

GRAU EN ENGINYERIA INFORMÀTICA

ESPECIALITAT EN COMPUTACIÓ

INFORME DE SEGUIMENT

Construcció d'un dataset per a l'extracció de models de processos de negoci a partir de textos

GUILLEM PLA BERTRAN

Director
JOSEP CARMONA VARGAS

Codirector
LLUÍS PADRÓ CIRERA

Tutora del GEP
OLGA PONS PEREGORT

24/03/2022



UNIVERSITAT POLITÈCNICA DE CATALUNYA
BARCELONATECH

Facultat d'Informàtica de Barcelona



Índex

1	Introducció	3
2	Context	4
2.1	Actors implicats	4
2.2	Business Process Model and Notation	4
2.2.1	Elements de BPMN	5
2.3	Problema a resoldre	8
3	Justificació	11
4	Abast	13
4.1	Objectius	13
4.2	Possibles obstacles	14
5	Metodologia	15
5.1	Mètode Kanban	15
6	Planificació temporal	17
6.1	Planificació alternativa	17
6.2	Canvis als objectius	17
6.3	Canvis als costos	18
6.4	Punt de la planificació en què es troba el projecte	18
7	Anàlisi d'alternatives	19
7.1	Parser BPMN	19
7.2	FreeLing	19
7.3	SimpleNLG i jBPT	20
8	Integració de coneixements	21
9	Lleis i regulacions	22
	Bibliografia	23

Índex de figures

2.1	Representació dels diferents tipus d'Esdeveniments. Font: Elaboració pròpia.	5
2.2	Representació dels diferents tipus d'Activitats. Font: Elaboració pròpia. .	6
2.3	Representació dels diferents tipus de <i>Gateways</i> . Font: Elaboració pròpia. .	7
2.4	Representació dels diferents tipus de Connectors. Font: Elaboració pròpia.	7
2.5	Representació d'una <i>Swimlane</i> . Font: Elaboració pròpia.	8
2.6	BPMN d'exemple. Font: camunda.com	9
2.7	Descripció del BPMN d'exemple. Font: Elaboració pròpia.	9
4.1	Subobjectius de cada part del projecte. Font: Elaboració pròpia.	13
5.1	Metodologia Kanban. Font: Pixabay.	16

1

Introducció

El diccionari de l'Institut d'Estudis Catalans [1] defineix un procés com una manera de descabdellar-se una acció progressiva. És a dir, que qualsevol conjunt d'accions que realitzem en el nostre dia a dia es pot definir com a un procés. De la mateixa manera, en una empresa es duen a terme una llarga llista d'accions que permeten aconseguir els objectius marcats.

Evidentment, en una empresa els processos no són tan simples com els que realitzem nosaltres diàriament. En conseqüència, és molt important que cada una de les tasques estigui ben definida.

La definició formal de processos de negoci en estàndards com BPMN 2.0 és un pas crucial per a la transformació digital de les organitzacions, ja que permet gestionar els processos que es duen a terme en una empresa. Moltes organitzacions tenen els seus processos documentats en text, i crear-ne els models formals té un elevat cost de personal especialitzat.

Automatitzar la creació de diagrames BPMN seria molt útil per a les empreses. Una manera per a aconseguir això seria utilitzar models de *Deep Learning*. Malauradament, per entrenar aquests models es necessita una gran quantitat de dades que no sempre és fàcil d'aconseguir. Per evitar l'escassetat de dades, en aquest treball es pretén crear un programa que generi un *dataset* de diagrames BPMN enllaçats amb la corresponent descripció en llenguatge natural.

2

Context

Aquest treball de fi de grau (TFG) del Grau en Enginyeria Informàtica (GEI) de l'especialitat de Computació es fa amb modalitat A (centre), i es realitza en la Facultat d'Informàtica de Barcelona (FIB). El projecte pretén aprofundir en la recerca per aconseguir una millora en la creació de models BPMN a partir de textos en llenguatge natural.

2.1 Actors implicats

Els actors implicats en aquest projecte són el personal del projecte, és a dir, els directors Josep Carmona Vargas i Lluís Padró Cirera, la tutora del GEP Olga Pons Peregró i l'autor Guillem Pla Bertran.

