

Innovation credit scan - pr.co

Describe the project in a short summary

The goal of this project (called Pulse) is to create an engine that helps companies to communicate better.

Pulse does this by analyzing the text of all the news articles published in a particular industry (e.g. Tech, Travel or Automotive) and stores the analysis in a database on which we can build different applications that help improve communication.

The Pulse database will have information on all articles published by our pre-defined list of publishers in each industry, what the articles are about, understands who the author is, where they're from and how to contact them. This way we will have an always up-to-date graph of the media landscape.

Text classification will be brought in house using the latest deep learning techniques and tools. The aim is to use TensorFlow to experiment and train a Convolutional Neural Networks for article classification for each industry. Keyword and entity extraction and enrichment will also be brought in house using Google's SyntaxNet, Spacy and datasets from various sources including CrunchBase, index.co, wikimedia and dbpedia.

Planned applications are:

- **Match:** analyzing your message on topics and keywords in order to match it with similar articles in our database. Via these matching articles we can provide you with a top 20 list of relevant journalists, influencers and experts.
- **Precognition:** show related topics and articles real-time while you are writing; this will help you improve the quality of the news you generate and understand its significance in the industry.
- **Clipping monitoring:** monitoring all articles going indexed by pulse on mentions of our customers. Once we find a clipping related to a customer, we ask them to review it and add it to their reports.

- **Trend research:** based on chosen strategic themes, Pulse will be able to provide research on articles that cover your themes so that your team can plan campaigns around these topics. Using the data extracted from SyntaxNet and Spacy we will generate a tag cloud, keep a tally of them and rank them in real time using algorithms similar to hacker news and reddit.

Which activities are scheduled?

The following activities will be necessary to achieve the project's goal:

- **Pulse Engine programming:** the majority of this project is programming and development of the Natural Language Processing (the Pulse Engine) and the algorithms that drive this. The engine part will focus on gathering the news articles through various sources, analyse it using NLP, extracting the author's and publisher's information.
- **Application development programming:** for each of the applications on top of the Pulse engine there will be a distinct sub-project to develop customer facing applications that perform the said functions as outlined above.
- **Project management:** the development and alignment of the different components will need to be managed between the development teams as well as coordinated with the marketing work involved.
- **UX Design & Customer research:** all work required to create a great interaction design. This includes the UI design (including customer research) and to research correct UX as well as the product market research to determine feature value and prioritization.
- **Marketing:** all activities related to the successful launch of each of the various Pulse applications to existing and new customers. For each application there might be a different target customer and value proposition so each launch would have a different marketing approach.

Which risks do you foresee?

Technically

The Natural Language Part of the Pulse Engine needs to correctly identify keywords, concepts, entities, persons, companies and geographies based on our pre-defined taxonomy list. We are currently testing this with 3rd party API's for subfunctions of Pulse, for example, IBM's Watson for NLP and Clearbit for data enrichment. Once the concept of Pulse is proven we will be creating the matching algorithms ourselves to make it fit our purpose.

The risk there is that it could be hard to get the right accuracy because of the large batch sizes needed to use for training data. If our batch sizes are too small we might be overfitting our training data and the results will be below the desired standard.

Another risk would be that the methodologies we will choose to set up our NLP neural network will not produce the desired results. Right now we're looking to use Convolutional Neural Networks for article classification, this methodology is currently predominantly known for its image classification capabilities but there are successful examples implementing it for text classification from the community as well.

Commercially

The potential risk is that we don't create the trust for brands to switch from their current methods (PR agencies or expensive tools) or that our marketing channels do not reach the intended market. Another risk could be that customers don't feel comfortable with using research results provided by artificial intelligence and machine learning.

Financially

If the creation of the Pulse engine is more complex than expected we might need extra developers to reach the projected product milestones.

The various industries will have a different volume, complexity, and nomenclature of news. Scaling the system up to more industries could prove to be more expensive than anticipated, due to higher costs of computation and storage.

Describe the market, your commercial opportunities, and the business model.

We target the communications teams of SMB (small medium businesses) and SME (small medium enterprises). These are companies with over 25 people that have a communications director and a team for external communications.

We believe this market is massive and we are already seeing many communications teams moving their workflow to the cloud, we believe to have a very compelling value proposition. We will charge for the applications built on top of the Pulse engine. The planned pricing strategy will need to be tested and probably tweaked to better match the target market.

How will you finance the project?

Own means: None. We are currently close to breaking even but will need external financing to further grow our business through the above tool.

Private investments: €400.000 - currently raising an investment round to pursue the above-planned project.

Innovation Credit: €325.000 - ~45% of the total project costs.

Describe relevant experience, expertise and your track record.

Relevant experience and expertise

Pr.co is a collaborative tool for PR professionals and communications teams to get meaningful exposure for your company. We have helped hundreds of companies get exposure and streamline the planning, writing and media outreach workflow so that the companies could focus on what they do best: crafting their story.

Our team is filled with experts in their field: developers, marketers and communication strategists.

We are currently a partner of AWS and Softlayer receiving over €150.000 in terms of free hosting and computational credits from these two companies combined.

Trackrecord

pr.co has existed since 2009 as an independent company and has over time becoming a global player in the PR industry. Together with a number of other companies we are revolutionizing a space that is begging for change!

ABOUT PR.CO

PR.co is a collaborative tool for PR professionals and communications teams to get meaningful exposure for your company. We streamline your planning, writing and media outreach workflows so you can focus on what you do best: crafting your story.



pr.conewsroom