

# CA Technologies

“El Netflix Científico”

Alberto López Sánchez, Gerard Otero, Adrià Garcia, Ruben Ajenjo

# Problemas

Problemas iniciales:

- Leemos demasiado.
- Difícil recordar todo lo leído.
- Hay que estar a la última por lo que hay que leer más.
- Complicación para encontrar lo que quieres con precisión.

# Primera Solución

Crear un software que detecte el tema de un Paper antes de empezar a leerlo evitará las lecturas absurdas y haría recordar lo que has leído.

Software utilizará deep learning y word2vec para aprender a clasificar los papers según los temas que nos interesen. (Primer prototipo con framework magpie).

Recomendación de papers similares que traten sobre el mismo ámbito que un paper dado.

# Plan de trabajo.

- Tendremos citas con la empresa cada dos semanas para poner en común el trabajo que hemos ido haciendo y que ellos nos den el visto bueno.

## DISTRIBUCIÓN DE TAREAS.

Frontend web y Documentación. (Adrià)

Sistema clasificador, recomendador. (Alberto, Gerard)

Base de datos, recogida de datos (crawler). (Rubén)

# Base de datos

- Se guarda la estructura categorias/clusters/papers en una tabla.
- En caso de que el registro sea un paper se almacena la ruta donde está archivado en el sistema.
- Se almacena en otra tabla información básica de los papers.
- Se guarda un pequeño historial de papers visitados por el usuario para futuras recomendaciones al entrar al sistema.

# Siguientes pasos

- Sistema clasificador automático: word2vec + doc2vec para obtener topics
- Front end web: navegación en forma de árbol, tarjetas mostrando información básica del paper, papers relacionados cuando seleccionas uno.