



UNIVERSITAT POLITÈCNICA DE CATALUNYA
BARCELONATECH
Facultat d'Informàtica de Barcelona



Grau en Enginyeria Informàtica
Especialitat de Computació
Treball de Final de Grau

Learning Process Models from Images

INFORME DE SEGUIMENT

Aleix Dalmau i Amezua

Director

Josep Carmona Vargas

Codirectors

Lluís Padró Cirera

Enrique Romero Merino

Dimarts 29 de Maig del 2020
Barcelona

Learning Process Models from Images	0
0. Context	3
1. Introducció	4
1.1 Definició de conceptes	4
1.1.1 Procés de negoci	4
1.1.2 Flux de treball	4
1.1.3 Minería de processos	5
1.1.4 BPMN	6
1.2 Descripció del problema	7
1.3 Actors implicats	7
2. Abast del projecte	8
2.1 Estat de l'art	8
2.2 Objectius i sub-objectius	8
2.2.1 Objectes de flux	9
2.2.2 Tipus d'activitats	9
2.2.3 Tipus de gateway	9
2.2.4 Reconeixement de text	10
2.2.5 Objectes de connexió	10
2.2.6 Carrils	10
2.2.7 Artefactes	11
2.2.8 Construcció del diagrama i del XML	11
2.3 Obstacles i riscos	11
2.4 Metodologia i rigor	11
2.4.1 Mètode àgil	12
2.4.2 Eines	12
2.4.3 Validació del projecte	12
3. Planificació temporal	13
3.1 Definició de les tasques	13
3.1.1 Gestió del projecte	13
3.1.2 Treball previ	15
3.1.3 Elaboración del software	16
3.2 Gràfic de Gantt	17
3.3 Gestió del risc	17
3.3.1 Risc temporal	17
3.3.2 Implementació	18
3.3.3 Fiabilitat i robustesa	18
3.4 Anàlisi d'alternatives i decisions	18
3.5 Estat del projecte en la planificació	19
4. Gestió econòmica	21

4.1 Costos de personal per activita	21
4.2 Costos genèrics	22
4.2.1 Material	23
4.2.2 Desplaçament	23
4.2.3 Espai	23
4.2.4 Ús d'internet	24
4.2.5 Ús corrent elèctric	24
4.3 Altres costos	24
4.3.1 Contingència	24
4.3.2 Imprevistos	24
4.4 Resum	24
4.5 Control de gestió	25
5. Sostenibilitat i compromís social	26
5.1 Autoavaluació	26
5.2 Dimensió econòmica	26
5.3 Dimensió ambiental	26
5.4 Dimensió social	27
6. Integració de coneixements	28
7. Lleis i regulacions	29
Bibliografia	30

0. Context

L'alta competitivitat que existeix entre empreses i organitzacions en una societat regida pel capitalisme i el mercat lliure en un món globalitzat, ha fet que aquestes siguin cada cop més sofisticades i eficients ja no només per optimitzar guanys, sinó pel simple fet de sobreviure. Existeixen innumbrables mètodes per fer que una empresa sigui eficient, recollits així en l'enginyeria de negocis o també anomenat enginyeria empresarial. Aquesta, aplica coneixements científics, tecnològics i empresarials i optimitza l'ús dels recursos dels que es disposa ja siguin recursos humans, tècnics o informatius. És d'aquesta manera essencial per a una empresa gestionar de la millor forma possible els seus recursos si aquesta vol ser exitosa.

Dins de les organitzacions es descriuen processos de negoci. Aquests bàsicament són una col·lecció d'activitats o tasques que en una seqüència específica produeixen un producte o servei per a un client concret [1]. Exemples de processos de negoci poden ser la simple seqüència d'events que segueix l'acte de demanar i servir un cafè o processos molt més complicats en els quals hi intervenen moltes més entitats i factors com pot ser per exemple el procés de renovar tota la flota d'ambulàncies del Servei d'Emergències Mèdiques (SEM), on hi poden intervenir nombroses institucions públiques i empreses privades. És vital per a les empreses, doncs, tenir molt ben estructurats aquests models de negoci, ja que analitzant-los, es poden detectar colls d'ampolla i aplicar-hi millores d'eficiència pot significar importants estalvis de temps i/o de diners beneficiant notablement l'empresa en qüestió.

Aquest treball de fi de grau es situa en el marc de la Facultat d'Informàtica de Barcelona (FIB) i més concretament en el Process and Data Science UPC Group (PADS-UPC), grup multidisciplinari amb fronts oberts en els àmbits de la mineria de processos, la gestió de processos de negocis i la ciència de dades [2]. Donat que PADS-UPC es dedica a fer recerca en aquestes àrees aplicant les seves tècniques per desenvolupar plataformes open-source o col·laborant amb projectes industrials, aquest Treball de fi de Grau (TFG) també seguirà en aquesta línia de treball.

1. Introducció

Aquest informe correspon al lliurament de l'informe de seguiment del Treball Final de Grau. El document està basat en l'últim informe de la fita inicial. L'informe mostra l'estat del projecte exposat a data de 29 de Maig de 2020. A part de les modificacions i correccions pertinents que se li han fet al document de la fita inicial que es mostren en aquest document, també se li han afegit els següents apartats:

- Apartat 3.4: Anàlisi d'alternatives i decisions.
- Apartat 3.5: Estat del projecte en la planificació.
- Apartat 6: Integració de coneixements
- Apartat 7: Lleis i Regulacions

1.1 Definició de conceptes

A continuació es defineixen els termes i conceptes propis del tema que és objecte d'estudi. Els primers punts serveixen d'introducció per finalment entendre l'últim punt que defineix l'objecte d'estudi d'aquest treball, justificant i explicant quines són les seves utilitats.

1.1.1 Procés de negoci

Ampliant la breu definició que s'ha fet a la introducció, un procés de negoci, a més, és sovint representat com un diagrama de flux d'una seqüència d'activitats amb punts de decisió. La principal avantatge dels processos de negoci és oferir una gran agilitat per reaccionar als constants i ràpids canvis del mercat.

Un procés de negoci comença amb un objectiu i acaba quan aquest objectiu es compleix, generant així un benefici per al client. A més, un procés de negoci complex es pot dividir en subprocessos, que tenen els seus propis atributs, però també contribueixen a complir l'objectiu general de l'empresa. L'anàlisi dels processos de negoci típicament inclou el mapeig del modelat de processos i sub-processos fins al nivell d'activitat o tasca. Aquests processos, a l'hora, es poden modelar a través d'un gran número de mètodes i tècniques. Al final, tots els processos formen part d'un resultat àmpliament unificat, que com s'ha dit abans és la creació de valor per al client. Aquest objectiu s'accelera amb la gestió dels processos de negoci que tenen com a objectiu analitzar, millorar i promulgar els processos de negoci.

1.1.2 Flux de treball

El flux de treball és l'estudi dels aspectes operacionals d'una activitat de treball. Té en compte com s'estructuren les tasques, com es realitzen, en quin ordre es duen a terme, de quina manera es sincronitzen, el flux de la informació que generen les tasques i el seguiment del compliment de les tasques [3]. Una aplicació de flux de treball automatitza la seqüència d'accions, activitats i tasques utilitzades per l'execució del procés, incloent-hi també el seguiment de l'estat de cada una de les seves etapes i aportant les eines

necessàries per gestionar-lo. El propòsit dels sistemes de fluxos de treball és apropar persones, processos i màquines amb l'objectiu de reduir temps i accelerar la realització d'algun treball. D'aquesta manera també estableixen els mecanismes de control i seguiment dels processos organitzatius, agilitzen el procés d'intercanvi d'informació així com la presa de decisions de l'organització. Com a resultat s'optimitza el servei i es millora la gestió del coneixement.

