**Llista de propostes de projecte**

Aquesta secció recopila un conjunt d'idees que poden servidor com a projectes finals de carrera per estudiants de la Facultat d'Informàtica de Barcelona.

L'objectiu de cada una de les propostes no és la de representar un enunciat tancat, sinó oferir pistes sobre diferents implementacions possibles que puguin inspirar als estudiants a modificar-les, combinar-les, reduir-les, ampliar-les o crear-ne de noves.

Com es podrà veure, moltes d'aquestes propostes giren al voltant d'estudis històrics o la validació de les dades, emmagatzemades en els sistemes de FamilySearch, respecta la realitat. El motiu, és que més enllà de la funcionalitat de cerca, la principal preocupació sobre aquestes dades és amb quin grau d'exactitud representen la realitat que les envolta.

A més a més, recordar que l'API de FamilySearch es troba sempre en constant moviment, i que per tant, és una bona idea revisar la viabilitat de cada proposta abans de decidir, amb total certesa, el projecte que es vol realitzar.

**Comparació popularitat de noms**

L'objectiu d'aquesta funcionalitat és comparar la popularitat d'un o més noms, en un període concret del temps i amb la possibilitat de fixar la regió demogràfica a consultar.

La idea principal és que donada la introducció d'un o més noms, el sistema busques el nombre d'instàncies de persones nascudes amb el nom especificat, en els deu anys anteriors o posteriors a la data indicada, per la regió geogràfica especificada.

D'aquesta forma, es podria comparar quin dels dos noms ha estat més popular, segons les dades de FamilySearch, any a any.

Al mateix temps, esdevé interessant permetre la cerca d'un sol nom per observar si certs esdeveniments històrics han pogut influenciar el nombre de nadons amb un cert nom. Per exemple, suposa l'elecció d'Obama com a president d'Estats Units un increment en el nombre de persones amb aquest nom durant els següents anys?

El segona raó de ser de l'eina es ajudar a decidir, per exemple, el nom dels fills d'una persona, comparant, d'aquesta forma, la popularitat actual dels noms que s'estiguin avaluant.

Possibilitats d'extensió:

* Ampliar la cerca a diferents països, on per cada país, es cercarà mitjançant el nom introduït localitzat. Per exemple, si l'usuari introdueix Alexander, la cerca a Espanya utilitzes el nom d'Alejandro o Alex. En cas de no trobar aquesta base de dades enlloc, sempre es podria crear una taula manual d'exemple, amb unes quantes llengües i utilitzar-la.
* Comparar instàncies de noms a Catalunya extrets de les dades de FamilySearch, amb la comparació real extreta del institut nacional d'estadística.

**Portal de cerca localitzat al Català**

En aquesta memòria ja hem parlat de la funcionalitat de localització habilitada per part de FamilySearch i encara que aquesta suporta la localització a la llengua espanyola, no ho fa per la catalana.

L'idea d'aquesta funcionalitat és oferir a l'usuari un portal de cerca en català sobre les dades de FamilySearch, que no només faciliti la comprensió de la cerca a aquelles persones que vulguin utilitzar el català, sinó que també localitzi, en la mesura que sigui possible, la resposta retornada per l'API.

Per exemple, la informació relativa a l'estat actual d'una persona: *living* o d*eceased*, podria ser mostrada com a *viva* o *difunta.* De la mateixa forma, es pot localitzar tot el contingut de la resposta, des del nom dels camps d'informació al contingut d'aquests en algunes situacions.

Possibilitats d'extensió:

* Restringir la cerca del portal a només Catalunya, amb la possibilitat de desactivar la funció. A més a més, oferir ajuda de refinació a la cerca, de cara a introduir les diferents províncies o ciutats, evitant que l'usuari introdueixi valors invàlids.
* Posar-se en contacte amb l'organització FamilySearch, per convertir la localització realitzada a la llengua Catalana, en una localització oficial acceptada pel sistema.

**Geolocalització d'un cognom en diferents nivells**

Aquesta funcionalitat pretén ampliar l'exemple programat en aquest projecte final de carrera.

L'exemple implementat només permet la visualització d'instàncies d'un cognom al nivell de país. Aquesta nova funcionalitat hauria de permetre, com a mínim, les visualitzacions a nivell de continent i estat o comunitat autònoma. Com més nivells de profunditat diferents poguessin ser utilitzats, millor.