També hi estan implicades les persones a les quals va dirigit el projecte. Aquestes són totes les persones que fan recerca en el camp de NLP o els que treballen amb BPMN en el món empresarial, principalment *Business Analysts*. Segons la Northeastern University [2] aquestes persones són les responsables d'entendre què es fa en el negoci, quins passos cal seguir per a completar els objectius, avaluar els processos per eficiència, cost i resultat, desenvolupar plans del projecte, entre altres coses.

2.2 Business Process Model and Notation

BPMN vol dir Business Process Model and Notation, en català Model i Notació de Processos Empresarials, és un estàndard per a la modelització de processos empresarials[3]. Aquest proporciona una notació gràfica per descriure un procés en un Business Process Diagram (BPD). Està basat en una tècnica de diagrama de flux molt similar als diagrames d'activitats de l'Unified Modeling Language (UML).

L'objectiu de BPMN és donar suport a la gestió de processos empresarials, tant per als usuaris tècnics com per als usuaris empresarials, per això proporciona una notació

intuïtiva, però que alhora és capaç de representar una semàntica de processos complexa.

Recentment, la quantitat d'empreses que han començat a utilitzar BPMN ha crescut molt. Segons una enquesta del 2018 [4] el 64% dels negocis estan interessats en utilitzar aquest mètode per a simplificar els seus processos empresarials per a poder estalviar diners i millorar la productivitat.

A part de millorar l'eficiència hi ha altres motius que expliquen el perquè és necessari [5]. Augmenten la satisfacció dels clients, milloren la capacitat de resposta organitzativa i ajuden a complir amb noves regulacions, entre d'altres.

2.2.1 Elements de BPMN

Els models BPMN fan servir diagrames simples construïts a partir d'un conjunt d'elements gràfics. Els cinc elements bàsics que es tindran en compte en el projecte són els *Events*, les *Activities*, les *Gateways*, els *Connectors* i els *Swimlane*.

Esdeveniments

Un *Event* o Esdeveniment es representa amb un cercle i significa que alguna cosa passa durant el transcurs d'un procés. Poden tenir símbols que determinin el tipus d'esdeveniment que són. Principalment, es fan servir tres tipus d'esdeveniments:

- **Start event** o Esdeveniment inicial: És el que inicia tot el procés. Tots els processos comencen amb un esdeveniment d'aquest tipus.
- **Intermediate event** o Esdeveniment intermedi: Representa que alguna cosa passa entre l'inici i el final del procés.
- **End event** o Esdeveniment final: Representa el resultat final del procés. Tot procés té com a mínim un esdeveniment final.



Figura 2.1: Representació dels diferents tipus d'Esdeveniments. Font: Elaboració pròpia.

Activitats

Tal com explica Visual Paradigm [6], una *Activity* o Activitat és un "treball" que una empresa duu a terme en un procés de negoci. Les activitats poden ser atòmiques (Tasques) o descomponibles (Subprocessos). N'hi ha tres tipus bàsics:

- **Task** o Tasca: És una activitat atòmica dins d'un flux de processos. Es creen quan l'activitat no es pot desglossar a un nivell més detallat. En general, una persona o aplicacions duran a terme la tasca quan s'executi.
- **Sub-Process** o Subprocés: És una activitat composta que representa una col·lecció d'altres tasques i subprocessos. Per facilitar les comunicacions, no es vol que un diagrama de processos empresarials sigui complex. Mitjançant l'ús de subprocessos, és possible dividir un procés complex en diversos nivells, la qual cosa permet centrar-se en una àrea determinada d'un diagrama de procés.
- **Call Activity** o Activitat de trucada: És una activitat definida en un procés extern a la definició del procés actual. Permet crear una definició de procés que pot ser reutilitzada en altres definicions.

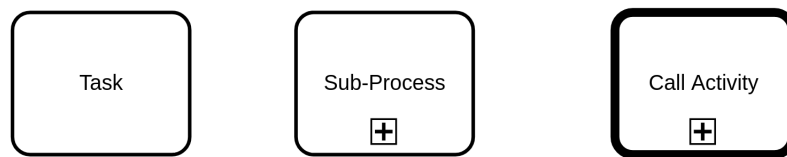


Figura 2.2: Representació dels diferents tipus d'Activitats. Font: Elaboració pròpia.