1.1.3 Minería de procesos

La minería de procesos és una tècnica d'administració de processos que permet analitzar els processos de negoci d'acord amb un registre d'events [4]. A través d'aquesta activitat es desitja extreure coneixement des dels registres d'events dels processos emmagatzemats pels sistemes. A més, la minería de processos és utilitzada generalment quan no existeix una descripció formal dels procesos o quan la informació existent és de mala qualitat. Essent el descobriment de models de processos una etapa inicial, el pas següent de la minería de processos és aconseguir una millora en els processos estudiats mitjançant la detecció de colls d'ampolla en el procés, millors rutes per complir amb l'objectiu, prevenció de congestions, etcètera. Podem veure un esquema en la Figura 1.

La minería de processos, en el camp de l'enginyeria de processos de negocis, es fa servir en l'anàlisi de processos, el disseny de processos i la validació de processos. L'objectiu és tenir el control dels processos, permetre el descobriment de nous processos, controls, informació i estructures d'organització partint dels registres d'events i per altre banda la seva aplicació ajuda a identificar colls d'ampolla, anticipar problemes, identificar violacions de normes que s'han de respectar i simplificar processos.

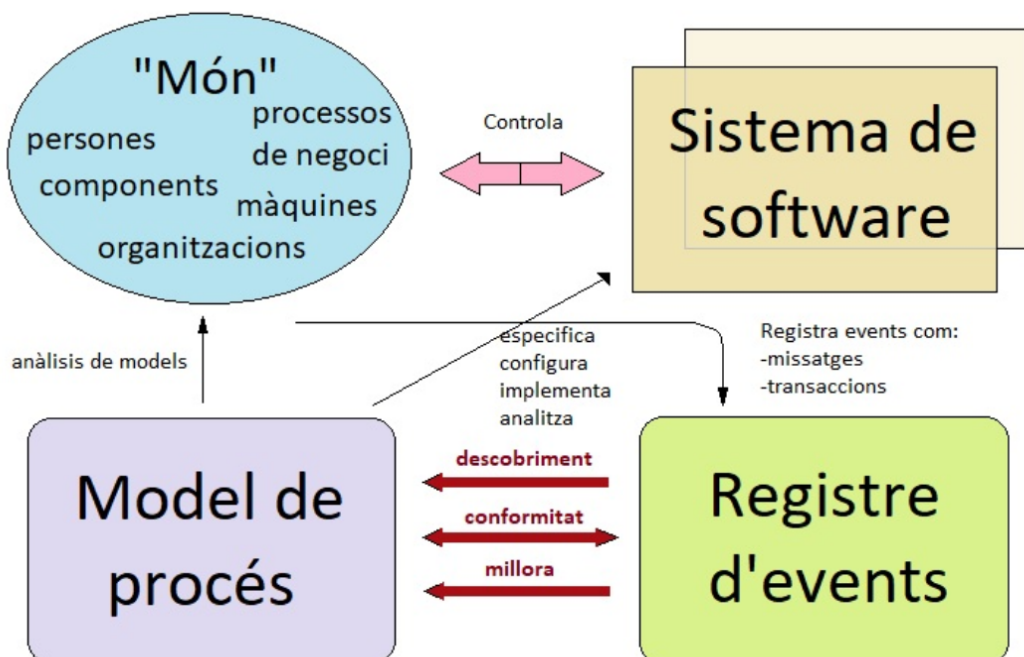


Figura 1: Esquema conceptual de la minería de procesos [5]. Elaboració pròpia.

1.1.4 BPMN

De l'anglès *Business Process Model and Notation* (BPMN), és una notació gràfica estandaritzada que permet el modelat de processos de negoci. L'objectiu de BPMN és proporcionar una notació gràfica que sigui fàcil de llegir i usar i s'entengui per part de tots els involucrats i interessats del negoci [6]. A més, acabant amb la bretxa de comunicació que freqüentment es presenta en el món empresarial entre el disseny de processos de negoci i la seva implementació actuant com a pont entre aquests dos factors i fent que es beneficiïn ambdues parts, els dissenyadors del procés de negoci i els que l'implementaran.

Aquesta notació va ser inicialment desenvolupada per la *Business Process Management Initiative* (BPMI) davant la necessitat d'unificar els diferents tipus de notacions que feien servir les organitzacions en el camp de la gestió de processos de negoci (en anglès: *Business Process Management* (BPM)) en un sol tipus de notació global i estandaritzada. Des de que aquesta notació és mantinguda per la *Object Management Group* (OMG), l'any 2005, BPMN s'ha convertit de-facto en la notació estàndard pels diagrames de processos de negoci, facilitant i normalitzant notablement l'intercanvi d'informació entre les diferents empreses i organitzacions [7].

Una BPMN està pensada per donar suport únicament a aquells processos que siguin aplicables a processos de negoci. Això vol dir que qualsevol altre tipus de modelat realitzat per una organització que tingui una finalitat diferent a la dels negocis no estarà en l'àmbit de BPMN, com ara un model d'estructures organitzatives o models de dades. A més a més tot i que en una BPMN es mostra el flux de dades (missatges) i la associació de dades amb activitats, de cap manera és un diagrama de flux de dades.

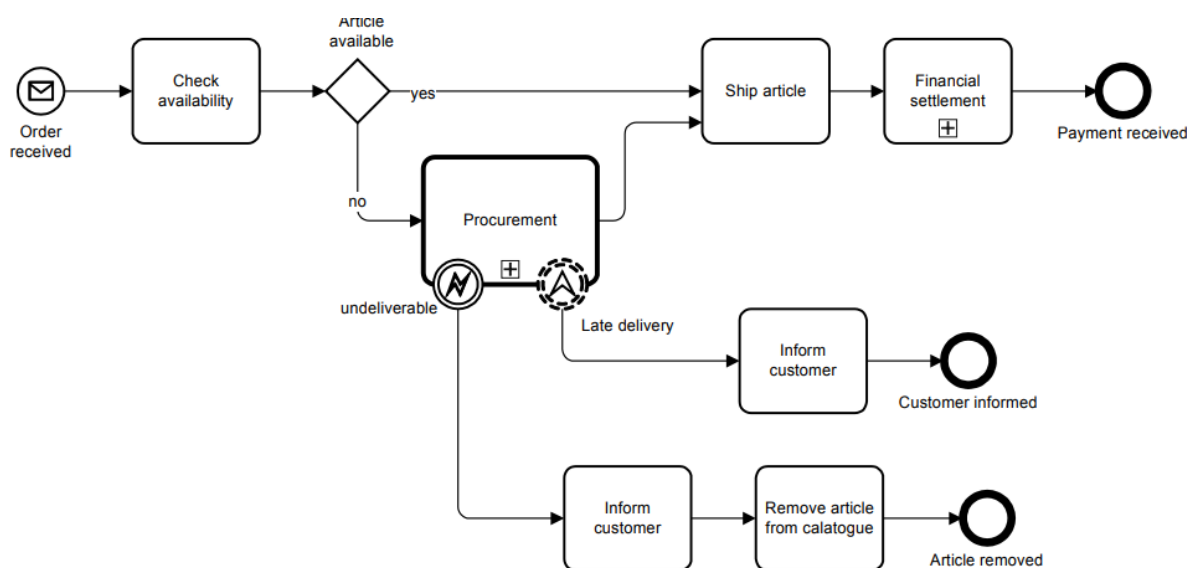


Figura 2: Exemple d'un diagrama BPMN [8]

1.2 Descripció del problema

El camp de la BPMN és un camp molt ben definit i acotat. Existeixen nombroses eines per crear nous processos de negoci o editar-ne d'existents. Inclús existeixen programaris que et permeten obtenir el seu codi font en diferents formats per posteriorment manipular-los. Però no existeix cap programari que donada la imatge d'un diagrama pugui reconèixer el contingut del mateix. En aquest aspecte doncs, si es vol manipular un diagrama BPMN s'ha de disposar del fitxer font o del codi font, no n'hi ha prou amb la imatge del mateix diagrama ja que no hi existeixen aplicacions que mitjançant tècniques de visió per computador reconguin el contingut d'aquests diagrames. Això presenta un problema, ja que a l'hora de treballar amb un diagrama BPMN s'ha de disposar sempre de les metadades del mateix, impossibilitant la manipulació informàtica si no es disposa d'aquestes.