L'usuari haurà de ser capaç de navegar per aquests diferents nivells de forma interactiva.

**La recerca genealògica i l'heràldica**

Aquesta funcionalitat vol relacionar les dues ciències que sempre han caminat de la mà, la genealogia i la heràldica.

El primer objectiu del projecte passaria per identificar quina font de dades podria ser utilitzada per obtenir els diferents escuts d'armes. Probablement, no existeixi cap API en línea de la que es puguin obtenir i caldrà realitzar un parcejat automàtic d'alguna pàgina web o aplicació d'escriptori.

Un cop es disposin dels diferents escuts d'armes, es proposa realitzar una implementació de cerca simple per veure, en els detalls d'una persona, l'escut d'armes del nom de família, conjuntament amb les seves dades personals.

En cas de que només es disposes d'un grup limitat d'escuts, sempre es podria restringir la cerca, als noms de família dels que es disposa l'escut d'armes o a regions concretes per les que disposem d'un elevat nombre d'aquests escuts.

Possibilitats d'extensió:

* Implementar una API que serveixi els diferents escuts d'armes i la nostre aplicació web l'utilitzi per complementar les dades de les persones trobades a l'API de FamilySearch.

**Projectes d'indexació**

Tot i que aquesta proposta no pretén interactuar directament amb l'API de FamilySearch, volíem realitzar al menys una proposta, que estès relacionada amb el procés d'indexació.

Existeixen dos processos d'indexació diferents, els que es realitzen sobre fitxers formatats i els que es basen en la transcripció d'imatges a través del software de FamilySearch. Aquesta proposta de projecte, es divideix en realitat, en dues diferents.

La primera, aconseguir accés a algun registre genealògic local o posar-se amb contacte amb alguna organització que vulgui pujar el contingut de registres formats, al núvol. Sobre aquest registre, implementar un sistema d'automatització que transcrigui les dades i les prepari per ser enviades a FamilySearch.

La segona possibilitat, és realitzar un programa que llegeixi les seccions de la imatge sobre les que s'ha d'extreure la informació i intenti informar a l'usuari, que no omplir automàticament, sobre el contingut dels camps.

Aquestes propostes esdevenen amb un alta probabilitat, bastant complexes, per aquest motiu, es prega a l'estudiant que realitzi un bon estudi previ sobre l'abast i viabilitat del que vol realitzar abans d'inscriure el projecte.

**La història de l'església mormona a través de FamilySearch**

En una de les primeres seccions de la memòria, s'ha exposat per sobre la història del col·lectiu de l'església mormona. Com s'ha pogut observar, aquesta ha estat marcada per nombroses expulsions de diferents territoris i conflictes.

L'objectiu d'aquesta funcionalitat seria intentar traçar una relació entre els diferents emplaçaments o seus principals del col·lectiu, en els diferents moments del temps i les dades emmagatzemades per FamilySearch.

És a dir, destaca FamilySearch per tenir dades sobretot d'aquelles zones en les que s'ha trobat especialment pressent en comparació a la resta de localitzacions?

**Diversitat geogràfica d'un cognom: Els nostres avantpassats**

Aquesta funcionalitat ha estat inspirada per la campanya, de l'agencia de viatges Momondo, anomenada "The DNA Journey". Recomanem la visualització del vídeo per entendre millor les motivacions darrera aquesta funcionalitat.

https://www.youtube.com/watch?v=tyaEQEmt5ls

Totes les persones creiem que som bastant autòctones del lloc on hem nascut, però els estudis d'ADN sobre els nostres gens ens poden deixar molt sorpresos i demostrar que gairebé tothom te una diversitat ètnica i geogràfica important en els seus gens.

L'objectiu d'aquesta funcionalitat, de forma similar a l'exemple implementat sobre l'evolució geogràfica d'un cognom, és trobar una forma eficient d'oferir a l'usuari un llistat del nombre o percentatge d'instàncies d'un cognom, per cada país i en diferents èpoques. Podria ser una bona idea, restringir la cerca als primers vint països més grans de cada continent.

El repte del projecte, no és només la consulta de les dades, sinó trobar una forma intel·ligent d'extrapolar l'influencia d'un cognom en una regió determinada, en comptes de basar-nos en el total d'instàncies com ha fet l'exemple implementat.

