIs that participated as interviewees.

<u>Section 4:</u> Raw interview notes.

<u>Section 5:</u> Suggested papers grouped by theoretical theme

<u>Section 6:</u> Additional contacts

<u>Appendix A</u>: Secondary interviews

<u>Appendix B</u>: Interview request email

1.2. Purpose

In January 2021, Dr. Ghassemi performed interviews with 15 Neuroscience Principal Investigators (PI). The interviews were performed to refine the specific data extraction and representation methodology for the *BRAINWORKS* Platform proof-of-concept and helped clarify: (1) the definition of the term "theory" in neuroscience research community, (2) the current neuroscience theoretical landscape and (3) the ways in which neuroscientists tend to describe their theories in natural language.

1.3. Approach

Unless stated otherwise, each interview was 30-minutes in duration. In an effort to guide the discussion, each PI was provided with the following three questions in advance of the meeting:

1. Specific features that should be considered when extracting and/or representing theories embedded within neuroscience texts.
2. An example project-related text you've generated, and the succinct statement of the theory embedded within that project-related text.
3. Information on existing efforts that are complementary to, or may be informative for, the project's initial proof-of-concept.

Note that the PIs were not required to provide direct answers to the above questions. The interview trajectory was adjusted to accommodate the individual points raised by the PIs.

2. Executive Summary

2.1. “Theory” encompasses a broad range of scientific statements

A theory is a testable (i.e. falsifiable) set of mathematical and/or natural language statements that answers a why/what/how question involving two or more variables. For example: “Conjunctive input processing drives feature selectivity in hippocampal CA1 neurons”.

2.3. The theoretical landscape is complex, diverse and expansive

This is due, in part, to the variety of distinct scientific disciplines (Bioengineering, Physics, Mathematics, Biology, etc.) that work on Neuroscience problems, all of which have distinct analytical frameworks, styles of communication and preferences for disseminating results. Hence, the unification of the theoretical landscape is beyond the capabilities of any single investigator. This provides motivation for the development of automated approaches that can account for the heterogeneity of the field when unifying the theoretical landscape.

2.1. Feedback on BRAINWORKS was overwhelmingly positive

Keeping up to date with the literature is time-consuming for experts and immensely challenging for trainees. 14 of the 15 PIs interviewed felt that BRAINWORKS may help them keep better track of theoretical developments in the literature and inform future research directions; all 15 PIs felt that the platform would be useful for trainees.

2.4. Conversations inform BRAINWORKS development approach

2.4.1. Data Collection: Approximately half the PIs suggested that we choose a theoretical sub-domain for the platform’s initial instantiation. The most common sub-domain suggested (3/15) was neural plasticity. BRAINWORKS should use perspective pieces and review articles (as opposed to research papers) to obtain clusters of related papers to train/tune embedding models.

2.4.2. Information Extraction: BRAINWORKS should collect theories as well as the methodological frameworks and data used to support (or falsify) the theory. It should use a Probabilistic Context Free Grammar supported by an established Neuroscience ontology to parse documents into their theory, framework, and data components.

2.4.3. Natural Language Representation: The embedding approach should start with a pre-trained model (e.g. PubBERT) that be tuned using the paper clusters extracted from the reviews and perspective pieces. Relationships between papers (e.g. shared co-authors) should be represented as a graph. Learning algorithms should be developed to predict longitudinal changes in graph properties. The performance of the approach should be validated on held-out data selected by the PIs.

2.4.4. User Interface Considerations: BRAINWORKS should allow users to (optionally) define sets of papers they deem as “similar” and dynamically warp the embedding space based on these user-provided definitions of similarity.