Gateways

Les *Gateways* o Portes d'entrada determinen quin camí s'empra a través d'un procés que controla el flux tant de fluxos de seqüències divergents com convergents. És a dir, una única porta podria tenir múltiples entrades i múltiples fluxos de sortida en els quals es permet o no l'entrada o la sortida. N'hi ha de diversos tipus, a continuació se n'explica tres dels més importants:

- **Exclusive Gateway** o Porta exclusiva: Se segueix un únic camí.
- **Parallel Gateway** o Porta paral·lela: Se segueixen tots els camins.
- **Inclusive Gateway** o Porta inclusiva: Se segueix un camí o més.



Figura 2.3: Representació dels diferents tipus de *Gateways*. Font: Elaboració pròpia.

Connectors

Els elements anteriors es connecten entre si en un diagrama per crear l'estructura bàsica d'un procés empresarial. Hi ha tres objectes de connexió que proporcionen aquesta funció. Aquests connectors són:

- **Sequence Flow** o Flux de seqüències: S'utilitza per mostrar l'ordre (la seqüència) en què es realitzaran les activitats en un procés.
- **Message Flow** o Flux de missatges: Representa el flux d'informació a través dels límits de l'organització. Els grups, les activitats i els esdeveniments de missatges es poden associar amb el flux de missatges. El flux de missatges es pot personalitzar amb un sobre que mostri el contingut del missatge.
- **Association** o Associació: Les anotacions permeten afegir més informació al diagrama que ajudi a documentar el procés.

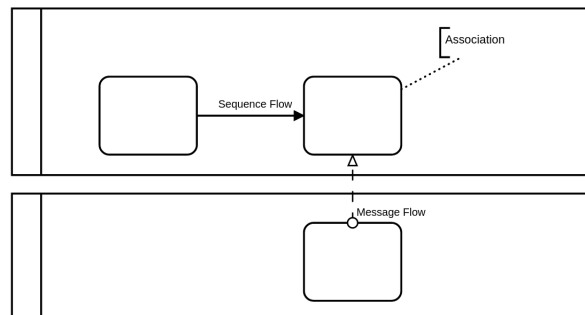


Figura 2.4: Representació dels diferents tipus de Connectors. Font: Elaboració pròpia.

Swimlane

En BPMN existeix un concepte anomenat *Swimlane*, aquest es divideix en dos tipus els *Pool* (en català, piscina), i el *Lane* (en català, carril). Un *Pool* representa un actor que pren part activa en un procés. Es representa amb un contenidor rectangular que pot contenir elements com els descrits anteriorment. Un *Lane* és una subdivisió del *Pool*, es fa servir per categoritzar activitats segons el rol o la funció de l'actor. En la figura 2.5 el *Bank* és el *Pool* i conté dos *Lane* que representen els rols de *Director* i de *Sales Department*.



Figura 2.5: Representació d'una *Swimlane*. Font: Elaboració pròpia.

2.3 Problema a resoldre

Com s'ha comentat anteriorment, el modelatge dels processos empresarials és una part vital a l'hora d'establir el funcionament d'un projecte. Moltes empreses tenen textos que expliquen com estan definits les accions a realitzar, però no ho tenen modelat en un llenguatge formal. El modelatge pot arribar a comportar fins al 60% del temps gastat en els projectes de gestió de processos.

Una possible solució per a automatitzar la creació dels diagrames a partir de descripcions en llenguatge natural és fer servir tècniques modernes d'Intel·ligència Artificial. En aquest cas concret es podria entrenar un model de *Deep Learning* a partir d'un conjunt de diagrames que tenen associat la seva descripció en llenguatge natural.

Lamentablement, tal com Matthew Stewart [7] explica, un dels reptes de l'aprenentatge automàtic és la manca de dades i la falta de bones dades. Aconseguir els conjunts de dades adequats és complex, ja que per motius de privacitat i confidencialitat molt poques

empreses alliberen les descripcions dels seus processos.

En aquest projecte es pretén abordar el problema de l'escassetat de dades. Es farà generant descripcions sintètiques, que semblin mínimament reals, i que permetin entrenar algorismes de forma fiable i controlada.

A continuació, en la figura 2.6 es mostra un exemple d'un model BPMN que representa el procés de preparació d'un àpat. Per a il·lustrar el que es pretén fer al treball el model està acompanyat d'una descripció en llenguatge natural generat a mà, però l'objectiu és aconseguir generar una descripció similar automàticament.