Es deix a decisió de l'usuari intentar realitzar una interpretació de la influencia d'un cognom en cada regió i evidentment, de forma aproximada. L'objectiu final és demostrar la diversitat geogràfica dels nostres possibles avantpassats o relatius.

Aquesta proposta pot ser complicada degut al control sobre el temps d'execució. En cas de no veure viable realitzar una proposta que extregui les dades en temps real, sempre es pot realitzar un estudi específic sobre alguns cognoms d'interès i mostrar-ne els resultats.

**Les col·leccions de dades de FamilySearch**

L'objectiu d'aquesta aplicació es llegir les diferents col·leccions o fonts de dades de FamilySearch i llistar-ne de quines regions i sobre quins període de temps contenen dades.

L'objectiu, és oferir als usuaris un portal d'informació que permeti, mitjançant la introducció d'un país, període de temps o les dues condicions, visualitzar les diferents col·leccions disponibles i la informació específica d'aquestes.

Per exemple, donada la introducció del país Espanya i el període 1500-2000, la funcionalitat hauria de llistar totes les col·leccions que contenen dades que compleixen aquestes condicions, quants registres totals suposa, quants d'aquests estan accessibles a FamilySearch i quans pendents d'indexar, etcètera, etcètera.

Aquesta funcionalitat pretén respondre a un dels problemes principals de l'API i és la falta d'informació sobre la informació emmagatzemada. D'aquesta forma, mitjançant un petit anàlisi previ, podríem esbrinar si la informació que desitgem es probable que existeixi o no, sense perdre el temps en l'exploració de registres.

**FamilySearch i la segona guerra mundial: Natalitat i defuncions.**

Durant el transcurs del temps han succeït un gran nombre d'esdeveniments que han afectat a la població mundial de diferents formes. Un dels conflictes que ha causat més reporti ha estat la segona guerra mundial.

L'objectiu d'aquesta proposta de projecte és observar el impacte que va tenir aquest esdeveniment en les tasses de natalitat i defuncions, dels diferents països implicats, durant els anys del conflicte. Es recomana ampliar la finestra de temps estudiat més enllà dels anys del conflicte per observar quins eren els valors normals, previs i posteriors, a l'esdeveniment.

L'estudiant haurà de trobar la forma d'escalar les dades de cada país segons el volum de registres disponibles.

Una altre tasca que pot realitzar l'estudiant, és comparar els valors obtinguts a través de FamilySearch amb les dades oficials de la segona guerra mundial i respondre preguntes com: Es corresponen els països amb un increment de defuncions més elevat amb els que van patir més durant la segona guerra mundial?

**FamilySearch i la segona guerra mundial: Increment en els casaments**

Un estudi realitzat per Randal S. Olson, demostra que quan la segona guerra mundial va esclatar als Estats Units, es va registrar un increment enorme del nombre de casaments al llarg del país.

L'objectiu d'aquesta funcionalitat és intentar replicar aquest estudi pels estats units i estendre'l als països més implicats en la segona guerra mundial.

http://www.randalolson.com/2015/06/15/144-years-of-marriage-and-divorce-in-1-chart/

**FamilySearch i la segona guerra mundial: La llista de Schindler**

Aquesta proposta pretén realitzar un estudi sobre un dels col·lectius que es va veure més afectat durant la segona guerra mundial, els jueus.

Aquesta proposta de projecte ofereix a l'estudiant, parcejar els cognoms coneguts dels que van formar la llista de Schindler, i realitzar un estudi d'aquests sobre les dades de FamilySearch.

El projecte pot intentar respondre preguntes com Van disminuir en gran quantitat el nombre de registres amb els cognoms indicats? Van realitzar aquestes persones un procés d'emigració a diferents indrets del món? Etcètera, etcètera.

**La gran recessió o altres esdeveniments històrics.**

Aquesta funcionalitat s'aprofita del fet que es disposa d'un nombre major de dades per Estats Units que par la resta de països.

L'objectiu, estudiar els impactes de la gran recessió en la població d'Estats Units a través de les dades de FamilySearch i respondre a la pregunta de si la realitat observada és semblant a la que va succeir en realitat.

Es suggereix a l'estudiant que estudií el nombre de morts i emigracions al voltant d'aquest període com a punt d'entrada.

