Research meeting guidelines

Jan van Gemert
Delft University of Technology
Full Research Guidelines are:
`https://jvgemert.github.io/ResearchGuidelinesInDL.pdf`

1 Research meeting questions

Questions during a research meeting often are specific, special case instantiations of more general questions. It is not always clear what goal the specific questions have, which might confuse, frustrate, or even demotivate. Here, I aim to give the generalized questions that lie behind specific question types, with the goal of making it clearer why, and what the goal of a question is. Additional benefits might include that these questions can be asked even without a meeting to get unstuck, and might help in preparing and focusing the meeting. Questions are grouped per stage; where each stage is typically revisited often, see Fig **??**.

Research meeting questions		
Research question	**Related work**	**Method/approach**
RQ1 Why interesting?	RW1 Who will use it?	MA1 Why this method?
RQ2 What storyline?	RW2 How different?	MA2 Explain each step?
RQ3 Formalize/simplify	RW3 Builds on what?	MA3 Formalize/simplify
RQ4 What problem?	RW4 Baselines?	MA4 Alternatives?
		MA5 Align with RQ?
Experimental setup	**Analyzing**	**Conclusions**
ES1 What Qs?	AR1 Validate?	DC1 Exhaustive?
ES2 How answer Q?	AR2 Baseline?	DC2 Expectations?
ES3 Baselines?	AR3 Understand all?	DC3 Align with RQ?
ES4 Expected outcome?	AR4 When fail?	DC4 Simplify?
ES5 Simplify?	AR5 Link to Q?	DC5 New hypotheses?

1.1 Questions about the Research Question (RQ)

RQ1: Why is the RQ interesting? Why do you care? Who else cares? Why should others care? Should we change the RQ? It can be useful to rephrase the RQ to better align with the problem: it can become more general, or more specific. Changing the RQ is normal during the research process because finding the suitable RQ often takes a large part of the process.

RQ2: What is the 10-15 bullet point main storyline? Does the story still make sense? Which point in the storyline are we now discussing? Is the point still valid? Should we change the storyline? The storyline is the key motivational driver and changing the storyline is common during the process because any new result can invalidated other parts of it.

RQ3: What is the RQ precisely? (Formalize/Simplify) Why can't we ask a simpler question? How to formalize it? Which part is the most uncertain? Should we remove that part or should we focus on that first? Formalizing the RQ makes it more precise, possibly revealing hidden assumptions. Simplifying the RQ can make it more general and/or easier to explain.

RQ4: Does the problem exist? How often? When (not)? What is the simplest example of the problem? How to convincingly demonstrate that the problem exists? Creating a simple, fully controlled, setting of the problem reveals hidden assumptions. It simplifies and gives focus. Coming up with a good problem example is difficult, and often takes several iterations.

1.2 Questions about Related Work

RW1: For which work is the RQ interesting? Why can't this approach be used for X? Why can't Y also make use of this approach? Should we change the research question to better facilitate X or Y? Don't wait for others see links to your work; actively link them yourself.

RW2: How is existing work different/similar? Why can't method X already answer the research question? What assumptions are different from method Y? What other related papers are there? Should we change the research question to clearly discriminate our setting from X or Y? Or should align better with X and Y? Motivate where the work fits, how it's different and how it's similar.

RW3: Building on what existing work? What is the motivation to build on component X? Why don't we use Y? Can't we use a simpler or more common building block? If the building block is not the focus of the work, then it should be as standard as possible.

RW4: How does existing work (baselines/competitors) work? What do baselines/competitors do? Why? How do they solve part X of the research question? How are we different? Why are we different? There should be good motivation to do something different.

1.3 Questions about the Method/Approach

MA1: Can you motivate why this method? Each step? Why not another method? Is each step essential and why? (is there evidence? ablation?).

MA2: What's going on? (Visualize output for each step) What is it really doing? Please explain this step in detail? Please show only the outcome of this part, and keep the rest constant. The goal is to validate that the method is behaving as expected.

MA3: What precisely? (Formalize/Simplify) How to make it simpler? How to formalize (math?) to describe what exactly is going on? Simpler is stronger.

MA4: Does method make sense? (Alternatives?) What are other options? Can we motivate why we do not use them? These choices might not be 'obvious' and may require an empirical experimental ablation.

MA5: How well is the alignment of the method with the RQ? Match the method with RQ: Which part of the RQ aligns with this step? Match the RQ with the method: Where does this part of the RQ come back in the method? Does it do what we think it does, and if not, should we then change the RQ to match this? It can be a game changer if results interestingly deviate from expectations.

1.4 Questions about the Experimental Setup

ES1: What empirical questions belong to the RQ? What questions do we wish to answer? Why? What would answering this question give? Do the questions align with the main RQ? Should we change the RQ to align better?

ES2: What exact question is answered by this experiment? Why does this experiment answer this question? What other experiments are possible? Which experiment to do? Each experiment should test an hypothesis or answer a question.

ES3: What are relevant baselines? Are there very simple (non learning? or simple averaging?) baselines to compare to? To which (existing) methods do we compare? Why? Why not more/less? Should we make the RQ tighter?

ES4: What exact outcome is expected? What outcome is wanted? Before running the experiment, answer what outcome you would like? What are the exact numbers that you expect as an outcome? Does doing the analysis on those numbers give the wanted outcome? If not, should we change the RQ?

ES5: What is the minimal setting? (Simplify) How to use a smaller problem? How to use a less complex setting? Why can't part X be removed from this setting? Simpler is stronger.

1.5 Questions about Analyzing Results

AR1: How to validate results? How to validate there are no bugs? How to validate if your method does what you claim it does (semantic debugging)? Do we have stddevs? Do results consistently align with our previous results? Can we do a small test to validate? Can we do an independent experiment to validate? Bugs are normal, and neural networks are notoriously difficult to debug; they might seem to work, but there might still be a mistake. Start with the assumption that there is a mistake somewhere, and then write code to find it.

AR2: How to verify correctness of baseline? Do not assume baselines directly work. How well do we match the reported results in their paper? Are these results expected? How to best optimize (hyperparameter tuning) the baseline so that the baseline is fairly evaluated?

AR3: Do we understand all results? When looking at the result table, can we explain each pattern in the table? Can we also look at some individual data samples?

AR4: When does it fail? Can we systematically predict when it fails? Can we look at some individual mistakes? Do these failures make sense?

AR5: Do results answer the question of the experiment? Each experiment has a question to answer. What was the question? Do these results answer that question?

1.6 Questions about Drawing Conclusions

DC1: What are all conclusions we can draw? Can we list all patterns that we see? Can we explain all patterns? Are there patterns that we have missed?

DC2: How well do results align with previous expectations? What were the previous expectation? How to explain deviations? Are all results internally consistent?

DC3: How well do results align with RQ? What results did we want? How much are these results what we wanted? Should we change the RQ? Should we redo a different variant?

DC4: Is there a simpler experiment with same conclusions? Which properties are not essential? Which properties should be more emphasized? Results/patterns that are not relevant distract from the main message.

DC5: To which new hypotheses lead these results? How well do these align with the RQ? Should we change the RQ to include these new hypotheses or write them as future work?