

## Resumo do Traballo de Fin de Grao

### Introdución, interese e propósito

A vacinación, consistente na inoculación controlada de preparados compostos de axentes infecciosos desactivados para estimular a resposta inmunitaria do corpo, é unha práctica médica obrigatoria en plans de saúde de países desenvolvidos. Dende os seus orixes aló polo século X, en forma de variolización rudimentaria sen controis de seguridade, foi obxecto de gran investigación científica e desenvolvemento, ata o punto de chegar a existir un gran consenso na actualidade acerca da súa efectividade en previr enfermidades infecciosas [1][2].

Porén, o feito de que as vacinas estean compostas por patóxenos implica preocupacións. Por unha banda, as autoridades sanitarias e científicas deben regular estritamente os ensaios clínicos aos que someter novas vacinas, e facer cumprir as súas directrices de conservación e notificación de eventos adversos. Por outra banda, a existencia de síntomas secundarios, sumada a motivos relixiosos, de liberdade persoal, descoñecemento ou desconfianza, constitúen razóns suficientes para que algúns individuos rexeiten esta práctica [3]. Así pois, é imprescindible que todas as persoas teñan acceso á información que necesiten para formar o seu propio criterio, pois así poderán tomar mellores decisións para elas e para a sociedade.

Neste senso, é coñecido que os medios sociais convertéronse no que posiblemente sexa o maior vehículo de intercambio de opinións a nivel global. Grazas a eles, calquera persoa que o desexe pode compartir experiencias e novas relativas á vacinación en tempo real, xerando grandes cantidades de información de interese sobre o tema que aínda está infraempregada para campañas de concienciación.

O presente traballo pretende, entón, afondar na utilización práctica deses datos, que serán extraídos de Reddit, un medio social popular e cunha API ao respecto pública e liberal. Primeiramente aplicaranse técnicas de minería de grafos, coa finalidade de construír un grafo de coñecemento representativo desa marea de información. A continuación, como exemplo de aplicación, desenvolverase un prototipo de chatbot que, a partires da ontoloxía proporcionada polo grafo e técnicas de procesado de linguaxe natural, sexa capaz de interactuar con humanos sobre o tema, tendo consciencia da posición que toman os participantes no debate. Os chatbots, aparte de estar en auxe como asistentes en aplicacións web, protagonizan recentes liñas de estudo, e poden resultar ser un potente método de divulgación, debido a súa proximidade e plena dispoñibilidade ao usuario final.

### Obxectivos

En consonancia co propósito xeral do traballo, cabe definir unha serie de obxectivos a desenvolver:

- *Extraer automaticamente información sobre o caso de estudo de Reddit.* Empregando as API para tal efecto, integraranse os posts e comentarios de Reddit nunha representación amigable para futuro procesamento, realizando un filtrado preliminar por relevancia, tema e similares características no proceso. A intervención do usuario neste proceso debería de ser reducida.
- *Integrar e organizar os datos nunha base de datos relacional.* Modelarase e porase en marcha unha BD que dea soporte ao almacenamento coherente, normalizado e automatizado da información, así como a súa recuperación de maneira eficiente.
- *Modelar e implementar unha ontoloxía do dominio obxecto de estudo para soportar a operación dun chatbot.* Por último, os datos representaranse de forma que se poida deseñar con eles un bot capaz de participar en conversacións con axentes humanos acerca do debate da vacinación, tomando unha postura a favor ou en contra da práctica.

### Arquitectura

O sistema seguirá unha arquitectura modular e orientada a obxectos, polo que os diferentes compoñentes representaranse con clases, favorecendo a cohesión e reutilización. Estes compoñentes comunicaranse conforme ao modelo de repositorio, compartindo acceso a unha BD común. Este deseño ten vantaxes en aplicacións intensivas en datos respecto a outros alternativos, como un modelo cliente-servidor ou baseado en capas.

A continuación sintetízanse visualmente as relacións entre subsistemas, a través dun diagrama de bloques. Adicionalmente, indícanse nel algunhas ferramentas relacionadas co desenvolvemento de cada un.