3. PI Contact Information

PI	Email	Title	Field	Institution
Dr. Bill Lytton	bill.lytton@downstate.edu	Prof.	Biomed. Engineering	SUNY Downstate University
Dr. Fidel Santamaria	Fidel.Santamaria@utsa.edu	Prof.	Biology	UT San Antonio
Dr. Florian Engert	florian@mcb.harvard.edu	Prof.	Biology	Harvard University
Dr. Marc Howard	marc777@bu.edu	Prof.	Brain Sciences	Boston University
Dr. Brent Doiron	bdoiron@uchicago.edu	Prof.	Statistics	University of Chicago
Dr. Ivan Soltszel	isoltesz@stanford.edu	Prof.	Neuroscience	Stanford University
Dr. Ilya Nemenman	ilya.nemenman@emory.edu	Prof.	Physics	Emory College
Dr. David Kleinfeld	dkleinfeld@ucsd.edu	Prof.	Physics	UC San Diego
Dr. Carina Curto	ccurto@psu.edu	Prof.	Mathematics	Penn State University
Dr. Sam Gershman	gershman@fas.harvard.edu	Prof.	Psychology	Harvard University
Dr. Robert Froemke	Robert.Froemke@nyulangone.org	Prof.	Neuroscience	New York University
Dr. SueYeon Chung	sueyeonchung@gmail.com	PostDoc	Neuroscience	Columbia University
Dr. Elizabeth Buffalo	ebuffalo@uw.edu	Prof.	Neuroscience	University of Washington
Dr. Greg Deangelis	gdeangelis@ur.rochester.edu	Prof.	Biomed. Engineering	University of Rochester
Dr. Jan Drugowitsch	Jan_drugowitsch@harvard.edu	Prof.	Neurobiology	Harvard University

Table 1: Biographical information of the interview participants.

4. Interview Notes

This section contains notes taken by Dr. Ghassemi in the course of his interviews with the Principal Investigators (PI) listed in Section 3; **these notes are based on Dr. Ghassemi's personal recollection of the interactions and should not to be taken as a precise or complete transcript of the interviews.**

4.1. Dr. Bill Lytton and Dr. Fidel Santamaria:

Dr. Lytton and Dr. Santamaria had a positive reaction to the project's vision and potential utility. More specifically, they agreed that the BRAINWORKS Platform (if successfully developed) could inform new avenues for brain research and may assist with theory integration efforts.

The PIs indicated that the development of holistic theories in neuroscience (and biology more generally) is challenging because the number of variables that influence the system are very high, but investigators are inevitably only observing a very small fraction of those variables in any given experimental setup. That is to say, if our tool is data-driven – and the data consists of theories constructed with incomplete data, it is unlikely that the tool will be able to generate holistic theories, but it may be able to point investigators towards papers/experiments that, when extended, may inform the creation of more holistic theories by the investigators.

The PIs felt that validation of the tool would be essential for longer-term community use; they suggested that we identify a specific theoretical domain where there is a clear and well known evolution in the theoretical trajectory to use as a validation set for a preliminary version of the tool. As an initial integration target, he suggested looking at the set of theories that explain how neurons encode information: firing rate, spike timing, etc.

4.2. Dr. Florian Engert

Dr. Engert indicated that Neuroscience is lacking a unified theoretical direction to start with and that this problem has grown worse as the community continues to shift its efforts "from explaining to describing". He indicated that a significant proportion of the contemporary efforts involve the creation of detailed catalogs and maps, not theoretical frameworks which he described as "a dilemma and a conundrum".

Dr. Engert was skeptical that data science would be able to assist with the generation or integration of theories in Neuroscience; he doesn't think that the data or the methods are the problem, rather the theory just isn't being done. To support this, he cited that Darwin wasn't very accomplished by data science standards – but was a brilliant theorist: "data science is descriptive not theoretical". He was skeptical that his community would ever be able to synthesize a unified theoretical direction because most papers, including his own, are lacking a solid theoretical framework to begin with: "There may not be a unified theory."

His advice for the project was to narrow the scope of BRAINWORKS as much as possible "Even trying to understand how Neurons encode information is still too broad". He suggested investigating how the neural representations of measurable physical information such as how the brain represents the distance of an object*:* "if an animal is some distance from it's prey how does it know the distance to it's prey?"

He emphasized that neuroscience papers have three levels of theoretical explanation:

1. **Why** is it useful for an animal to solve a problem?
2. **What** are the algorithms that are being used to solve a problem?
3. **How** does the organism optimize the algorithm using available information?

For a theory to be "integrative", he said it would need to explain all three levels. He noted that most papers contain only one level of explanation and thus they do not contain theories in the complete sense of the word. He postulated that this may be because there is not a single theory that explains all three levels; he thinks of the development of the brain as a sequence of (somewhat) disjoint evolutionary "hacks" that organisms use to accomplish survival tasks.

I asked him what role (if any) data science could play in advancing Neuroscience. He indicated that data science could help with compression of complex data into simpler representations that are easier to theorize with: "to help us extract salient information from irrelevant factors".

With regards to the BRAINWORKS platform specifically, Dr. Engert thought the tool could be useful as a literature search assistant. Specifically, he was interested in finding papers that were the "nearest neighbor" to a given paper of interest: "If I could find all the papers that relate to my theory at an abstract level, not a key-word level, that would help us perform more effective research. For instance, if we are looking at the Zebrafish, and monkeys and other fish – it would be nice to know similarities and or differences in algorithms, behaviors, etc."