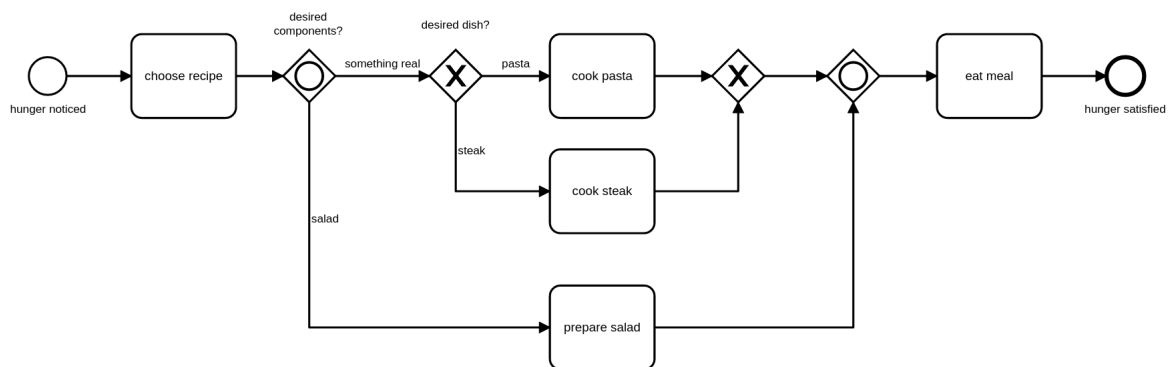


Figura 2.6: BPMN d'exemple. Font: camunda.com

When hunger is noticed, a recipe is chosen. Then the question 'desired components?' is asked.

If the answer is salad, then a salad is prepared.

If the answer is something real, then the question 'desired dish?' is asked.

If the answer is pasta, then pasta is cooked, otherwise if the answer is steak, then a steak is cooked.

Then the meal is eaten and hunger is satisfied.

Figura 2.7: Descripció del BPMN d'exemple. Font: Elaboració pròpia.

En aquest cas la descripció de la figura 2.7 és senzilla perquè l'exemple utilitzat ho era, però no sempre serà així. Un model pot comptar amb moltes més activitats i ramificacions. Això pot complicar la generació de la frase per què cal decidir quan

s'ha de fer un nou paràgraf. També cal estructurar bé el text perquè quan diverses ramificacions s'ajuntin s'entengui el conjunt sencer del paràgraf. També s'hauria de vigilar que no sembli un text molt artificial, però això ja és més complicat. Aquests reptes són els que s'intentaran superar en el transcurs d'aquest projecte.

3

Justificació

Donat que el modelatge de processos amb l'estàndard BPMN és molt específic l'estat de l'art no és molt extens. Tot i això, es poden trobar diverses aplicacions similars, encara que no fan ben bé el que es pretén en aquest treball. Aquests projectes fets anteriorment els podem classificar entre els que transformen d'un diagrama BPMN a text natural i els que ho fan a l'inrevés.

Abans, desenvolupar una eina que automatitzés qualsevol procés requeria molt esforç i dedicació, ja que, s'havien de tenir en compte totes les possibles combinacions en els possibles *inputs*. Això feia que en certs camps, com ara el de convertir un text natural, fos pràcticament impossible d'aconseguir. Per sort, mètodes com el *Deep Learning* permeten generar contingut de forma ràpida i eficient. Però, per aconseguir un model funcional es requereix un conjunt de dades molt gran i divers, i això no sempre és fàcil d'aconseguir.

Per a solucionar el problema plantejat en aquest treball no disposem d'un bon *dataset* que ens permeti generar un model de *Deep Learning* fiable. És per això, que la part principal d'aquest treball consisteix a crear una aplicació que generi un conjunt de dades variat i amb prou registres.

Això es farà convertint diversos diagrames a un llenguatge natural que sigui entenedor per qualsevol persona sense coneixement de BPMN. En aquest cas s'ha decidit que es farà servir l'anglès, ja que és un dels més utilitzats al món.

Per a fer aquesta primera part, hi ha un treball del 2016 titulat *Transformant Models BPMN a Llenguatge Natural* [8] que fa quelcom semblant. En el treball es llegeix el fitxer BPMN, després es crea un RSPT (Refined Program Structure Tree) per a guardar la informació necessària per solucionar el problema i també es fa servir una xarxa de Petri per a extreure l'estructura del procés. Finalment, es fa la generació del text analitzant cada element i concatenant els diferents missatges seguint l'estructura dels arbres creats anteriorment.

