The R Series: Roundtables on Mental Health and Data Science[1]

Magnus Boman and Sumithra Velupillai

8 October 2020

In a world in which researchers and clinicians are overloaded, is it possible to work and toil together, with just pen and paper and a very large table to assist the conversation? We have found the answer to be in the positive: it counters the myopia of our daily tasks, and allows for a digital detox.[2] A notion that became important to us is that of an R—a roundtable in our series—happening in the midst of everything else, in a particular location chosen to distance it from everything else. Lessons learned from four roundtables with psychologists, psychiatrists, data scientists, artificial intelligence researchers, brain researchers, molecular biologists, and ethics researchers were collected, and are presented here.[3] A total of seven days of intellectual immersion is reported on, spread over 25 sessions with person names and session titles, but without presentations, papers, or slides. All events were invite only and participants worked clinically or academically in Sweden or the UK. The only object was to investigate how data science could help reduce human suffering. Simple.[4]

Introduction

The People

ON DECIDING WHO TO INVITE, we looked for researchers and practitioners in either data science or mental health. We also identified people in the intersection, as there is a clear interest in data science and machine learning technology from mental health professionals. Perhaps less obvious is the willingness of data scientists to focus on mental health problems. Many participants stated the relatively large impact that data science could have on clinical work as a motivating factor.

The most common profession was psychologist or psychiatrist, with data scientist coming in second. All data scientists used machine learning methods regularly, with natural language processing being the most used family of methods.[5] An example application that several attendants were studying would be text analysis of electronic health records to better understand mental health symptoms and behaviours, such as suicidal ideation.[6]

[2] Nir Eyal. How to be indistractable. *Psyche*, 2 September, 2020.

[3] "The brain is a prediction machine, preparing you for what is about to happen in the world. We are tailor-made for our own world. So, what is the overlap with the worlds of others?" (R4)

[4] "Mental health is inherently difficult, approximations of reality, filtered through subjective observations." (R3)

Figure 1: During a break at Chandos House, and the third R, no one is looking at the person taking the picture.

[5] "I was asked to *do some NLP on this data*." (R3)

[6] "Suicidality is imprecise and rare." (R2); "If you apply NLP on EHRs, perhaps the last line reads *Patient is not suicidal now.* and the patient is sent home, but the reason the patient came in was an attempt." (R4)

Figure 2: In the inner yard of Chandos House, everyone is looking at the machine taking the picture.

The Format

ON DECIDING HOW TO MEET, we deliberately avoided slides and designated presentation slots. We also decided not to use these meetings for any measurable outcomes such as grant proposals or scientific publications.[7] Instead, we focused on curating and nourishing an agora of entirely conversational nature. We provided moderation in terms of suggested agendas and suggested speakers on given topics, but allowed for plenty of time for open discussions. This led to an open atmosphere where participants could reflect on their own research, and also gain new perspectives from other participants. Even in the breaks, the discussion often stayed in plenary format, as the group never exceeded 20 in number. Lunches were long and with walk+talks. Subgroups formed serendipitously; no meals were planned or included.

[7] "Hype funding: making money rather than helping. At the end of my days, how am I helping?" (R3)

The Workflow

THE IDEA OF MAKING THE R's analogue events was influenced by a thought experiment called Analogue Friday. It popped up inside a European project on mediated presence:[8] to finally be able to meet meaningfully and efficiently without physically being in the same room. The experiment asked you to consider every Friday as free from digital devices. No phones, computers, digital music playing,... It invited us to contemplate how hard this would be for work, leisure, and social networking, and to ponder the positive effects.