He felt that the ability of the investigator to specify what they meant by "similar" would be essential for the tool to return functional results. He suggested having the user define sets of papers as "similar" up front, and for the tool to dynamically account for that when constructing the embedding space.

4.3. Dr. Marc Howard

Dr. Howard felt it would be useful to systematically organize the contemporary (and historical) ideas in Neuroscience and that insofar as BRAINWORKS could assist with this, it would be very helpful. He believes that there could be a single theory that unites all neuroscience observations but that the community is 25-30 years away from finding it. He has an intuition for what the uniting theory might look like, and said that he's been attempting to make the first steps in that direction by uniting models formally in his research group: "Identifying higher order ways of representing what populations of neurons are doing".

He noted strong parallels between neuroscience and NLP indicating that in both fields, the primary challenge was one of representation: "People assume that if they're recording from one hundred thousand neurons, that each of these neurons are a separate basis vector; that's probably wrong. It's a dimensionally reduction problem in neuroscience, we need a way to represent the data so it's easy to see the connections".

He noted that any models in the literature that don't tie behavior to brain activity are almost sure to be wrong: "It has to be reversable". For validation of the BRAINWORKS tool, he suggested that I look into evidence accumulation models from Dr. Gold.

4.4. Dr. Brent Dorion

Dr. Dorion opened the conversation by sharing that he was “very happy you are doing this. It would be very interesting to see how the theories flow through the literature. The field is far too broad, and no one neuroscientist knows all the theories.”

He felt that theorical work has benefited greatly from the BRAIN initiative and was complementary of the BRAIN initiatives’ overall organization and funding philosophy: “It used to be hard to get funding for this kind of work.”

With respect to parsing the Neuroscience literature, Dr. Dorion indicated that there will be unique challenges when approaching the Neuroscience literature: “Neuroscience, compared to physics, has unique challenges. We don’t have a unified way to even discuss ideas, or unified ways to approach problems.”.

He dichotomized most researchers in neuroscience as either focused on “math or biology”. He felt that the investigators with a mathematical focus would be easier to convince that BRAINWORKS will have value and that biologically-focused investigators would be harder to convince. He indicated that the co-existence of these two research factions results in heterogenous publication formats: “Some people don’t want to see the equations.”

He emphasized that the BRAINWORKS Platform may have to consider the audience a paper was written for when performing information extraction. He suggested that we focus our efforts on papers with a mathematical focus.

4.5. Dr. Ivan Soltszel

Dr. Soltszel indicated that BRAINWORKS is “Exactly what needs to be done.” recalling that “When I was a postdoc, I could read 10 papers and understand a field. Today, it’s impossible to read everything – there is just too much. Not only have the number of papers grown but they’re also significantly more dense. These days, a paper may have 7 figures. Once could spend their whole lifetime just trying to catch up on the literature. That’s why I think what you’re doing is very important.” He stressed that the tool would be especially useful for trainees “If somebody doesn’t know the literature – this tool you are building would be amazing.”

With regard to analysis of the literature, he encouraged a breadth-first approach “many neuroscience projects have a very narrow focus, but to understand how the brain works you need to look at as many things as possible first.”

Dr. Soltezel also emphasized that any method we develop must account for the writing style of the authors: “Each paper is always written by a PI. Each PI has a style for guiding the reader through to the conclusion. There is a lot of artform in the papers. Sometimes the data is good, but the results are a farse. If you are going to solve this – you need to understand the narrative tension.”

He felt that an essential element of the tool would be to score papers with respect to the “credibility coefficient” of the authors. We may construct such a score by looking at how often the authors release code/data, or if their work seems to go unused despite being published.

He encouraged us to take advantage of structural properties in the papers and to scale back the analysis to just abstracts if it poses too great a challenge: “There are certain pseudo structures that exist within the papers, and you should take advantage of that information. The POC could be done just using abstracts – you may not need to parse the entire document.”

He encouraged the use of reviews to extract papers for embedding space training: “There are primary research papers, and then there are reviews. A good review is the right place to start. Maybe you can use those to seed the papers for your work.”

He suggested opinion pieces as another avenue: “you might want to look at those. That’s where you will get people’s honest opinions . It’s less constricted than research.”