En definitiva, donat el format estàndard d'un diagrama BPMN es pot obtenir la imatge del diagrama, però donada la imatge d'un diagrama no es pot obtenir el format que el construeix.

1.3 Actors implicats

Els fruits d'aquest treball, principalment, van dirigits al grup PADS-UPC encarregat d'investigar en el camp, desenvolupar i promocionar l'eina, que es veuria beneficiat a l'hora d'usar els resultats del treball per a continuar amb la investigació i, potser, crear un producte que posteriorment es podria vendre a empreses del mercat laboral.

Si finalment això passés, les beneficiats de l'eina en dites empreses serien els implicats en l'ús de diagrames de processos de negoci (*stakeholders*). En aquest grup de gent podem distingir els analistes de negoci (que són els encarregats de definir i redefinir els processos), els desenvolupadors tècnics (que són els que posteriorment implementaran els processos) i els gerents i administradors de negoci (qui monitoritzaran i gestionaran els processos de negoci).

D'aquesta manera, també es beneficiarien els clients, donat que el procés de fabricació del producte o el temps en donar el servei del qual estaria interessat el client, s'hauria abaratit gràcies a les millores d'eficiència introduïdes, potencialment, pels resultats d'aquest treball. En conseqüència, també es veuria beneficiada el conjunt de la societat en tant que client potencial de qualsevol procés de negoci que s'hagi vist impulsat per l'eina que es pretén desenvolupar en aquest TFG. Així doncs, el col·lectiu que en pot sortir beneficiat és potencialment molt ampli.

2. Abast del projecte

2.1 Estat de l'art

Per dur a terme aquest sistema de reconeixement d'imatges de processos, es començarà una aplicació des de zero, fent servir llibreries i eines ja existents de visió per computador. A PADS UPC són experts en el camp de la mineria de processos i en el *Business Process Management* fet que propicia un entorn ideal per desenvolupar dit software amb la seva ajuda i supervisió en tot moment.

Per altre banda, com que aquesta és una línia d'investigació nova, gairebé no hi ha informació sobre altres intents de desenvolupament d'un software capaç de reconèixer diagrames de procesos de negoci mitjançant tècniques de visió per computador, de fet no en consta cap a internet. Amb això, s'hauran de crear noves solucions per resoldre aquest problema i, si escau, adaptar soluciones existents que resolen problemes o sub-problemes similars. El que sí que existeix, en canvi, són altres programaris que s'encarreguen de reconèixer altres diagrames que no tenen res a veure amb els BPMN, és possible que l'estratègia a l'hora d'enfocar el problema i algunes sub-tasques siguin semblants a les del problema que s'està plantejant resoldre en aquest TFG.

També, en referència a l'estat de l'art del món de la visió per computador, existeixen multitud d'eines i llibreries que es poden fer servir per desenvolupar els propis softwares, com ara: OpenCV, Matlab, Tensorflow, SimpleCV, CUDA, Accord.NET Framework, ImageUltimate, libdwt, BoofCV, etcètera.

2.2 Objectius i sub-objectius

L'objectiu principal d'aquest treball és desenvolupar un software per transformar imatges raster de models de processos en objectes XML manipulables amb editors de modelat. Aquest format XML, és el format estàndard marcat per la Object Management Group i és el que usen les organitzacions. Tenint en compte que és una tasca que no s'ha fet abans i que és una àrea d'investigació nova, si l'objectiu final no s'aconsegueix, el que es vol fer és obrir camí en aquest sentit de l'investigació establint les bases, tècniques i conceptes per, en un futur no molt llunyà, continuar la investigació fins arribar a l'objectiu principal, que és desenvolupar un software de reconeixement de diagrames de processos de negocis.

Per altre banda, al estar parlant de desenvolupar un software de visió per computador sobre diagrames BPMN, s'ha de tenir en compte també que aquest software treballarà sobre un entorn molt concret i molt acotat. Els diagrames BPMN consten de diferents tipus d'elements, així doncs, les sub-tasques o sub-objectius d'aquest treball seran identificar correctament els elements que formen un diagrama BPMN. Aparentment, més enllà de reconèixer la varietat d'elements que pugui contenir un diagrama BPMN no hi ha altres requeriments funcionals.

A més, a part dels següents sub-objectius que es descriuen a continuació, que són els principals que conformen un diagrama BPMN, existeix una llarga llista d'elements que podem trobar en aquests diagrames que aquí no detallarem. Aquest TFG es centrarà en reconèixer correctament els elements descrits a continuació, ja que són els elements que formen el contingut real d'un diagrama BPMN i que es troben en el cent per cent d'aquests. Després, si aquests ja es reconeixen correctament s'aniran afegint altres elements donant prioritats als que siguin més freqüents. Cal dir també, que en aquesta llista d'elements que podem trobar, també hi trobem text, que va relacionat amb arestes o els objectes de flux.

2.2.1 Objectes de flux

Un diagrama de processos de negoci consta d'una sèrie d'elements principals, aquests s'anomenen Objectes de Flux (coneguts en anglès com a *Flow Objects*). El principal sub-objectiu d'aquest treball és reconèixer aquests elements. Donat que els diagrames BPMN es centren en ser clars i concisos, només existeixen tres tipus d'elements principals: els events, les activitats i les portes o *gateways*.



Figura 3: Representació gràfica dels objectes de flux. [6]

2.2.2 Tipus d'activitats

A la vegada, les tasques poden ser de diferents tipus, hi ha fins a vuit tipus de tasques. Aquests tipus són: *Send Task*, *Receive Task*, *User Task*, *Manual Task*, *Business Rule Task*, *Service Task*, *Script Task*. Totes aquestes tasques tenen una petita imatge associada a la cantonada superior esquerra. Per últim, també podem trobar les tasques per defecte, sense cap imatge associada a elles. Aquest sub-objectiu es basa en identificar correctament els diferents tipus d'activitats nomenats.



Figura 4: Representació gràfica de 3 tipus de tasques. [6]

2.2.3 Tipus de gateway

De la mateixa manera, les *gateways* també es poden trobar de diferents maneres. Hi ha cinc tipus de gateways, cada una d'elles és formada per un rombe amb una figura al seu interior. Podem trobar cinc tipus de *gateways*: *exclusive*, *parallel*, *inclusive*, *complex* i *event based*.



Figura 5: Representació gràfica de 3 tipus de *gateways*

2.2.4 Reconeixement de text

Els objectes de flux poden tenir text que hi va associat. Per exemple, en el cas de de les activitats, podem trobar text en l'interior de les mateixes que ens descriu amb més detall en què consisteix aquella activitat. En el cas dels events, aquests texts els trobem sota els mateixos objectes. En el cas de les gateways, el text el podem trobar sobre les diferents arestes que surten d'aquests objectes. El reconeixement d'aquests textos no és essencial per a la llegibilitat del l'estructura del BPMN, però facilita enormement a l'usuari entendre quina és la funció d'aquell BPMN.

2.2.5 Objectes de connexió

Els objectes de flux, estan connectats entre ells mitjançant objectes de connexió (coneguts en anglès com a *Connecting Objects*) creant l'esquelet bàsic del procés de negoci. Reconèixer aquests elements juntament amb els objectes de flux és crucial, ja que com són els principals elements que defineixen la geometria, són els que donen forma i sentit al diagrama, així que és un sub-objectiu important. Només existeixen tres objectes de connexió, per simplificar el contingut dels diagrames, aquests són: connexió de seqüència, connexió de missatge i connexió d'associació.

