

## Lamedoc. Next stage.

At first, I wrote a long text with a bunch of technical peculiarities, graphs and equations, which maybe can be interesting und useful to read. However at the last moment I decided to concentrate your attention on the main problem we have. So this essay is very terse, with the only one diagram. All pure scientific parts I moved to another document.

## LameDoc for the DealRoom imlementation

Being an universal text meaning extractor, Lamedoc can be implemented for DealRoom workflow as follows(diagram.1). So, at the moment we have a pre trained M&A specific model (though not with a big enough corpus, but it will be fixed soon, as we get more texts). Recently it has shown not so bad performance and we can go along with it. Also we have an "Request-answer inference module" which is a set of the simple weighting function, and we can play around with them.

However, to start the real experiments and tuning the cooperation between "Requist-answer inference module" and "Machine learning engine" as a whole, we need the real examples of "Requests". As we know, all models should be evaluated on the real data, without such kind of evaluation we do not know where we are, and where to go. We have to somehow evaluate the real DealRoom scenario with real data. They can be fake, emulated by human data. I see the solution like this. Someone who knows the M&A field could write them by hand. One hundred of such kind of request will be enough to start with. They also should be intermixed with those that DealRoom has yet in DB. Also I can try to generate texts using the recurrent neural network, but the result is unpredictable. In any case to start it I need some real corpus of requests as the teaching example.

So as it seems to me, here we can bump at the "chicken and egg" problem. To collect more data and improve algorithm performance, the feature must be useful to our clients so they would started actively use it and leave more requests. To make the feature useful the algorithm must be trained on the requests.

So the only, thing that stops us from starting the testing of the real DealRoom work scenario is the absence of the real requests and data. I would like to hear your opinion on the matter.

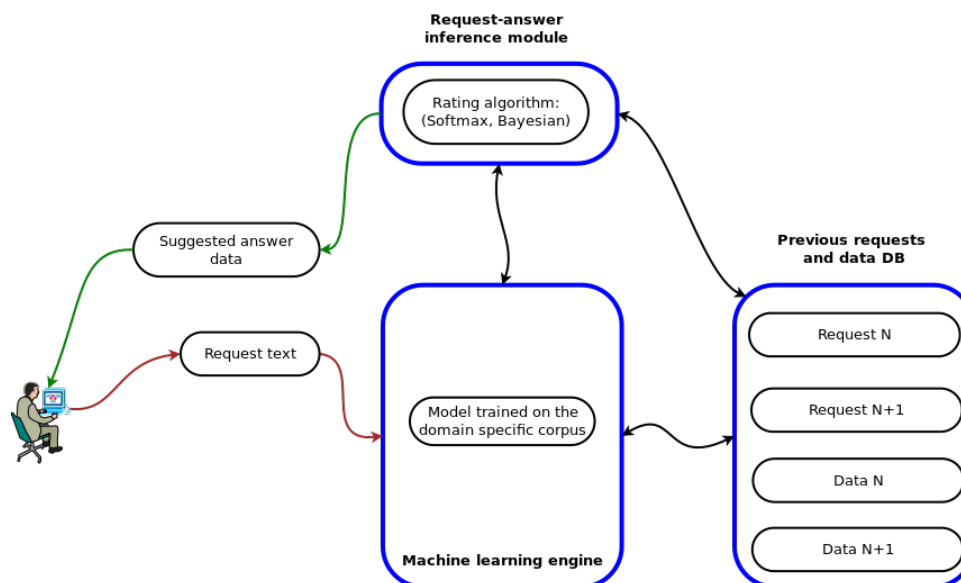


Diagram.1

## **Conclusion**

So the only, thing that stops us from starting the testing the real DealRoom work scenario is the absence of the real requests and data. I would like to hear your opinion on the matter.