4.6. Dr. Ilya Nemenman

Dr. Nemenman felt that BRAINWORKS would be useful for traditional approaches but would fail to provide value for novel methodologies: “What you are doing will be straight forward and will work in the context of traditional models – but it might not work in the context of novel theoretical frameworks... detecting a paradigm shift is not going to be possible with a machine learning approach.”

Despite that, he felt the project would be useful to help him keep track of developments in the literature: “your project could be very useful to me personally. Google scholar is a little too good at trying to figure out what I want and can create scientific tunnel vision. If you can design the tool to cluster papers by theory, it would be interesting to then explore the periphery of that space – that’s probably where the ideas worth reading are.”

4.7. Dr. David Kleinfeld

Dr. Kleinfeld was enthusiastic about the BRAINWORKS platform and felt it may serve as a starting point for a broader discussion in the community about what the core theories are – and what directions investigators should invest their research efforts to advance those theories.

Dr. Kleinfeld was optimistic that a unified theory of the brain would one day be discovered, but not without strong theoretical foundations: “Neuroscience is like material science – the equations are always imperfect and following the formula will never yield perfect crystals – but despite that, there are overarching principles and at large enough scales – those principles hold and are important to understand... The brain of the fly is very structured – it’s not a complete mess. In order for our field to progress, we need to approach research with the belief that these structures exist – that will only happen with strong theoretical foundations.”

He stressed that the Neuroscience literature may be dichotomized into: (1) models and (2) analyses: “Keep in mind that theory has formed two branches - those that model systems, at various levels of abstraction, and those that analyze data, with dimensionality reduction as the current mantra.”

He pointed out that computational methods inform Neuroscience which in-turn inform computational methods: “Much of the computational advances are informed, at least in part, by thinking about neuroscience problems. Hopfield networks were a metaphor for many things in neuroscience. The ring model is another that’s been proposed.”

4.8. Dr. Carina Curto

Dr. Curto felt that consolidating information across the extracted theories would be challenging: " I don't think that you'll be able to consolidate information across equations and models."

She urged us to distinguish between works that were theoretically similar from those with similar frameworks: "Theoretical similarities have to do with what's being explained. This is separate from the framework you are using – are you using a neural network, or large matrices. Theory is different from theoretical framework. Theory is the specific equation that explains y, using x. But there are frameworks as well – some people like to use Bayesian approaches, others Neural networks, etc."

She encouraged us to focus on language, and avoid analyses of equations: "Better to focus on the language aspects rather than on matching equations. You may also want to look for papers that have common references. It's likely that papers with citation overlap will be theoretically related."

With regard to an initial target theoretical domain, she proposed continuous attractor networks: "Continuous attractor networks are a generic proposal for how a network could be structured. It's been applied to controlling eye movements, hippocampus place cells, and several other applications, but the same theoretical framework is being shared across the specific problems. It would be important to distinguish the framework from the theory because the frameworks are transmissible, and the theories tend not to be."

She proposed that we structure our extraction work to collect:

1. **Theory**: What is the question you want to answer?
2. **Frameworks**: What method are you using to answer the question?
3. **Data**: What data did you use to support or falsify the theory?

4.9. Dr. Sam Gershman

Note: Sam was recommended to speak in place of Dr. Bernardo Sabatini

Dr. Gershman felt that the field of NLP has not advanced enough that it would be possible to meaningfully extract theories from scientific papers: "How do you represent the research space? It's a worthy problem to solve - but I doubt that what you come up with will be useful to researchers; it sounds more like a tool created by NIH administrators, that will be used by NIH administrators for making funding decisions."

Dr. Gershman felt that BRAINWORKS would be completely irrelevant to Neuroscience research: "There is no way that I would use an NLP tool to inform my research, but I understand that there is a politics to this and that knowing what others are up to could be informative for some trainees, which is something interesting, but not from the perspective of a practicing scientist like myself."

Dr. Gershman felt that asking PIs to provide clusters of theoretically-related papers would be meaningless without taking human similarity judgements into account when creating the theoretical embedding space: "It may very well be the case that different PIs mean different things when they provide you with a cluster of *similar* papers. I'm personally interested in the use of similarity judgements with respect to highly structured objects. Theories are structured objects that have discrete parts, and if you don't take that structure into account, your method probably won't work."

He did not agree with the extraction structure proposed by Prof Curto: "I'm not sure that it makes sense to represent papers according to dimensions you've decided on ahead of time (e.g. Theory, Framework, Data)"

4.10. Dr. Robert Froemke

Note: Robert was asked to speak in place of Dr. Richard Tsien

He's interested in neuronal plasticity or rather, understanding precisely how neurons change themselves as a function of action potentials and how those changes relate to memory, behavior, etc. He believes that plasticity is "the holy grail of neuroscience" research - how does the brain change itself?