Several R attendants have spoken on the amazement they have felt at their own presence in the room, physically and mentally. This was never our explicit intention. Instead, we had sought to eliminate ego by telling people to abstain from thinking of publishing a paper, presenting slides, recruiting for projects, and the rest that goes with typical academic and business interaction. We gave people a slot under a session heading, nothing else. We moderated proceedings so that no single person and no single issue took up too much time, and the total number of moderator interventions with the flow of any kind could be counted on one hand, over four R's. In short, the R's had wondrous intellectual self-hygiene. The detoxing experience could at least in part be explained by the workflow itself: to boot up a computer or check your phone for notifications would interfere. Not so much because people would notice and be distracted by your sounds or motions, but to yourself: five or ten seconds of not paying attention would kick you off the train of thought, and then you would have to catch up running. Not worth it.

If those of us, privileged as we are, enjoying unabated intellectual exchange did immerse in one or more R's, should not similar arrangements spread like wildfire? Are the investments or efforts too large, the logistics too complicated, the possible attendants too busy? For our last R, we had more than twice the 20 or so that attended on a list of interested and probably very interesting people. This interest was built word-of-mouth, mostly, but we also found people we thought we should have already invited, and some people that seemed always too busy to set aside two full days. But there was no search for profiles of possible participants, and there was no marketing. In hindsight, the bias is huge: the people that did come proved their control of their own time by being there. The people that repeatedly told us how sorry they were for not coming probably were not right for an R. But had they come, after all, something might have changed; in the R, and in them.

[8] A one-minute film might be the fastest way to approach the outcome of that project: www.youtube.com/watch?v=5cDeINxzYRk
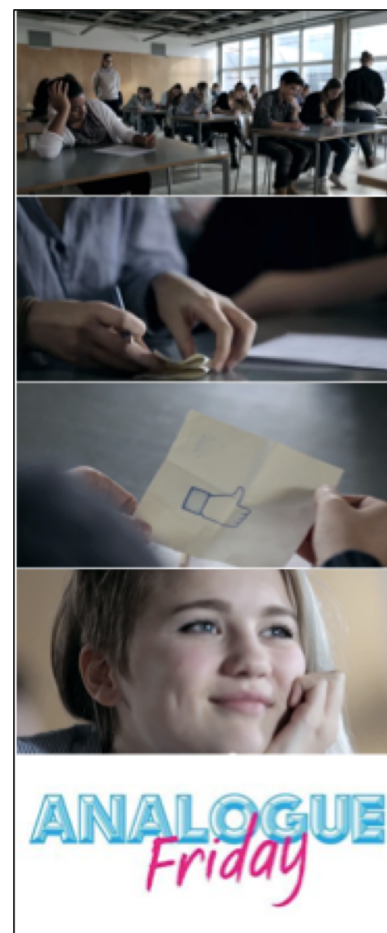

Figure 3: Screen dumps from an Analogue Friday video used as a firestarter for discussions in a very large European project on mediated presence and countless other topics.

The Outcome

The Clinical Challenge

VIRTUALLY ALL OF THE ROUNDTABLE PARTICIPANTS work in, or with, clinics.[9] Crystalizing thousands of patient-therapist dialogues into one:[10]

- Are you reading my mind?
- No. Is this hard for you?
- I want to engage with you, it looks like you're interested in my mind.
- What do you need?
- A home, a job, and a friend.

The first question the patient asks is about many things, one being trust.[11] Some patients have been shown to read fear in other people as anger.[12] The therapist answers in the negative, but was in fact reading the patient's mind. This is part of the process of mentalization.[13]

In an interdisciplinary research environment, it is unrealistic for clinicians to learn data science or for data scientists to understand all aspects of the clinical work. What can be achieved is a ping pong match where the burden of evidence, usefulness, and possible impact take turns:

C: What can machine learning do for me?
D: I don't know, tell me more about the data sets you have.[14]
C: I have databases and a pipeline for getting knowledge out of these databases, which support clinical insights, but also scientific papers, students, ...
D: Are your databases unimodal or multimodal?
C: I have text, images, and more.
D: Let's start unimodal, by looking only at text first. What do you want to know?
C: I want to know which patients have shown self-harm behaviour.
D: Is this information not in your database already?
C: Yes, but not readily available. There is no database column with Yes/No indications of self-harm. Instead, it is documented in the notes.
D: Which notes would that be?
C: It could be in event notes, correspondence notes, or some other notes.
D: I am unfamiliar with those.
C: Event notes are taken at the point-of-care, correspondence notes are summaries, referral letters, and other means to communication between clinicians. There are also notes in forms, which might have free-text fields, and different clinics might have different forms for different kinds of situations.
D: What about electronic health records?[15]