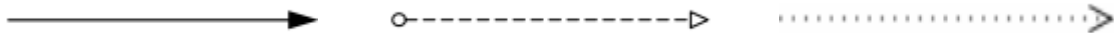


Figura 6: Representació gràfica dels objectes de connexió. [6]

2.2.6 Carrils

Molts models de processos utilitzen Carrils (en anglès *Swimlanes*) com un mecanisme per organitzar activitats i separar-les visualment en categories per il·lustrar diferents capacitats funcionals o responsabilitats. El següent sub-objectiu d'aquest treball serà identificar quines activitats hi ha en cada carril.



Figura 7: Representació gràfica d'un carril amb dos sub-carrils. [6]

2.2.7 Artefactes

BPMN va ser dissenyat per permetre als modeladors i a les eines de modelatge tenir algun tipus de flexibilitat per estendre la notació bàsica i disposar d'elements per donar el context apropiat a les diferents tasques o activitats del procés de negoci. Donat un diagrama, se li pot afegir el número d'artefactes que el modelador consideri. La versió actual de BPMN té en compte tres tipus d'artefactes, que són: Objectes de dades, grups i anotacions. L'últim dels sub-objectius d'aquest treball és identificar els artefactes, ja que, a la pràctica, només donen valor situacional i de context, no aporten informació al procés de negoci.



Figura 8: Representació gràfica dels artefactes que podem trobar en una BPMN. [6]

2.2.8 Construcció del diagrama i del XML

Per últim, evidentment, després d'haver reconegut els diferents *items* del graf, s'haurà de procedir a la reconstrucció del mateix obtenint un graf equivalent i amb aquest, per últim, s'haurà de procedir a la creació del fitxer XML que tindrà, si el procés s'ha fet satisfactòriament, tota la informació del diagrama d'entrada així representant un diagrama calcat al de l'entrada. Per construir aquest XML correctament s'han de seguir estrictament les directives de format dels BPMN, marcades per la *Object Management Group* [6].

2.3 Obstacles i riscos

Com s'ha dit abans, en aquest projecte s'obre un front d'investigació que no s'ha obert abans, és un front d'investigació nou. Això comporta, evidentment, que no se sap el que un es pot trobar de cara a arribar a l'objectiu proposat. De totes maneres, com en qualsevol altre projecte, s'ha de tenir en compte que poden sorgir errors de disseny, errors d'implementació i hi pot haver una falta de temps per assolir els objectius proposats.

També, un dels principals riscos és emprar una estratègia per resoldre el problema que no sigui la correcta des d'un primer moment, invertint així recursos en una idea morta des d'un principi. De totes maneres, al principi, es farà una avaluació d'opcions rigorosa per evitar-ho.

2.4 Metodologia i rigor

La metodologia de treball es basarà en l'autoaprenentatge i la investigació seguint les pautes que els tutors aniran marcant a mesura que es vagi avançant amb el contingut del

treball i es vagin obtenint resultats. Adaptant-se en cada moment als canvis necessaris que s'hagin d'adoptar per tal d'encarar el treball cap a la solució buscada.

2.4.1 Mètode àgil

Degut que no es disposa de molt temps per dur a terme aquest projecte, es durà a terme una estratègia *àgil*. Això vol dir que es mantindrà un constant contacte entre els tutors i l'alumne encarregat de fer el treball i s'anirà posant petites fites en petits períodes de temps, per així poder adaptar-se amb més facilitat als imprevistos que puguin sorgir i per planificar la feina en funció de l'evolució del progrés fins al moment.

2.4.2 Eines

Les eines que s'utilitzaran per fer aquest projecte són, a banda de les convencionals ofimàtiques i del mateix ordinador de sobretaula convencional, l'entorn amb el que finalment es decideixi treballar (Matlab o OpenCV per exemple). A més, també s'usarà l'eina GitHub que servirà per portar un correcte control de versions i per poder compartir còmode i fàcilment el codi font amb les parts interessades, tutors i alumne.

2.4.3 Validació del projecte

La validació del projecte s'anirà fent en les periòdiques reunions que s'anirant tenint al llarg del quadrimestre en que aquest projecte es du a terme. Tenint en compte i comprovant que, abans d'assolir l'objectiu d'aquest treball, s'hauran d'assolir correctament tots els sub-objectius del treball, i serà doncs, una vegada vist que els sub-objectius han estat assolits, quan es podrà comprovar si es compleix l'objectiu principal del treball. Fer aquesta comprovació és relativament senzilla, doncs només caldria veure si per una àmplia varietat d'exemples el software en qüestió es comporta com hauria de comportar-se. De fet, si escau, es pot fer un altre programari per testejar la correctesa del software del treball en qüestió.

Amb això doncs, el mètode que s'emprarà és el mètode àgil. El projecte s'anirà validant a mida que es vagi desenvolupant. I serà al final, un cop aquest estigui acabat, que es farà l'última validació amb les corresponents proves per determinar la fiabilitat del software que en surti.

3. Planificació temporal

Aquest TFG prèn l'equivalent a 18 crèdits ECTS. D'acord amb el marc del Sistema Europeu de Transferència de Crèdits[9] cada crèdit equival a una activitat d'estudi d'entre 25 i 30 hores, per tant, previsiblement aquest TFG tindrà una càrrega de treball d'entre 450 i 540 hores. En aquesta càrrega de treball, a l'hora, s'inclou el temps de planificació, les reunions amb els tutors, el temps invertit en l'autoaprenentatge i busca d'informació, l'elaboració del software solució i l'elaboració de la memòria així com també la lectura de la mateixa entre d'altres tasques. Per altre banda, aquestes hores de treball estaran repartides entre el dia 8 de gener de 2020 i el dia de la lectura de la memòria, que previsiblement serà a principis de l'estiu de 2020.

3.1 Definició de les tasques

En aquest primer apartat es procedeix a descriure cada una de les tasques que conformen aquest TFG, també es dona una previsió en hores de feina que requerirà cada tasca i se'n dona una justificació. A més, les tasques estan descrites en una seqüència lògica que es respectarà a l'hora d'elaborar el treball.

3.1.1 Gestió del projecte

Abans de començar amb l'elaboració del treball, és vital planejar i mantenir una bona gestió del mateix. En aquest punt s'explica els diferents punts de gestió que requerirà aquest TFG.

Reunions

Les reunions amb els diferents tutors són essencials per mantenir la coherència de la feina feta en tot moment. Donat que aquest treball va dirigit a PADS-UPC i que es fa amb la col·laboració de membres de la mateixa organització, serà de vital importància que els progressos fets siguin supervisats pels mateixos membres, per evitar així que desviacions en la línia de treball que interessa. És per això que com a mínim es farà una reunió per setmana, ja sigui presencial o per via telemàtica. A part de les reunions de seguiment també hi haurà reunions per consultar i demanar opinió als tutors sobre característiques tècniques en el camp de la visió de computadors, ja que és un camp on aquests són experts. Amb això s'estimen 25 hores.

Abast del projecte

Com s'ha dit abans, és important abans de començar amb la recerca i l'elaboració de la solució definir l'abast del projecte. Aquest s'ha definit juntament amb els membres de PADS-UPC, ja que com s'ha dit abans, el software que en sortirà d'aquest treball serà, en primera instància, per ells. S'estima que el temps requerit per realitzar aquesta tasca serà de 25 hores.

Planificació temporal

Una vegada definit l'abast del projecte, s'ha de fer una planificació temporal per ajudar a una millor planificació del mateix. La planificació associa la càrrega d'hores de treball que tindrà cada tasca que compon aquest TFG. Es preveuen unes 5 hores en l'elaboració d'aquest punt.