La segona part del projecte consisteix a entrenar una xarxa neuronal que sigui capaç d'interpretar una descripció en llenguatge natural i que ho converteixi en un model BPMN. Per a fer això, es farà servir el dataset que hem creat a la primera part. De treballs anteriors i aplicacions existents que converteixin de text a diagrames n'hi ha algunes que s'esmenten a continuació. Però cap d'elles fa servir un model de *Deep Learning*. Per tant, en aquest cas caldrà entrenar el model des de zero.

Process Talks [9] és una aplicació web que permet modelar processos a partir d'una explicació informal, sigui escrivint o parlant. L'eina fa servir regles i patrons que analitzen l'estructura sintàctica de les frases i treuen informació de qui realitza cada acció, o quina acció precedeix una altra. Aquesta implementació és costosa de mantenir i no és escalable, ja que, és complicat traspasar-la a nous idiomes. Per això, l'objectiu del projecte és crear una alternativa basada en *Machine Learning*.

A part d'aquesta aplicació web, també podem trobar un treball de recerca anomenat *Process Model Generation from Natural Language Text* [10]. Podem veure que es combinen algunes eines existents en el processament del llenguatge natural per a obtenir uns millors resultats. L'avaluació mostra que per a un conjunt de 47 parells de text-model de la indústria i els llibres de text, generen de mitjana el 77% dels models correctament.

Com s'ha pogut veure hi ha poques implementacions que serveixin de cara a plantejar una solució. Tanmateix, s'intentarà aprofitar el màxim de recursos disponibles per a reduir costos i temps.

4

Abast

4.1 Objectius

El principal objectiu del projecte és crear un conjunt de dades anotades que relacionin textos amb els models formals que els descriuen. Les dades han de servir per a ser usades per entrenar sistemes neuronals que donats un text en llenguatge natural el converteixin a un diagrama BPMN.

Com ja s'ha dit anteriorment a l'apartat de Justificació, el projecte es pot dividir en dos parts. La primera i principal és la **Generació del Dataset** i la segona és la d'**Entrenar un model** de Deep Learning com a demostració que el *dataset* és útil. En la figura 4.1 es pot veure les diferents parts del projecte i els subobjectius que té cadascuna d'elles.

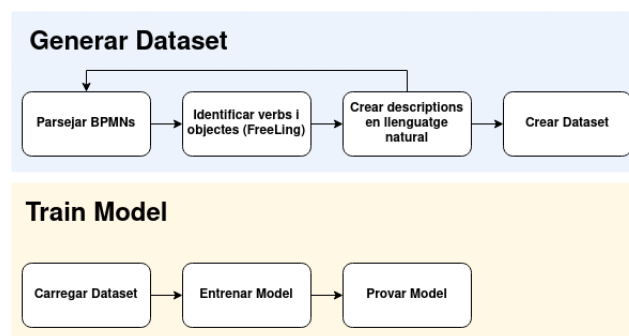


Figura 4.1: Subobjectius de cada part del projecte. Font: Elaboració pròpia.

En fase de generació del dataset el que es busca és tenir un conjunt de dades amb suficients registres i amb descripcions prou variades que serveixi per a poder entrenar després un model.

Per a generar les dades cal crear un programa que llegeixi (en anglès *parse*) els elements

d'un diagrama BPMN i emmagatzemi la informació necessària en un fitxer JSON. Caldrà extreure i modificar els elements que es necessitin. Un cop fet això només caldrà fer la tria dels atributs que es volen. Després s'identificaran els verbs, objectes i complements de les frases de cada element utilitzant un software d'anàlisi lingüístic anomenat *FreeLing* [11] [12] [13] [14] [15]. Tot seguit, es procedirà a crear les descripcions en llenguatge natural, utilitzant un software de generació de llenguatge natural anomenat *SimpleNLG*[16].

Un cop s'hagi desenvolupat el programa anterior, simplement caldrà repetir el procés de generació de text per a cada un dels models BPMN i emmagatzemar el conjunt de les dades en un format adequat per a poder-ho utilitzar després.