[9] "Clinicians often don't know what they want." (R2)

[10] "Symptoms change depending on context." (R3)

[11] "Trust breach, as in trauma or abuse from your care-giver, leads to revering the process of mind-mindedness. You learn that it's simpler not to learn from them: epistemic trust is replaced by epistemic hypervigilance. *People that have experienced trauma* is the new term for borderline personality. We know that trauma leads to more trauma, and repetitive patterns based on the question *Why did they do it?*" (R4)

[12] "Psychopaths may see frightened faces as angry faces." (R2)

[13] "*What is going on in another person's head?*, parametrised on anxiety, hopefully leading to certainty, as in *Oxytocin levels have dropped to zero.*" (R3); "The stance that we are able to explain people's intentions by reading their minds. Those of us that are good at it are good at it all the time. It is sometimes described as a uniquely human trait, but this is not helpful at all." (R4)

[14] "*You don't know what you don't know*-type data analyses." (R2)

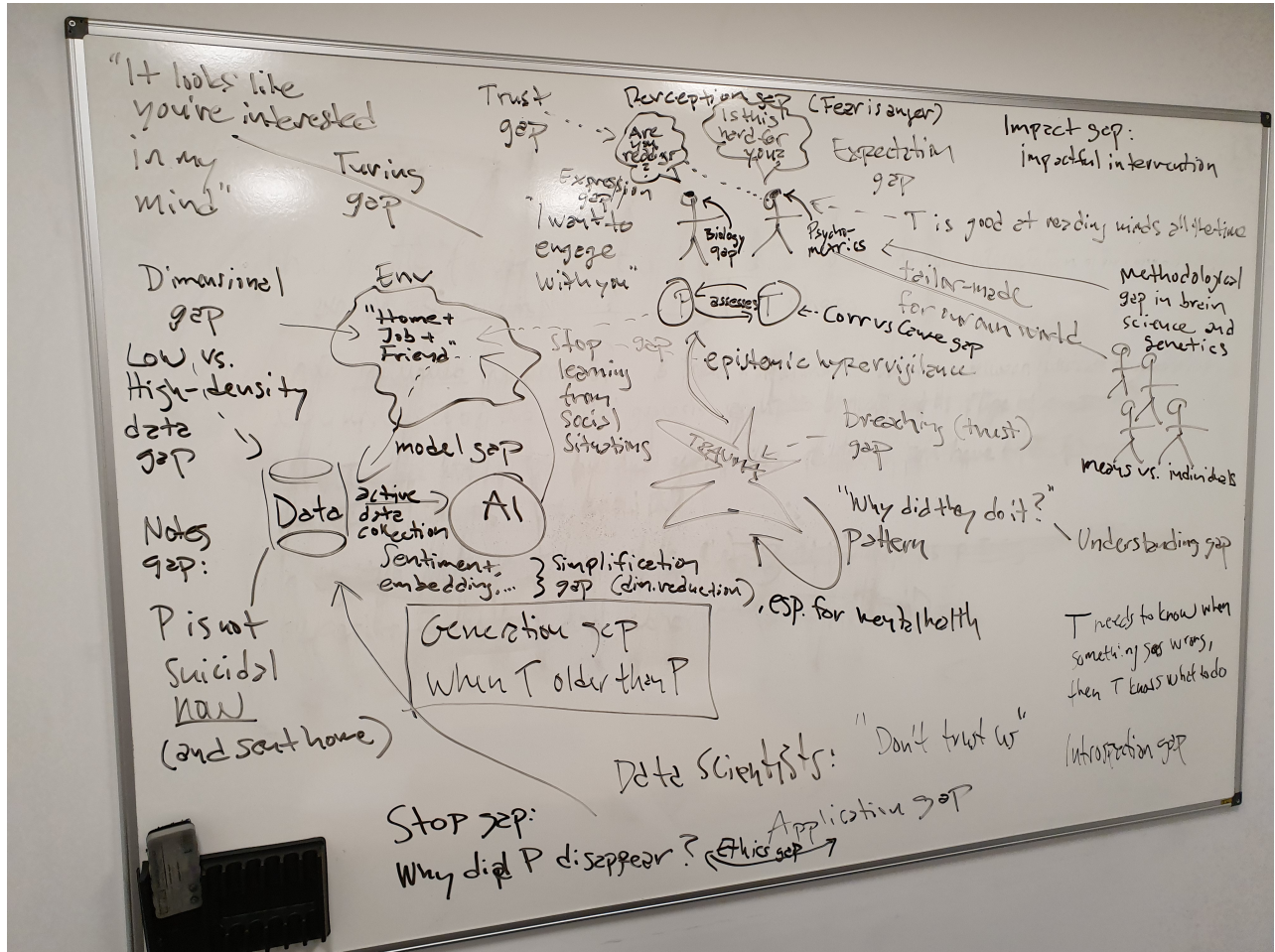[15] "There is a reluctance to use free-text fields, because people perceive of it as help." (R1)