També volem aprofitar aquesta proposta per suggerir als estudiants que qualsevol fet històric, amb un cert impacte, pot intentar ser estudiat a través d'aquesta API. Per aquest motiu es convida als estudiants a plantejar els seus propis cas d'estudi.

**Comparacions amb els amics o seguidors de Facebook i Twitter**

L'objectiu d'aquesta funcionalitat es proporcionar a l'usuari, mitjançant les API de FamilySearch i Facebook o Twitter, un conjunt d'eines per tal de comparar les dades d'uns i altres.

El nombre de comparacions o eines de cerca a FamilySearch que podrien ser utilitzades és il·limitada, però per donar alguna idea als estudiants del que tenim en ment, en citem algunes a continuació:

* Quin dels nostres amics o seguidors te el nom o cognom més popular en el país on viuen o arreu del món?
* Quin dels nostres amics és més probable que tingui ascendència als estats units (per instàncies del cognom en el país)?
* L'esperança de vida mitjana de les persones enregistrades a FamilySearch amb el mateix nom que algun dels nostres amics i diferents comparacions.

Les possibilitats són realment il·limitades i a més a més, l'aplicació podria permetre accions com, seleccionar només els amics que ens interessin per certes consultes, realitzar una cerca per localitzar de forma fàcil als nostres amics, etcètera.

**Comparació de dades genealògiques reals amb FamilySearch**

Aquesta proposta de projecte pretén validar, en certa forma, com de bé representen les dades de FamilySearch la realitat d'un país o països a través de diferents èpoques, o per una època determinada.

Aquesta proposta podria ser dividida en molts projectes diferents, un per cada país, concepte o època que l'estudiant vulgui explorar. Per ajudar a comprendre als estudiants, al que ens estem referint, a continuació citem una sèrie d'exemples:

* Comparació de mortalitat per infants: Global, per continents, països, èpoques, etcètera.
* Mitja de fills per família: Global, per continents, països, èpoques, etcètera.
* Proporció d'homes i dones: Global, per continents, països, èpoques, etcètera.
* Esperança de vida per les dones i homes: Global, per continents, països, èpoques, etcètera.

Aquesta proposta no pretén que l'estudiant abordi tots els conceptes diferents que es pugui imaginar, però si aquells que cregui que poden tenir més valor de cara a comparar entre diferents regions i èpoques.

Segons els factors a estudiar, la complexitat de com hauran de ser processades les dades canvia, i per tant, caldrà tenir-ho en compte a l'hora de definir l'abast del projecte.

Es recomana també als estudiants que cerquin estudis estadístics sobre la població del món, per poder inspirar-se de cara al plantejament de propostes de projecte.

**Estudi de profunditat dels arbres familiars**

Aquesta funcionalitat es basa en l'estudi de les genealogies disponibles a través de FamilySearch.

Es planteja a l'usuari un estudi de la profunditat d'aquestes, quantes generacions diferents estan emmagatzemades de mitjana, com de complertes solen estar les dues bandes de l'arbre familiar, etcètera.

Un segon conjunt de dades que l'estudi pot intentar respondre és quins són els períodes de temps en els que era més probable mantenir un arbre genealògic. S'està perdent la tradició? Ha anat en augment durant els últims anys?

**Algoritme per marcar duplicats**

FamilySearch te una funció que permet, donada una persona, aconseguir les persones de l'arbre que tenen una alta probabilitat de ser un duplicat.

L'objectiu d'aquesta aplicació, seria realitzar un algoritme, que donada una persona, busques en les persones marcades com a candidates a ser un duplicat, i avalués si aquesta marca és correcta o no.

L'estudiant podria cercar i comparar segons la diferent informació disponible de cada persona i emetre una conclusió final com per exemple:

* Insuficient informació sobre el possible duplicat
* Duplicat descartat per inconsistència en els naixements
* Duplicat real amb coincidències de dates de naixement.
* Etcètera

Un altre aspecte que el projecte podria intentar atacar és comparar la fiabilitat d'aquest algoritme amb el que utilitza FamilySearch per marcar duplicats.

Aquest projecte pot resultar bastant compleix, i a priori, és desconeix la precisió o condicions sota les que FamilySearch marca a una persona com a duplicada. Per tant, es recomana realitzar un bon estudi previ d'aquests conceptes abans d'embarcar-se en el projecte.