La segona part del projecte és la de crear un model de *Deep Learning* que serveixi com a demostració que el conjunt de dades generat és utilitzable. En aquest pas s'haurà de carregar el dataset generat anteriorment, entrenar el model i provar que retorna el resultat que es vol. Si no s'ha obtingut un resultat satisfactori, es pot polir el model canviant paràmetres.

4.2 Possibles obstacles

És molt comú que quan es comença a fer un projecte es trobin obstacles i problemes que no s'havien plantejat. Quan això passa pot ser que el projecte s'endarrereixi molt. Per tant, s'ha d'intentar tenir en compte el conjunt de riscos que poden sorgir.

Un dels obstacles que pot aparèixer és que no es disposi de prou diagrames BPMN i, per tant, no poder crear un *dataset* amb gaires registres. També pot ocórrer que la utilització de llibreries no sempre funcioni correctament. En aquest cas s'haurà d'invertir molt més temps a crear alternatives. Tenir suficient informació disponible per a fer un projecte és important, i el fet que no es pugui trobar prou documentació sobre un tema és un risc a tenir en compte. A més a més, s'ha de tenir present que el temps és limitat, per tant, és possible que no sempre es puguin assolir tots els objectius marcats.

5

Metodologia

Inicialment, s'havia decidit fer servir la metodologia *Agile*, però finalment s'ha optat per la metodologia *Kanban*. Malgrat ser una metodologia vàlida per a desenvolupar no és la més adequada. Ja que funciona bé amb projectes on no hi ha un objectiu final definit, en canvi, en el nostre cas sí. El nostre objectiu tampoc és proporcionar una integració contínua. A més *Agile* no permet comprovar la feina en progrés fàcilment, però amb les taules *Kanban* és molt senzill tenir un control del progrés de cada tasca. *Kanban* també permet saber de forma molt clara quines són les tasques prioritàries. A continuació, es mostra una breu explicació de la metodologia escollida.

5.1 Mètode Kanban

Per a dur a terme aquest treball s'ha decidit fer servir la metodologia *Kanban*. Tal com s'explica en la web APD [17], és un sistema de producció efectiu i eficient que està dins del grup de metodologies àgils [18]. La base de les metodologies *Agile* és permetre una planificació iterativa, això implica descobrir requisits i desenvolupar solucions mentre s'avança en el desenvolupament. El principal tret de *Kanban* és l'ús de targetes visuals que ens ajuden a organitzar les tasques.

Les targetes estan definides en una web de gestió de tasques. I aquests es poden moure per definir el seu estat. S'han fet servir els següents estats:

- **To do:** Significa que la tasca encara s'ha de realitzar.
- **Doing:** Implica que la tasca s'està duent a terme.
- **Review:** Vol dir que la tasca s'ha realitzat, però el resultat pot no ser definitiu.
- **Done:** Significa que la tasca ja es dona per acabada

A més a més, la metodologia es basa a acabar la feina que ja s'ha començat i en evitar tenir diverses tasques a mitges. També permet prioritzar cada tasca, per a intentar fer



Figura 5.1: Metodologia Kanban. Font: Pixabay.

de la millor forma aquelles funcionalitats que es consideren més importants.

6

Planificació temporal

Durant el transcurs del treball es va decidir endarrerir la data de finalització del treball. Aquesta va passar de ser del gener del 2022 a l'abril del 2022. Aquest canvi es va produir perquè la compaginació del projecte, amb la feina i una assignatura de la universitat va ser impossible. Això implica que s'ha tingut més dies dels que es van calcular inicialment, però el nombre estimat d'hores dedicades totals és el mateix.

També hi va haver un endarreriment en la primera etapa del projecte, la de *parsejar* els models BPMN. En la planificació temporal inicial no es va tenir en compte que aquesta etapa podia tardar tant, ja que semblava una part fàcil del projecte. Es va intentar buscar eines o programes que llegissin els models. Malauradament, l'eina escollida va resultar ser poc fiable i es va haver de començar de nou. També es va intentar programar de zero un *parser* que llegixi cada model, però finalment es va decidir utilitzar la llibreria *Camunda*. Els motius d'aquesta decisió s'expliquen més endavant en l'apartat d'Anàlisi d'alternatives 7.

6.1 Planificació alternativa

A causa dels obstacles esmentats anteriorment es va haver d'adaptar la planificació temporal. Es va canviar la data d'entrega i es va decidir dedicar menys hores setmanalment. També es va ampliar el nombre total d'hores dedicades a la tasca de llegir els models BPMN. Això va comportar que la tasca d'entrenar un model de *Deep Learning* amb les dades del *dataset* se suprimís de la planificació.