He stated that Neuroscience is "In desperate need of this tool" (referring to the BRAINWORKS Platform) his justification being that, despite significant research efforts, the practical and/or translational impact of Neuroscience works has been very limited; he believes this lack of impact is due to the fact that neuroscientists come from a variety of scientific disciplines and that the expertise, analytical frameworks and style of communication varies as a function of those backgrounds. Neuroscience is "equal parts biology, computer science, physics" and each of discipline has their own traditions around structure of experiments and how to use language to describe results.

He encouraged us to use neural plasticity as an initial theoretical test bed for BRAINWORKS because there is an established tradition of research along with clear points of evolution that the method *should* be able to identify: "the Neural Plasticity field has the advantage that both computational and experimental people have been in lock-step for many years, so the possibilities for integration of research might be stronger there than in other domains where theory outpaces experiment."

He explained that neural plasticity is fundamentally a question of re-weighting nodes in a neural network, and because of this, it's been easier for people from math and computer science backgrounds to enter the domain. He indicated that there is a clear trajectory of theories around neural plasticity in the literature.

He was enthusiastic about the BRAINWORKS platform. He emphasized more than once that he would love to see how the ideas of Neuroscience have evolved, and he thinks our tool might even allow him to extrapolate into the future. He suggest that we present at the COSINE meeting, BRAIN initiative meeting, and/or NeuroIPS to solicit feedback about the work

4.11. Dr. SueYeon Chung

Note: Robert was asked to speak in place of Dr. Larry Abbott

Dr. Chung is working in the center for theoretical neuroscience with Dr. Abbott as her mentor. Her work involves computational neuroscience and deep learning. More specifically, she is investigating how artificial neural networks can be used to better understand organic neural networks. Her current approach involves training neural networks on several tasks that are traditionally used in animal experiments; she then studies how the artificial neural network weights change during training. She postulates that some properties of the representations learned by the artificial neural networks may be similar to those learned by the organic networks.

Dr. Chung felt that BRAINWORKS was an interesting project although she had some questions about how we would be creating the embedding space from the natural language documents; she correctly surmised that the performance of our tool would depend on the properties of the embedding space. To enable continuous improvements to the embedding methodology, she suggested that the tool be open sourced, and that members of the community be allowed to compete to develop the best approach to the embedding problem.

She categorized the Neuroscience literature as being of two varieties: (1) “modeling work”, which involves how to develop models of the brain and (2) “Analytical work”, which involves the interpretation and/or extraction of meaning from Neural data. She felt that neural plasticity was an excellent initial domain to validate the tool.

4.12. Dr. Elizabeth Buffalo

Dr. Buffalo is interested in understanding the links between brain regions; how does activity in one area modulate activity in another? Her U19 project is investigating how the prior experiences of an organism change what they pay attention to during a novel task. She noted that the relationship between attention and experience is well-established in the psychological community, but the precise neuroscience mechanics are not well understood. She’s interested in understanding how place and memory fit together and the specific role of the hippocampus in this process.

Dr. Buffalo was excited by the vision of the BRAINWORKS platform and felt that it could be especially impactful for trainees, junior scientists, and even established PIs. More specifically, she thought it would valuable for quickly identifying new papers in a domain, and for understanding how theoretical domains have changed over time. Her overall impression was that the tool would be “very interesting and helpful”.

When asked to provide a theoretical domain for platform validation, she felt that neural plasticity had potential. She indicated that neural plasticity research has had well known contending theories that have since been harmonized.

She suggested that in addition to text, we consider processing conference proceedings posted on YouTube. She also felt that perspective pieces (as opposed to research papers) would be a useful way to obtain an overview of the theoretical landscape because the authors would be less conservative in their writing. She noted that there are journals that specialize in opinion pieces including Nature Reviews Neuroscience.

4.13. Dr. Greg Deangelis and Dr. Jan Drugowitsch

Dr. Deangelis and Dr. Drugowitsch felt that BRAINWORKS would be a challenging but worthy endeavor; Deangelis stated: “Most of us still do key-word searches on PubMed. I think the way we do literature search today is not very effective. In principle, there is a better way to understand the theories in the literature and BRAINWORKS could be the way to do that.”. They suggested that investigators be allowed to tune the tool to accommodate their individual use-cases.