Codi	Tasca	Temps	Dependència
T1	1. Gestió del projecte	170h	
T1.1	1.1 Reunions	25h	
T1.2	1.2 Abast del projecte	25h	T1.1
T1.3	1.3 Planificació temporal	5h	T1.2
T1.4	1.4 Informe de sostenibilitat	10h	T1.2
T1.5	1.5 Memòria	90h	T1.2
T1.6	1.6 Lectura	15h	T1.5
T2	2. Treball previ	85h	
T2.1	2.1 Familiarització amb el programa	25h	T1.1
T2.2	2.2 Cerca d'informació i autoaprenentatge	45h	T1.2
T2.3	2.3 Valoració d'opcions	15h	T2.2
T3	3. Elaboració del software	240h	
T3.1	3.1 Disseny	30h	T2.3
T3.2	3.2 Implementació	140h	T3.1
T3.3	3.3 Testeig	35h	T3.2
T3.4	3.4 Valoració de resultats	20h	T3.3
T3.5	3.5 Propostes de millora	15h	T3.4
	Imprevistos	25h	
Total		520h	

Taula 1: Llistat de tasques que componen el TFG amb les respectives càrregues de feina i les seves dependències. Elaboració pròpia.

Informe de sostenibilitat

Una altra part del treball és l'elaboració d'un informe de sostenibilitat. Aquest document comprendrà tant la càrrega de sostenibilitat que suposarà l'elaboració del treball com els

efectes sostenibles del mateix una vegada finalitzat. Per elaborar aquest document es preveuen 10 hores de treball.

Memòria

L'elaboració d'aquest punt és, sinó la més important, una de les més importants del TFG. La memòria comprèn tota la documentació referent al treball: el context, els objectius, la planificació, la documentació referent al programa desenvolupat, els tests, les possibles millores, etcètera. No és senzill fer una previsió acurada del temps que requerirà fer la memòria, però tenint en compte TFGs previs realitzats a la FIB s'estima que la realització d'aquest document necessitarà d'unes 90 hores de feina.

Lectura

La lectura del TFG és la tasca que conclou el mateix treball. Aquesta no només consisteix de la mateixa lectura, que pot oscil·lar els 30 minuts, sinó que també consta de la preparació de la mateixa. L'elaboració de les diapositives o material multimèdia per agilitzar la defensa i la preparació de la mateixa. La previsió per fer aquesta tasca és de 15 hores en total.

3.1.2 Treball previ

Un cop feta la planificació del TFG i abans de començar amb l'elaborar del programari és molt important fer una sèrie de tasques de treball previ. En aquest apartat s'explica cada una d'aquestes tasques prèvies que s'han de fer i la predicció de càrrega de feina en hores que aquestes comportaran.

Elecció de l'entorn

Com s'ha dit abans, en l'explicació de l'estat de l'art, en aquest projecte es faran servir llibreries de visió per computador ja existents. En aquest aspecte, és essencial conèixer bé les possibilitats que ofereixen cada conjunt de llibreries com també és indispensable conèixer les estratègies que són més adients emprar en els possibles sub-problemes en que ens enfrontarem en aquest treball, com ara l'ús d'algorismes de *machine learning* o *deep learning*. També s'ha de tenir en compte que existeix un ampli ventall d'opcions que es poden fer servir per resoldre el problema proposat en aquest treball. Saber quines eines existeixen i quin és el potencial de cada una és important a l'hora de fer una bona elecció que permeti el millor desenvolupament del projecte. Es preveu que aquesta tasca durà 25 hores de feina.

Cerca d'informació i autoaprenentatge

Un cop feta la familiarització de l'entorn de punt de partida, es procedirà amb la recerca d'informació i autoaprenentatge per adquirir el suficient coneixement com per proposar estratègies de solució i poder implementar les mateixes. Conèixer l'estratègia que es durà a terme és també essencial abans de dur-la a terme. Aquesta tampoc és una tasca fàcil de

preveure el temps que comportarà, ja que és un front d'investigació nou i previsiblement sense una abundància d'informació, a la vegada és una tasca molt relacionada amb l'anterior. Es fa una previsió de 40 hores.

Valoració d'opcions en l'estrategia

Un cop adquirits els coneixements necessaris, s'hauran de veure les diferents opcions i posteriorment escollir la millor estratègia per resoldre el problema. Quina és la millor eina per abordar aquest problema? Quines llibreries s'adeqüen millor a la resolució d'aquest problema? I el llenguatge de programació?. Això es farà també, evidentment, amb l'ajuda dels tutors. Es preveuen 15 hores de feina en aquest punt.

3.1.3 Elaboración del software

Arribats a aquest punt, havent adquirit els coneixements necessaris i amb una estratègia clara de com s'abordarà el problema i coneixent les eines que es faran servir per a fer-ho, és moment de començar amb l'elaboració del software que resoldrà el problema. En aquest apartat també s'enumeren i s'expliquen els diferents punts que hi ha en l'elaboració del software així com una relació d'hores que aquestes tasques necessitaran.

Disseny

La fase de disseny servirà per donar estructura a la línia de treball i per evitar camins sense sortida. Donat que aquesta es farà també amb les recomanacions dels tutors, experts en la matèria, es preveu un bon disseny. Es preveu que aquesta tasca tindrà una càrrega de treball d'unes 30 hores.

Implementació

La implementació del software juntament amb l'elaboració de la memòria formen les tasques troncal del TFG. En aquesta tasca, de la mateixa manera que l'elaboració de la memòria, també es fa difícil preveure quina càrrega de treball suposarà. La predicció que se'n fa és de 140 hores de feina.

Testeig

Un cop elaborat el software, s'ha de procedir amb el testeig del mateix. Aquest pot ser manual o es pot fer un segon programari que testegi la robustesa del software elaborat. En aquest aspecte es preveu que això requerirà 35 hores.

Valoració de resultats

Amb els resultats de la fase de testeig, es podrà valorar quins han estat els resultats de les proves en la fase de testeig. Això ens servirà per saber quin nivell de fiabilitat i robustesa ha obtingut el software elaborat a l'hora d'executar-lo. Per a fer aquestes valoracions es preveu que 20 hores seran suficients.

Proposta de millora

Per acabar amb aquest apartat, tenint la valoració dels resultats obtinguts i tenint en compte quin era el comportament esperat del programa, es procedirà a l'elaboració d'una llista de possibles millores per a un treball futur, si s'escau, per a solventar les possibles deficiències que pugui presentar el programa fet. Així com millores d'eficiència o suggerències de canvis estructurals o de l'estratègia seguida per abordar el problema. Aquesta tasca es preveu que requerirà 15 hores de feina.

3.2 Gràfic de Gantt

Per il·lustrar gràficament i per fer-ho més entenedor, s'ha usat la tècnica de Gantt per representar la realització de les tasques i el temps que comportaran al llarg del temps.

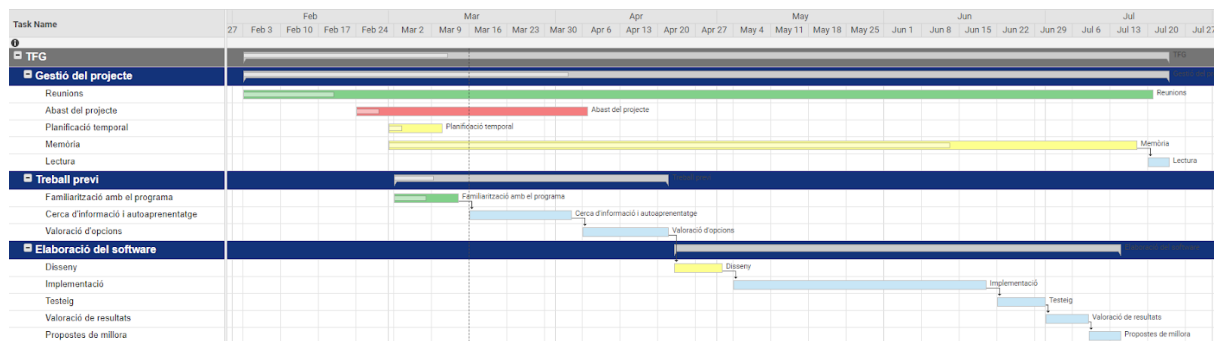


Figura 9: Il·lustració de Gantt de la planificació temporal del projecte. Elaborat al web app.smartsheet.com [10].
Elaboració pròpia.