6.2 Canvis als objectius

Com s'ha comentat prèviament, els objectius van variar lleugerament, ja que en la planificació inicial estava previst crear un model predictiu, però aquest objectiu es va haver de suprimir. Es considera que no és una gran pèrdua perquè no era l'objectiu principal, sinó una forma d'ensenyar la feina feta. Tot i això, l'objectiu principal era generar un conjunt de dades i aquest s'ha mantingut en la planificació alternativa.

6.3 Canvis als costos

No hi ha hagut canvis als costos finals del treball. Malgrat que s'ha desenvolupat el treball amb més dies no hi ha hagut un augment de les hores de feina i, per tant, els costos de personal s'han mantingut. Tampoc s'han incrementat els costos per software perquè no es tenia cap llicència mensual.

6.4 Punt de la planificació en què es troba el projecte

En aquest moment el projecte es troba en les últimes quatre setmanes de desenvolupament. Això vol dir que encara cal acabar de crear les explicacions en llenguatge natural per a cada BPMN, generar el *dataset*, acabar la memòria i preparar la presentació.

7

Anàlisi d'alternatives

En aquest apartat s'expliquen les decisions que s'han pres i el motiu de cada una d'elles.

7.1 Parser BPMN

Inicialment, es va pensar que seria millor desenvolupar el projecte en Python, ja que és més fàcil de fer servir i l'autor està més acostumat a utilitzar-lo. Per això es va començar a buscar *parsers* de BPMN que es poguessin fer servir en aquest llenguatge. Desafortunadament, no existeixen llibreries fiables i es va haver d'optar per un projecte de codi obert anomenat *python-bpmn-engine* [19]. Aquest no comptava amb una bona documentació i no feia exactament el què es volia i finalment es va descartar usar-lo.

Amb els directors del projecte es va proposar implementar un *parser* des de zero en Python, o bé fer servir la llibreria *Camunda*. Fer-lo de nou tenia els avantatges de ser en un llenguatge més tractat pel desenvolupador i que es podia dissenyar com es volgués. Per contra, implicava moltes hores de feina per a tenir una versió final. Finalment, després d'aconseguir instal·lar el *Camunda* es va fer servir aquesta llibreria. Està molt ben documentada i ja té molts mètodes creats. Per tant, els únics problemes van ser a l'hora d'instal·lar-la i d'aprendre el seu funcionament.

7.2 FreeLing

La llibreria *FreeLing* es va escollir perquè un dels directors va participar en el seu desenvolupament. Això implica que durant el treball s'ha tingut suport de primera mà per part d'un dels creadors. Ha sigut especialment útil quan s'ha trobat algun problema amb l'eina perquè ha sigut molt senzill de resoldre.

7.3 SimpleNLG i jBPT

En el cas d'aquestes dos llibreries també van ser propostes dels directors del projecte. Tot i que també s'ha estudiat si existeixen alternatives millors, no se n'ha trobat. En el cas de *SimpleNLG* també es va analitzar si era millor fer servir la llibreria en Python o Java. Al final es va decidir en Java, ja que era el llenguatge original i estava plenament documentat i desenvolupat.

8

Integració de coneixements

En aquest projecte s'ha aplicat coneixement que pertany a disciplines com enginyeria del software, tecnologies de la informació i sobretot de computació.

De tecnologies de la informació s'han aplicat les tecnologies necessàries per emmagatzemar, recuperar, transmetre i manipular dades. Això ha sigut molt útil per saber com emmagatzemar a disc la informació necessària de cada model BPMN, i també per a recuperar-la en les següents etapes del programa.

Per a dissenyar el programa també ha sigut necessari aplicar coneixements d'enginyeria del software. Gràcies a aquests s'ha sabut dissenyar, desenvolupar, mantenir i avaluar el sistema de software de forma fiable i eficaç.

I per últim, s'ha aplicat coneixements de computació que han permès desenvolupar algorismes eficients de forma efectiva. S'ha buscat sempre obtenir el millor rendiment basat en l'entorn d'execució i s'ha desenvolupat algorismes eficients que o bé generen els grafs dels BPMN o els recorren.