Dr. Deangelis suggested that we choose a narrow theoretical domain to start with. More specifically, he suggested sensory representations as the area for initial focus. With regards to approaches for parsing the literature, Dr. Drugowitsch suggested that we look at scientific output as either: “purely experimental, cross-section between theory and experiments, and theory only”. He emphasized that tackling the theory space would be harder because “when looking at theory, everything is possible but if you look at the experimental domain, there are more constraints. It will be easier to take away structure from the experimental papers compared to the theory papers.”

When asked to provide a specific definition of the term theory in neuroscience, the PIs emphasized that the precise definition of the term was likely to depend on the goals of the investigator: “There are two kinds of theories, one is a natural language articulation for what/how/why a brain network does something, the other is a mathematical formulation for why/how/why the brain does something; the language needs to be mathematical for it to be well defined.”

5. Suggested Papers:

All interviewees were asked to provide papers that they deemed to be theoretically related. In Table 2 we list the 46 scientific publications provided by the interviewees, organized by theoretical theme. These papers will be used to validate the performance of the BRAINWORKS platform. More specifically, we will assess if the BRAINWORKS platform is able to cluster these papers according to the theoretical themes provided by the PIs.

Suggested by	Theoretical Theme	Publication
Dr. Florian Engert	Evidence accumulation and perceptual decision making	Akrami, Athena, Charles D. Kopec, Mathew E. Diamond, and Carlos D. Brody. 2018. "Posterior Parietal Cortex Represents Sensory History and Mediates Its Effects on Behaviour." Nature 554 (7692): 368–72.
		Bahl, Armin, and Florian Engert. 2020. "Neural Circuits for Evidence Accumulation and Decision Making in Larval Zebrafish." Nature Neuroscience 23 (1): 94–102.
		Borst A & Bahde S Visual information processing in the fly's landing system. J. Comp. Physiol. A; 163, 167–173 (1988)
		Romo, Ranulfo, Adrián Hernández, Antonio Zainos, and Emilio Salinas. 2003. "Correlated Neuronal Discharges That Increase Coding Efficiency during Perceptual Discrimination." Neuron 38 (4): 649–57.
		Rossi-Pool, Román, Emilio Salinas, Antonio Zainos, Manuel Alvarez, José Vergara, Néstor Parga, and Ranulfo Romo. 2016. "Emergence of an Abstract Categorical Code Enabling the Discrimination of Temporally Structured Tactile Stimuli." Proceedings of the National Academy of Sciences of the United States of America 113 (49): E7966–75.
		Salzman, C. D., K. H. Britten, and W. T. Newsome. 1990. "Cortical Microstimulation Influences Perceptual Judgements of Motion Direction." Nature 346 (July): 174–77.
		Vázquez, Yuriria, Antonio Zainos, Manuel Alvarez, Emilio Salinas, and Ranulfo Romo. 2012. "Neural Coding and Perceptual Detection in the Primate Somatosensory Thalamus." Proceedings of the National Academy of Sciences of the United States of America 109 (37): 15006–11.
		Zylberberg, Ariel, Christopher R. Fetsch, and Michael N. Shadlen. 2016. "The Influence of Evidence Volatility on Choice, Reaction Time and Confidence in a Perceptual Decision." eLife 5 (October). https://doi.org/10.7554/eLife.17688.
Dr. Sam Gershman	Bayesian brain hypothesis	https://www.unige.ch/medecine/neuf/files/8415/7398/7414/MaBeckLathamPougetNN06.pdf
		https://www.nature.com/articles/nn1691
		https://www.mitpressjournals.org/doi/abs/10.1162/08997660460733976