El que es preveu doncs, segons els resultats que podem veure és una mitjana de

3.3 Gestió del risc

Al tractar-se d'un projecte d'unes dimensions considerables i al tractar-se d'un treball vertebrat en l'investigació, és possible i previsible que sorgeixin imprevistos. En aquest apartat es fa un recull dels imprevistos que poden sorgir i els plans alternatius que en derivarien. D'aquesta manera, a la predicció d'hores que s'invertiran en aquest treball, també li afegirem unes addicionals 30 hores per solventar possibles imprevistos que puguin sorgir.

3.3.1 Risc temporal

El primer de tots és el risc temporal, de no disposar suficient temps per implementar les la solució. Aquest problema pot venir donat o bé de la fase d'aprenentatge o bé de la fase d'implementació. En cas que es veiés que la predicció amb la primera fase no casés amb la realitat, s'hauria d'ampliar el número d'hores destinat treient-lo a l'hora dels altres apartats, principalment del de la fase d'implementació, ja que els altres apartats tenen molt menys marge. Amb això, el treball agafaria un enfoc més centrat en la investigació i recerca

d'informació i no tant en l'obtenció d'un software cent per cent funcional i robust. Pensant així en obrir camí en aquest front d'investigació per a possibles treballs futurs facilitant-los la feina.

3.3.2 Implementació

A l'hora de implementar el codi, al tractar-se d'un projecte relativament ambiciós, s'ha de tenir en ment que poden sorgir problemes i obstacles. Com haver partit d'un mal disseny o estratègia per resoldre el problema, encallar-se en algún punt en concret del programa o altres complicacions que poden sorgir i que naturalment sorgeixen a l'hora d'implementar aplicacions d'aquesta envergadura. A més, al Utilitzar tècniques relativament avantguardistes, com ara tècniques de deep learning aplicades a la visió per a computador, és previsible que el codi es pugui complicar en algun punt d'aquesta fase. De totes maneres no es preveu que es necessitessin recursos o eines addicionals més enllà del temporal.

3.3.3 Fiabilitat i robustesa

Tenint en compte que és possible que ni tan sols s'acabi el projecte degut a una falta de temps, encara que donat el previst cas de que aquest es pugui acabar en un estat usable, és possible que el software generat no sigui fiable o que funcioni bé per uns casos concrets i que per uns altres no ho faci. Això s'explicaria perquè uns sub-objectius serien fàcils de resoldre però altres hipotèticament es podrien complicar, o que els sub-objectius per separat fossin fàcils de resoldre però es compliqués a l'hora d'adjuntar les sub-solucions en una de sola.

3.4 Anàlisi d'alternatives i decisions

En aquest apartat es farà una valoració de les alternatives plantejades en els apartats del projecte i es justificarà, si escau, l'alternativa escollida finalment. Com és d'esperar, la intenció inicial és seguir la planificació temporal especificada en el gràfic de Gantt (Figura 9). Tot i així, s'han de tenir diferents factors que poden provocar desviaments en la planificació feta anteriorment.

El principal factor que pot propiciar canvis en la planificació és l'actual situació sota la qual es desenvolupa aquest TFG, aquest és l'estat d'emergència causat per la COVID-19. Donat que per fer aquest projecte no es necessiten eines més enllà d'un ordinador de sobretaula (apartat 2.4.2), l'estat d'alarma només hauria d'afectar al correcte desenvolupament d'aquest treball pel que fa les reunions periòdiques de seguiment, que s'hauran de fer telemàticament.

Pel que fa el disseny i el desenvolupament del projecte, hem de tenir en compte que es tracta d'un projecte amb un caràcter d'investigació fort. La idea principal és abordar tots els sub-objectius que s'han definit a l'apartat 2.2, però això no treu que el projecte és flexible i ens permet profunditzar en determinats sub-objectius en detriment d'altres. Una sèrie de tasques s'han designat amb una prioritat alta, unes amb prioritat mitja i unes altres amb prioritat baixa. En funció del temps que comporti dur a terme les tasques amb més prioritat i del temps que s'hagi designat a realitzar aquesta tasca, es decidirà si les altres tasques

amb menys prioritat es podran dur a terme. Es preveu que les tasques amb prioritat baixa no es puguin arribar a abordar. La classificació de les sub-tasques del desenvolupament d'aquest projecte s'ha fet així ja que les que tenen més prioritat són les que garanteixen que es reconeixeran amb èxit els elements més importants dels BPMN, fent així que l'algoritme sigui de més qualitat. Aquesta és la designació de les prioritats:

- Prioritat alta: [Objectes de Flux](#), [Objectes de Connexió](#) i [Construcció de l'XML](#).
- Prioritat mitja: [Tipus d'Activitat](#), [Tipus de Gateway](#) i [Reconeixement de Text](#).
- Prioritat baixa: [Carrils](#) i [Artefactes](#).

Pel que fa el testeig (i en bona part la valoració de resultats) aquest es pot dur a terme manualment, comparant visualment un per un el BPMN d'entrada amb el BPMN de sortida i anotant les possibles diferències que hi pugui haver amb un criteri propi, o per contra es pot automatitzar amb l'eina que ofereix RefMod-Miner, un comparador de BPMN online amb el seu propi criteri de comparació. Tot i que el testeig encara no s'ha fet, es preveu que la comparació es farà mitjançant aquesta eina online.

Fins ara, pel que fa la planificació temporal i tenint en compte les hores diàries que s'estan invertint en aquest treball (4 hores diàries) les expectatives de complir amb la previsió temporal detallada anteriorment són favorables.

3.5 Estat del projecte en la planificació

En aquest apartat es descriu l'estat de cadascuna de les tasques plantejades en l'apartat 3.1.3. Actualment el projecte es troba en fase de testeig i es compta amb un estat de la memòria avançat.

Abans però, respecte al treball previ (apartat 3.1.2) s'ha fet una valoració de les opcions que hi havia a l'hora d'escollir un l'entorn que facilités el correcte desenvolupament del projecte, finalment s'ha escollit Matlab per dur-lo a terme.

El disseny i la implementació de l'algoritme també es donen per tasques completades. Actualment, l'algoritme és capaç de detectar i distingir els objectes de flux, detectar els objectes de connexió continus, distingir els diferents tipus de gateway i activitats, reconèixer amb èxit relatiu depenent de la resolució del ràster d'entrada el text que hi ha en aquests i finalment convertir tota aquesta informació en un XML que genera el BPMN. És a dir que s'han completat les tasques amb prioritat alta i mitja, però no s'han abordat les tasques amb prioritat baixa.

Pel que fa el testeig, aquest finalment es farà híbridament entre l'eina de comparacions RefMod-Miner i una comparació manual amb una selecció de BPMNs que formaran el joc de proves. Aquesta avaluació de la fiabilitat i robustesa de l'algoritme obtingut es centrarà en aquells BPMNs que no tenen ni carrils ni artefactes, ja que són els elements que finalment no es tracten, però també s'evaluarà el comportament de l'algoritme en BPMNs on també hi hagi aquests elements. També s'ha pensat en afegir algun BPMN-art (dibuixos fets a partir d'elements dels BPMNs) per evaluar el comportament de l'algoritme en casos

extrems, pel que fa la visió per computador, que potencialment poden presentar aquests diagrames.