Figure 4: Our wyteboard in the Virginia Woolf building of KCL, the day after R4.

C: All of these notes are part of the electronic health record, the full database that in theory any clinician would be able to access. In practice though, each clinician might be given only partial reading rights to the full database. Which parts depend on the clinician's role in the health care system. If the clinician is a researcher, the full database could be available for a specific research problem, like self-harm.

D: So, I can safely assume I could do data-driven studies of the full database if I had your support?[16]

C: Yes.

D: In that case, data science can build learning systems that use different kinds of bias in your data to jump to inductive conclusions that might be hard to get to for humans, because of the extensive connections between different data points. If we stay with text, we can do deep learning for all kinds of data you've got that's annotated and shallow learning for the rest.

C: What kinds of bias do you mean?[17]

D: Machine learning thrives on bias.

C: What do you mean by that? Bias for me is a confounding factor. What would bias in the data be for data science?

[16] "Pre-processing in data-driven projects take 15-30 per cent of the total research time, but for clinical applications it could get up to 45 per cent." (R3)

[17] "There is a clinician bias in how they talk about patients." (R2)

D: When we talk about bias, we talk about surface structures and details invisible to the human eye.[18] For example, if different notes were written by different doctors, the machines will be able to learn this fact.

C: How?

D: Through stylometric analysis, things like verbosity, the use of terms, the mix of structured and unstructured sentences, the use of acronyms, the time points in which the notes were filed.

C: How is that bias?[19]

D: The data reflects something from the situation in which it is created, and this is something that the data science methods will capture.

C: What do you mean by annotated?

D: Your clinical practice produces a gold standard where you have made various classifications.[20] An example would be risk level of self-harm. it might be a dichotomy – risk/no-risk, but it could also be a spectrum.

C: OK, but how would you find those classifications?

D: If your assessments are structured, according to some classification system like ICD-10, we will look for ticked boxes and specific terms in your notes. If your notes are unstructured, we will derive the most probable conclusion from your assessments as expressed in the free text.

C: Clinical research has shown that most assessments are written in free text.[21]

D: We are aware that discrete measurements of a continuous clinical reality can never capture all aspects of that reality. We understand that this process and system is not ideal for you, but we try to make the best use we can of the data that has already been collected and produced. We need your help. We need to understand more about how these assessments are written then, to create a gold standard.