		https://www.mitpressjournals.org/doi/abs/10.1162/neco.2008.20.1.91
		https://science.sciencemag.org/content/sci/331/6013/83.full.pdf
		https://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.1002211
Dr. David Kleinfeld	Recurrent networks	Ring attractor dynamics in the Drosophila central brain; SS Kim, H Rouault, S Druckmann, V Jayaraman; Science 356 (6340), 849-853
		Theory of orientation tuning in visual cortex; R Ben-Yishai, RL Bar-Or, H Sompolinsky; Proceedings of the National Academy of Sciences 92 (9), 3844-3848
		Rate models for conductance-based cortical neuronal networks; O Shriki, D Hansel, H Sompolinsky; Neural computation 15 (8), 1809-1841
		Traveling waves and the processing of weakly tuned inputs in a cortical network module; R Ben-Yishai, D Hansel, H Sompolinsky; Journal of computational neuroscience 4 (1), 57-77
	Weakly coupled oscillators	Traveling electrical waves in cortex: insights from phase dynamics and speculation on a computational role; GB Ermentrout, D Kleinfeld; Neuron 29 (1), 33-44
		Frequency plateaus in a chain of weakly coupled oscillators, I.; GB Ermentrout, N Kopell; SIAM journal on Mathematical Analysis 15 (2), 215-237
		From Kuramoto to Crawford: exploring the onset of synchronization in populations of coupled oscillators; SH Strogatz; Physica D: Nonlinear Phenomena 143 (1-4), 1-20
		Phase dynamics for weakly coupled Hodgkin-Huxley neurons; D Hansel, G Mato, C Meunier; EPL (Europhysics Letters) 23 (5), 367
Dr. Ilya Nemenman	Bayesian modeling of sensorimotor systems	www.pnas.org/cgi/doi/10.1073/pnas.1713020115
		Shadmehr R, Smith MA, Krakauer JW (2010) Error correction, sensory prediction, and adaptation in motor control. Annu Rev Neurosci 33:89–108.
		Kording KP, Tenenbaum JB, Shadmehr R (2007) The dynamics of memory as a consequence of optimal adaptation to a changing body. Nat Neurosci 10: 779–786.
		Gershman SJ (2015) A unifying probabilistic view of associative learning. PLoS Comput Biol 11:e1004567.
		Wolpert DM (2007) Probabilistic models in human sensorimotor control. Hum Mov Sci 26:511–524.
		Wei K, Ko¨rding K (2009) Relevance of error: What drives motor adaptation? J Neurophysiol 101:655–664.

		Hahnloser R, Narula G (2017) A Bayesian account of vocal adaptation to pitch-shifted auditory feedback. PLoS ONE 12:e0169795.
Dr. Marc Howard	time and space and memory	Integrating time from experience in the lateral entorhinal cortex. Nature (2018)
		Remembering the past to imagine the future: the prospective brain (2007)
		What Is a Cognitive Map? Organizing Knowledge for Flexible Behavior. Neuron (2018)
		On the Integration of Space, Time, and Memory. Neuron (2017)
		The hippocampus as a cognitive map. Oxford university Press (1978)
		The hippocampus, memory, and place cells: is it spatial memory or a memory space? Neuron (1999)
		Temporal maps and informativeness in associative learning.}, TINS (2009)
Dr. Robert Fromeke	Neural Plasticity	Conjunctive input processing drives feature selectivity in hippocampal CA1 neurons Nature Neuroscience (2015)
		Long-lasting potentiation of Synaptic Transmission in the dentate area of the anaesthetized rabbit following stimulation of the prefrontal path. Journal of Physiology (1973)
		Heterosynaptic Plasticity Determines the Set Point for Cortical Excitatory-Inhibitory Balance. Neuron (2020)
		Spike-timing-dependent synaptic plasticity depends on dendritic location. Letters to Nature (2005)
		Common Forms of Synaptic Plasticity in the Hippocampus and Neocortex in Vitro. Science (1993)
		A Synaptically Controlled, Associative Signal for Hebbian Plasticity in Hippocampal Neurons. Science (1997)
		Postsynaptic Calcium Is Sufficient for Potentiation of Hippocampal Synaptic Transmission. Science (1998)
		Long-Term Potentiation--A Decade of Progress? Science (1999)
		Regulation of Synaptic Efficacy by Coincidence of Postsynaptic APs and EPSPs. Science (1997)
		Competitive Hebbian learning through spike-timing-dependent synaptic plasticity. Nature (2000)

Table 2: Papers suggested by the interviewed PIs. Papers are sorted by theoretical themes provided by the PIs.

6. Suggested Additional Contacts:

Name	Institution	Suggested By	Note
Dr. Aurel Lazar	Columbia	Dr. Bill Lytton	Performed an NLP analysis of the Neuroscience literature in fruit flies.
Dr. Bard Ermentrout	University of Pittsburgh	Dr. David Kleinfeld	A highly regarded theorist among neuroscientists
Dr. David Hansel	Centre de Neurophysique	Dr. David Kleinfeld	A highly regarded theorist among neuroscientists
Dr. Wolfram Schultz	University of Cambridge	Dr. Marc Howard	Has insights into temporal difference learning theory.
Dr. Markus Meister	Caltech	Dr. Florian Engert	Wants to translate information from natural language to numerical vectors

Table 3: Individuals suggested by the interviewed PIs for additional discussion.