Per altre banda, a la memòria, a part de la fase de desenvolupament i de l'avaluació de la fiabilitat i robustesa de l'algoritme, on s'explicarà amb detall les característiques de l'algoritme obtingut en aquest treball, també hi haurà un punt on es farà una sèrie de propostes de millora, futures ampliacions i una avaluació d'una migració del codi a Open Source, ja que l'entorn escollit per elaborar el treball és tancat (Matlab).

Finalment, tal i com es pot veure en la planificació, cal notar que l'estat del projecte és l'esperat i es troba en un punt tal i com s'havia previst en la planificació temporal i amb això s'estima que aquest es també es finalitzarà d'acord amb la planificació temporal.

4. Gestió econòmica

En aquest apartat es definirà el pressupost del projecte, detectant i analitzant també els diferents costos associats que en deriven del cost final. Em sembla pertinent, però, recordar que aquest projecte no disposarà de cap pressupost ja que tampoc té previst cap cost addicional a part del pagament de la matrícula (amb el que ja s'hi inclou el dret a utilitzar els equips informàtics de la facultat com també el dret a les hores de consulta amb els professors associats al projecte). De totes maneres, com a part de gestió de projectes, en aquest apartat es durà a terme un pressupost per veure quin seria el cost d'un projecte d'aquesta envergadura en un entorn de mercat laboral, tenint en compte els costos de totes les parts implicades com de tots els equips i recursos usats.

4.1 Costos de personal per activita

Es planificaran tantes partides pressupostàries com tasques figuren en el gràfic de Gantt (veure Figura 6), a més també s'associarà el cost humà que suposarà la realització de cada tasca. Novament, com que aquest pressupost està enfocat a una suposada realització del projecte en el mercat laboral, es tindran en compte els suposats rols humans que es veurien implicats i no els reals; que serien el director, els dos co-directors i jo mateix. En aquest aspecte doncs, es tindrà en compte un cap de projecte que s'encarregarà de l'elaboració de l'estratègia i cerca d'informació, així com de les tasques de gestió del projecte, un programador que s'encarregarà d'implementar el software demanat i un tester que s'encarregarà d'analitzar i testear el resultat final del software així com de donar una valoració detallada acompanyada d'una llista de propostes de millora. Per elaborar aquest pressupost, primer, s'ha de veure quin preu tindria cada persona implicada per hora de treball. Aquest, a la vegada, podria variar en funció de l'empresa, de la ciutat o fins i tot en el país en que es fes aquest projecte, de totes maneres es farà una estimació de la remuneració per a cada perfil (retribució en brut) segons indica la consultora PagePersonnel [11] per a projectes de software. Tot seguit podem veure la taula de preus per hora de treball en funció dels perfils que participen en aquest projecte (veure Taula 2):

Perfil	Retribució
Cap de projecte	20 €/h
Programador	15€/h
Tester	15€/h

Taula 2: Preus de les hores de treball de cada un dels participants del projecte. Elaboració pròpia.

En aquesta taula, ja s'inclouen les despeses de la seguretat social de cada treballador implicat. Amb aquestes dades, ara podem fer la predicció del cost de les hores de treball humanes així com el desgloss del cost de cada tasca definida prèviament en el gràfic de

Gantt. En la següent taula (veure Taula 3) podem veure en detall el preu de realitzar cada tasca així com les hores que hi ha invertit cada persona. Al final podem veure el cost total pel que fa el factor humà.

Tasca	Cap de p.	Programador	Tester	Cost
1. Gestió del projecte	170 h	0 h	0 h	3400 €
1.1 Reunions	25 h	0 h	0 h	500 €
1.2 Abast del projecte	25 h	0 h	0 h	500 €
1.3 Planificació temporal	5 h	0 h	0 h	100 €
1.4 Informe de sostenibilitat	10 h	0 h	0 h	200 €
1.5 Memòria	90 h	0 h	0 h	1800 €
1.6 Lectura	15 h	0 h	0 h	300 €
2. Treball previ	70 h	15 h	0 h	1550 €
2.1 Familiarització amb el programa	10 h	15 h	0 h	350 €
2.2 Cerca d'informació i autoaprenentatge	45 h	0 h	0 h	900 €
2.3 Valoració d'opcions	15h	0 h	0 h	300 €
3. Elaboració del software	60 h	155 h	40 h	3850 €
3.1 Disseny	30 h	0 h	0 h	600 €
3.2 Implementació	0 h	140 h	0 h	2100 €
3.3 Testeig	0 h	0 h	35 h	525 €
3.4 Valoració de resultats	10 h	0 h	10 h	350 €
3.5 Propostes de millora	10 h	0 h	5 h	275 €
Imprevistos	10 h	15 h	5 h	500 €
Total	300 h	170 h	40 h	9300 €

Taula 3: Predicció de costos de personal en funció de les tasques que realitzen. Elaboració pròpia.

4.2 Costos genèrics

Els costos genèrics són els que intervenen en les múltiples fases del projecte i que van a part del costos del factor humà. Aquests es calcularan globalment sense entrar al detall del càlcul de cada tasca o activitat, encara cada apartat anirà acompanyat de la pertinent justificació. Els costos genèrics que tindrem en compte en aquest projecte són els costos de material, de desplaçament, d'espai, d'ús d'internet i d'ús del corrent elèctric.

4.2.1 Material

Pel que fa els costos relacionats amb l'ús de material, només haurem de tenir en compte el hardware, no per adquirir-lo sino la part d'amortització al usar-lo. El software usat no comportarà cap cost ja que s'usaran versions gratuïtes del mateix tant pel que fa el sistema operatiu dels equips (distribucions OpenSUSE de Linux), com pel que fa l'ús de software tècnic (Python 3 i OpenCV) o l'editor de textos ofimàtic (LaTeX), programari gratuït en tots els casos.

Així doncs els costos relacionats amb el material són exclusivament associats al hardware usat, que en el cas d'aquest projecte consta únicament d'un ordinador de sobretaula de prestacions mitjanes-altes. L'amortització de l'ús d'aquest ordinador es calcularà de la següent manera:

$$\text{Amortització} = \text{Preu Producte} * \frac{\text{mesos projecte}}{\text{mesos vida útil del producte}}$$

Fent una estimació genèrica dels valors de les variables anteriors, podem dir que el preu de l'ordinador de sobretaula usat és de 900€ i que té una vida útil de 72 mesos (6 anys). Aleshores, tenint en compte que el projecte té una durada de 5 mesos, l'amortització en quant a l'ús d'aquest ordinador és de 62.5€.

4.2.2 Desplaçament

Pel que fa els costos de desplaçament del personal implicat amb l'elaboració d'aquest projecte, es requereixen dos trajectes (un d'anada i un de tornada) per mitjà de la xarxa de transport públic de Barcelona cada dia. Amb això tenim 10 transports cada setmana, uns 40 per mes. Sent conscients del marc tarifari existent que oferta TMB (Transports Metropolitans de Barcelona) [12], l'opció més econòmica és optar per 5 T-usuals, una per cada mes. Amb el preu de 40€ per cada T-usual, el cost total pel que fa els desplaçaments és de 200€.

4.2.3 Espai

Aquest projecte es durà a terme a les instal·lacions de la FIB i del Campus Nord UPC. S'usaran les aules d'informàtica lliures així com els espais de la Biblioteca Rector Gabriel Ferraté (BRGF) a banda dels despatxos dels professors on es faran les reunions.

Novament, l'ús d'aquests espais no suposarà cap cost per a fer aquest TFG, però de la mateixa manera que s'ha fet en l'apartat anterior, s'inclourà un cost d'amortització al pressupost del projecte. Per estimar aquest cost, direm que l'ús de les instal·lacions del Campus Nord de la UPC i de la BRGF tenen un cost d'amortització de 30€ al mes. Amb això ens queda que el cost de l'ús dels espais necessaris per dur a terme aquest TFG és de 150€.

