# peer review 1: Cohortney on OMS

Taras Khakhulin, Mikhail Pautov

April 23, 2021

Introductory note: indices of questions to consider refer to the ones from the corresponding assignment.

## Review

1. The problem of creating of the framework of modular structure for one-shot multi-speaker text2speech models is well defined. From the statements of work it is understood how exactly the problem is solved now (as a three-step pipeline) and what are the challenges to be faced (multi-voice scenario, one-shot learning perspective). Actually, the problem was discusses with the guys as initially we shared a project with them. Overall, the common thing, which is up to guys, is to add or not to add more advanced solutions (in terms of models) which are not presented in existing alternative repositories.

2. Proposed baseline solution is relevant to the problem of OMS; more than that, it is present in alternative repo what makes it convenient to be used as an example to construct an approximate architecture of the framework. Still, FastSpeech may be considered as a probable improvement (again, up to the guys).

3. It is well understood how Nikolay and Gleb will test their solution – as well as existing solution of the initial problem are *compilation* of models, as of blocks, it is expected to have reproducibility of baseline experiments as well as module-wise testing within the single API.

4. The section *Our Improvement* that was made by the guys make it clear how they are gonna work so far; initial guesses on what to do next – (i) implement a baseline solution with simple experiments, (ii) introduce the modular structure and (iii) start wrapping all it into the API.

5. Overall, the report is clear and transparent – it is stated what's the problem to solve, what are the existing solutions, with cons. and pros., its prospective improvements as well as workload distribution.

6. The project repo has quite a simple structure right now, so it is very easy to follow.