Appendix A: Additional Conversations

A.1. Dr. Martha Flanders

Dr. Flanders provided guidance on how to structure my interviews with the Neuroscience PIs. She indicated that Team E didn't yet have a unified definition of what a theory is in neuroscience. She indicated that a theory is a set of natural language statements that: (1) answers a what/how/why question involving two or more variables, (2) is testable and thus can be disproven. She suggested that we have the PIs define what a theory is and request a specific example of a theory from some of their papers.

Dr. Flanders scientific background is in motor control. She was interested in the use of control theory to model how organisms move from one distributed representation to another. She indicated that in fields others that motor control, there are not many solid theories and that for this reason – it might be useful to focus efforts on theories of motor control: "Anything that connects the neurons, their growth, their activity to the behavior of the organism – how do they move."

A.2. Christopher Belter:

Mr. Belter provided guidance on tools to download scientific papers into machine-readable formats. He informed me that PubMed is not the best resource for full-text articles, but that abstracts are widely available. PubMed Central provides access to full-text articles but contains only a subset of the articles from PubMed.

He indicated that extracting the text would be possible, but some curation would be required to unify the data formats: "Theoretically, you should be able to pull the XML – there are about 10 variants. The utilities allow you to pull it via an API. NCBI may be able to do a bulk download of all the data in PubMed. You may also want to see if there is a PubMed database already out there somewhere."

Mr. Belter provided the following references to assist with:

1. Downloading PubMed/PubMed central data

- PubMed/MEDLINE article metadata downloads: https://www.nlm.nih.gov/databases/download/pubmed_medline.html
- PubMed Central full text downloads: https://www.ncbi.nlm.nih.gov/pmc/tools/textmining/
- PubMed / PubMed central APIs (E-utilities): https://dataguide.nlm.nih.gov/

2. Obtaining full text data from publishers

- Elsevier developer portal: https://dev.elsevier.com/
- Springer Nature developer portal: https://dev.springernature.com/

3. Use of Scopus, NIH RePORTER, and PubMed APIs

- https://github.com/christopherBelter/scopusAPI
- https://github.com/christopherBelter/nih_reporter_api
- https://github.com/christopherBelter/textmining

A.3. Dr. Asiyah Lin

Dr. Lin provided guidance on neuroscience ontologies that might be useful for developing the data extraction pipelines in BRAINWORKS. There are several efforts related to the brain, but she wasn't sure which were the most relevant for the project: "Ontology of Biomedical Investigation may be useful for your purpose - it is actively developed and maintained and describes assets for studies. You may need to add some terminology to this. But, this would serve as a good starting point for your work. Machine learning methodologies is another ontology that might be helpful." Dr. Lin provided the following additional resources to learn more about ontologies.

Relevant Resources:

- https://www.gps.health/covid19_knowledge_accelerator.html
- http://www.obofoundry.org/ontology/stato.html
- http://www.obofoundry.org/ontology/obcs.html
- http://www.obofoundry.org/
- https://bioportal.bioontology.org/

Appendix B: Email Used to Contact PIs

Title: A Conversation About The NIH BRAIN WORKS Platform

Dear <INSERT PI HERE>,

I hope you're well.

I'm Mohammad Ghassemi, a Data and Technology Advancement Scholar at the NIH working on the "BRAIN initiative Workspace to ORganize the Knowledge Space" platform (BRAIN WORKS); I report to John Ngai, Grace Peng (CCed), and Jim Gnadt (also CCed).

Over the next year, I will be responsible for delivering a proof-of-concept for the BRAIN WORKS platform that uses Natural Language Processing to:

1. Automatically extract structured representations of theories from unstructured scientific documents and
2. Track the evolution and propagation of the extracted theories as a function of time, domain and other factors of interest.
3.

To be clear, I am defining a "theory" here as: *an explanation for an observed relationship between two (or more) variables*.

I am interested in scheduling a 30 minute meeting with you (and/or a relevant member of your research team) to discuss:

1. Specific features that should be considered when extracting and/or representing theories embedded within neuroscience texts.
2. An example project-related text you've generated, and the succinct statement of the theory embedded within that project-related text.
3. Information on existing efforts that are complementary to, or may be informative for, the project's initial proof-of-concept.

Please let me know a good time for us to speak in the upcoming weeks. To help us make the best use of the call, <u>please share a 1-page abstract on your BRAIN-funded research in advance of our meeting</u> and any publications or public resources that might be useful for this discussion.

Warm regards,