C: Is there an end to my efforts? Isn't it true that analysts always want more data? I don't want to be stuck annotating when I have work to do at the clinic.[22]

D: It is true that in data science for health, there can always be more data. This is because we throw away lots of data in pre-processing. On your side, you seem to create more through imputation when you run statistics on your data, but we try not to. We also divide into training, testing, and validation steps in our methods. First we train the system, then we'd prefer to test it on another subsample, so as not to overfit. And then we validate, folding the data five or six times, possibly doing hold-out sampling.

C: I understand the bit about the holding out, the rest was Greek to me. So, how do I know when to stop working for you, so that you can do your bits?

D: I do not need you to do any annotation for me. If there are annotations, I could probably use them. The validation I talked about is very different from your randomized controlled trials, I realise that. We validate the soundness of our model more than usefulness in reality, which you might get out of an RCT.[23]

C: I have yet to see an RCT that made sense in practice, but I see what you mean then. So you can work with whatever I've got, you say?

[18] "There isn't a whole lot of structure in the structured data." (R2)

[19] "Bias means thresholds for care. If you're black, you're not less likely to have ADHD, but less likely to get treatment." (R2)

[20] "There is no gold standard for ADHD, just talking to parents and their children. The child's psychological condition is persistent." (R2)

[21] "Clinical notes are a justification of what I, as a therapist, have already done. Phenomenologically important things should not be weeded out. The anomaly has significance." (R3); "A brick wall is felt between therapist and a borderline person, so for therapy to take effect, the person needs to take an enormous risk." (R4

[22] "Early onset means more data from patients, rather than from clinicians, retaining the variance." (R3); "What does it mean to intervene early?" (R3)

[23] "I have been involved with 60 RCTs, with no positive outcomes." (R2)

D: I can, yes. Most of data science is naïve though, in the sense of the more the merrier. The more adequate the model is, the more we can predict or classify.

C: What does adequate mean here?

D: Matching up with reality.

C: But Borges taught us that the map becomes useless if it's too complicated?[24]

D: Haha, yes, he did indeed, and so there's definitely a limit to how much information that goes into the model.

C: Data or information?

D: Information. Data science is also about information science, and I meant information.

C: How do you know if it's working?

D: Well, I can vouch for the soundness of the model, but we need to define the model together so that we can solve the problems you want to solve. I do not even want to try to come up with any problems myself, I'll stick to the modelling.

C: And I'll stick to my clinical research and tell you what I need?

D: I think we understand each other.

The Data Science Narrative

WHAT IS THIS NEW THING, being exploratory and hypothesis-less, what does it mean?[25] A found diary (sitting at the end of the round table after the end of a roundtable) helps us understand what this narrative entails, and what its bearing on reality might be in the future.

I think about the $n = 1$ scientific studies and what they entail. Big Data means really really large n's. Most data-driven studies, deep learning excluded, is pretty small n, however. And the $n = 1$ ones are the ones about individuals (or dividuals, perhaps):[26] there are no population statistics involved. Big Data was promised to take us to individuals by means of statistics and clever induction. But observational studies where $n = 1$ are different. They are often anecdotal, but details and interaction, together with many other factors, can make results generalizable. And in mental health, it seems the uniqueness of each person, each case, each history, each story is as unique as the clients that walked into Holmes's office on Baker St.[27] If we learn all we can, will it even matter in practice, if environment is 90 per cent of the cause of mental unhealth, and amino acids and wet lab-detectable stuff is the rest? Should we not then spend more time on psychodynamics, a fancy way of saying we should spend more time talking to people that are unwell?[28] Epigenetics does go some way towards dialogue, even if the patient rarely gets fully involved.

[24] Wikipedia: On Exactitude in Science. Accessed: Oct 8, 2020.

[25] "A philosophical question almost, is it not, can we use data?" (R2)

[26] Postscript on the Societies of Control. *October* 59 (Winter, 1992):3–7.