9

Lleis i regulacions

L'única llei que cal tenir en compte en aquest projecte són els drets d'explotació sobre el software [20]. Aquests estan contemplats en l'article 99 de la *Ley de Propiedad Intelectual*. En aquest article s'hi explica que només pot exercir els drets d'explotació del software qui tingui la consideració d'autor. Però, també s'hi recull que es pot dur a terme una cessió de llicència d'ús. Això implica que el creador d'un programa pot autoritzar a un altre utilitzar el seu programa sense perdre'n la propietat. En aquest cas, per evitar problemes legals, sempre es fa servir codis i programes que permetin explícitament el seu ús i modificació.

Bibliografia

- [1] Institut d'Estudis Catalans. Diec2. <https://dlc.iec.cat/>. Accedit el 26/09/2021.
- [2] Ashley DiFranza. What does a business analyst do? <https://www.northeastern.edu/graduate/blog/what-does-a-business-analyst-do/>. Accedit el 26/09/2021.
- [3] Wikipedia. Business process model and notation. https://en.wikipedia.org/wiki/Business_Process_Model_and_Notation. Accedit el 20/08/2021.
- [4] Beyond BPMN. Beyond bpmn. <https://beyondbpmn.com/>. Accedit el 24/08/2021.
- [5] Marin Perez. The beginner's guide to using bpmn in business. <https://www.microsoft.com/en-us/microsoft-365/business-insights-ideas/resources/the-guide-to-using-bpmn-in-your-business>. Accedit el 24/08/2021.
- [6] Visual Paradigm. Bpmn activity types explained. <https://www.visual-paradigm.com/guide/bpmn/bpmn-activity-types-explained/>. Accedit el 16/10/2021.
- [7] Matthew Stewart. The limitations of machine learning. <https://towardsdatascience.com/the-limitations-of-machine-learning-a00e0c3040c6>. Accedit el 15/03/2022.
- [8] Genís Martín Coca. Transformant models bpmn a llenguatge natural, 2016.
- [9] Josep Carmona, Lluís Padró, Marc Solé, and Sànchez-Ferreres Josep. Process talks. <https://www.processtalks.com/>. Accedit el 26/09/2021.
- [10] Fabian Friedrich, Jan Mendling, and Frank Puhlmann. Process model generation from natural language text. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 6741 LNCS:482–496, 2011.
- [11] Lluís Padró, Miquel Collado, Samuel Reese, Marina Lloberes, and Irene Castellón. Freeling 2.1: Five years of open-source language processing tools. In *Proceedings of 7th Language Resources and Evaluation Conference (LREC'10)*, La Valletta, Malta, May 2010.

- [12] Lluís Padró. Analizadores multilingües en freeling. *Linguamatica*, 3(2):13–20, December 2011.
- [13] Lluís Padró and Evgeny Stanilovsky. Freeling 3.0: Towards wider multilinguality. In *Proceedings of the Language Resources and Evaluation Conference (LREC 2012)*, Istanbul, Turkey, May 2012. ELRA.
- [14] Xavier Carreras, Isaac Chao, Lluís Padró, and Muntsa Padró. Freeling: An open-source suite of language analyzers. In *Proceedings of the 4th International Conference on Language Resources and Evaluation (LREC'04)*, 2004.
- [15] Jordi Atserias, Bernardino Casas, Elisabet Comelles, Meritxell González, Lluís Padró, and Muntsa Padró. Freeling 1.3: Syntactic and semantic services in an open-source nlp library. In *Proceedings of the fifth international conference on Language Resources and Evaluation (LREC 2006)*, Genoa, Italy, May 2006. ELRA.
- [16] Ehud Reiter. simplenlg. <https://github.com/simplenlg/simplenlg>.
- [17] Max Rehkopf. What is a kanban board? — atlassian. <https://www.atlassian.com/agile/kanban/boards>. Accedit el 18/12/2021.
- [18] Wikipedia. Agile software development. https://en.wikipedia.org/wiki/Agile_software_development. Accedit el 28/09/2021.
- [19] Nikola Tanković. python-bpmn-engine. <https://github.com/ntankovic/python-bpmn-engine>. Accedit el 23/03/2022.
- [20] Carmen Blázquez Rodríguez. Derechos de explotación del software. licencia de uso. <https://convelia.com/derechos-de-explotacion-de-un-software-licencia-de-uso>. Accedit el 22/03/2022.