[27] "Human beings as evaluators seems an outdated notion. A patient in diagnostic interview fell silent after five minutes. A single wrong word could shatter all hope of help." (R3)

[28] "*Spontaneous recovery* does not exist. I did things with my parents that made me mentally sane." (R2); "I used to tell my parents what I'd done (wrong) when they were driving." (R3)

A Slow Merger

MENTAL HEALTH IS A VERY BROAD TERM and topic, covering areas from mild temporary mental illness to severe chronic illness.[29] The history of psychology and psychiatry is long and winding, and full of variety.[30] One of the great challenges in this field is the lack of (strong) evidence for genotyping: very few aspects of mental illness can be easily measured with fixed variables (as one can, for instance, with a physical condition like a fracture).[31] That does not mean that mental health research relies only on non-measurable or incommensurable phenomena.[32] Numerous theories, scales, and instruments have been proposed, validated, and employed in clinical practice, and with success. Treatments ranging from medications to cognitive behavioural therapy have been shown to work really well for many patients, throughout history as well as now.[33] But there are still many unknowns. And society, clinical practice, research methods, science, have all co-evolved. Their interplay is perhaps more vivid, varied, and intertwined than ever before.

Computer science entered the arena slowly at first, but has now evolved into an omnipresent field, combining mathematics, statistics, linguistics, engineering, physics, social sciences, and more. Information theory is holding the social and medical sciences hostage, relentlessly moving forward with papers, algorithms, methods, and patents. It has reshaped society, clinical practice, research methods, and in our current era where the abundance of data coupled with computational approaches calls for rethinking what we can learn, how we claim that what we learn has value, and how we best make use of this, for the benefit of society.[34]

The Future

TRENDS COVERED IN THE R series include developments that in turn depend on even larger trends, like precision psychiatry, and advances in interpreting genetic data. We also jointly came up with new concepts and terms, like *the tablet tablet*, for instance. Mobile apps seem to in general lack the precision needed to augment mental health assessments, so in spite of the ubiquity of smart phones, apps might see less use in data collection, and more for providing additional support.[35]

The young now face having their digital records accumulate throughout their lives. Information previously only available in restricted quantities and expensive to collect can now be gathered

[29] "Chronically ill people should have better products. Proper biochemical science." (R2)

[30] "*Psychological* takes P4 medicine into P5 medicine." (R3)

[31] "No genetic test is predictive enough." (R3)

[32] "Are we measuring the capability of the healthcare staff to transmit enthusiasm to the patients, or are we measuring something else?" (R3)

[33] "Variability in amygdala determines responsiveness to CBT." (R3)

[34] "If we are to be judged on how we take care of the weakest citizens, mental health is where we want to be." (R4)

[35] "The six-fold increase in child mental health is not reflected in any questionnaire." (R2); "Blank out their records is how to react to an opt-out." (R2)

at scale, and potentially be used to the benefit of many more than before. This offers an opportunity for change in how effective treatments are discovered and delivered. Some therapies are already offered digitally, successfully, but there might be room and opportunity for more. Digital lives offer the possibility for novel discoveries, through co-creation and co-development, also in scientific fields.

There are, however, important considerations and challenges ahead, particularly related to socio-technical and ethical issues. We need to make sure—through regulation if necessary—that the data used for developing novel solutions is representative of those it covers and affects, and that there is no inherent bias or discrimination produced by algorithms. Any solutions developed to be used in practice need to be interpretable, at least in the sense that some form of explanation of an output can be given. This is where co-development and interdisciplinarity are crucial: for society to accept modern data science methods, they need to be transparent and understandable.

The community has so far been more keen to share models and ontologies than data. Rather than blaming data sensitivity and regulations, we would say researchers need to migrate to new contexts, new patient populations, and new laws.[36] Only then can transfer learning happen and only then can less myopic interpretations of data be made.[37]

[36] "Assent and consent." (R1); "Consent can sometimes be derived from EHRs." (R2)

[37] "I am working for the greater good." (R2)