4.2.4 Ús d'internet

Seguint les línies de l'apartat anterior, i tenint en compte la difícil predicció real d'aquest aspecte farem una estimació de la següent manera. Suposem que la tarifa d'accés a internet és de 60€ al mes. Si tenim en compte que s'estarà accedint a internet 4 hores al dia (mitja jornada laboral) i que el projecte té una durada de 5 mesos, el cost d'accés a internet d'aquest projecte és $60 \text{ €/mes} * 5 \text{ mesos} * (4 \text{ h} / 24 \text{ h}) = 50\text{€}$.

4.2.5 Ús corrent elèctric

Per últim, ens queda fer una estimació del cost de proveir amb corrent elèctric l'equip que farem servir per aquest TFG. Per a fer-ho només ens cal saber la potència que té l'ordinador de sobretaula abans esmentat i el preu del kiloWatt hora (kWh) per a fer el càlcul del cost del subministrament energètic. L'ordinador té una potència aproximada de 600 Watts, i el preu del kWh és de mitja entre les 16h i les 20h és de 0.0945 €/kWh, oscil·lant entre un màxim de 0.098 €/kWh a les 20h i un mínim de 0.089€/kWh a les 16h de la tarda [13]. Sabent que se li donarà un ús d'unes 520 hores (veure Taula 1) a aquest ordinador i amb les dades anteriors, podem estimar que el cost de l'ús del corrent elèctric serà de 29.484 € ($0.6 \text{ kW} * 0.0945 \text{ €/kWh} * 520 \text{ h} = 29.484 \text{ €}$).

4.3 Altres costos

4.3.1 Contingència

En projectes d'aquesta envergadura, sempre és bo preparar un pla de contingència per fer front a possibles imprevistos que puguin sorgir. Normalment en projectes tecnològics com aquest, es preveuen plans de contingència de entre el 10% i el 20% del pressupost. En aquest cas doncs, hi aplicarem un pla de contingència del 15%. Donat que el pressupost és de 9792€ ($9300 \text{ €} + 62.5 \text{ €} + 200 \text{ €} + 150 \text{ €} + 50 \text{ €} + 29.5 \text{ €}$), el pla de contingència constarà de 1468.8 €.

4.3.2 Imprevistos

Per últim, a part de les hores de treball que ja s'han previst prèviament degut a possibles imprevistos per als diferents perfils dels treballadors implicats en el projecte, que sumen 500€ (veure Taula 3), també hem de tenir en compte altres imprevistos com el d'adquirir un nou ordinador degut a una incidència amb el mateix. Direm que el probabilitat de que ocorri una incidència que ens obligui a comprar un nou ordinador és del 10%. El cost associat a cobrir aquesta despesa doncs és de 90 € ($900 \text{ €} * 0.1$). Amb això, tenim que els costos totals per cobrir els possibles imprevistos són de 590 € ($500 \text{ €} + 90 \text{ €}$).

4.4 Resum

En aquest apartat es fa sumari de les despeses totals que en deriven d'aquest projecte.

Apartat	Cost
Personal	9300 €
Costos genèrics	492 €
Material	62.5 €
Desplaçament	200 €
Espai	150 €
Internet	50 €
Corrent elèctric	29.5 €
Altres costos	1558.8 €
Contingència	1468.8 €
Imprevistos	90 € ¹
Total	11350.8 €

Taula 4: Sumari dels costos associats al projecte. Elaboració pròpia.

4.5 Control de gestió

Per fer que aquest pressupost s'ajusti lo màxim possible a la realitat i que no hi hagin desviacions importants, es farà un estricte control d'hores que inverteix cada treballador en cada tasca. Si es detecten errades en la predicció d'hores de feina, s'haurà de jugar amb les hores extres d'imprevistos ja pressupostades, i si amb aquestes no n'hi ha prou s'haurà de jugar amb el pressupost del pla de contingència. De la mateixa manera que si amb determinades tasques no s'arriba al número d'hores previst, dites hores passaran a formar part al paquet d'hores extres poguent-se fer servir en altres tasques que ho requereixin. Aparentment no cal fer cap gestió més, ja que només es prediu l'ús d'un sol ordinador de sobretaula i no es necessiten de més recursos. El cas que l'ús d'aquest ordinador falli ja està previst a l'apartat 4.3.2.

5. Sostenibilitat i compromís social

5.1 Autoavaluació

L'enquesta dissenyada per EDINSOST [14] serveix per analitzar el nivell de formació de sostenibilitat de professors i estudiants en el sistema universitari espanyol i definir propostes de capacitació per ambdós col·lectius. Aquest projecte el finança el Ministerio de Economía, Industria y Competitividad.

Després de contestar-la, m'he donat compte que en general, de les diferents dimensions que hi ha, tinc bons coneixements pel que fa les dimensions social i econòmica. Per contra, pel que fa la dimensió ambiental sí que és cert que em manquen coneixements. Les preguntes de l'enquesta es poden agrupar en petits grups, les que fan referència a les dimensions social i econòmica com ara les que es refereixen a la innovació, planificació, treball col·laboratiu o impacte social són preguntes reafirmen el meu bon coneixement de la matèria, cal dir, a més, que ja havia fet a la mateixa facultat assignatures sobre el tema en qüestió. Per contra, les preguntes que refereixen a l'impacte medioambiental posen en evidència la meua manca de coneixements per a fer una avaluació rigorosa de l'impacte ambiental de projectes de gran envergadura, així com també posa en evidència la poca èmfasi que s'hi ha fet en aquest tema a les assignatures impartides a la facultat. De totes maneres l'impacte ambiental d'aquest projecte no és gens significatiu, ja que l'ús del transport públic per arribar a la facultat ja té més impacte ambiental que l'ús de l'equip necessari per realitzar el projecte.

5.2 Dimensió econòmica

Pel que fa la dimensió social del projecte, i tenint en compte els pressupostos de projectes similars, podem afirmar que no és un projecte car, mantenint-se en la mitja. Cal recordar també, que l'únic material que es fa servir (que no sigui gratuït) per fer aquest projecte és un ordinador de prestacions mitjanes-altes. Els altres aspectes com l'ús d'internet, del corrent elèctric o del transport són cent per cent essencials i no es poden retallar. També cal remarcar que el gruix del pressupost del projecte es destina al factor humà, a les tasques destinades al cap de projecte, al programador i al tester. Naturalment, al tractar-se d'un projecte d'investigació on es demana un software que resolgui un problema i tal i com es faria en un altre projecte de l'estil, que el gruix del pressupost es destini a aquest aspecte es justifica. Aquestes tasques, són essencials i no es pot retallar el pressupost per aquesta via si es vol realitzar el projecte amb garanties i sense presses. El pressupost en aquest aspecte no té marge de maniobra, no es pot reduir. Per altre banda, tampoc s'usen instal·lacions cares o equipament extra, el pressupost ja és eficient.

5.3 Dimensió ambiental

Sobre l'impacte ambiental, cal tenir present que en l'elaboració d'aquest projecte només s'usarà una eina que consumeixi recursos energètics que tinguin impacte ambiental.

Aquesta eina és un ordinador de sobretaula relativament potent (600W). De totes maneres, l'impacte ambiental del projecte és ínfim, ja que segurament només amb l'ús del transport públic durant una o dues setmanes ja superaríem la petjada ecològica que produiria l'ús d'aquest ordinador durant els 5 mesos que durarà el projecte. En aquest sentit és impossible reduir l'impacte ambiental pel que fa l'ús de les eines necessàries per dur a terme el projecte. Si realment volguéssim reduir l'impacte ambiental del projecte ens hauríem de centrar en el mètode de transport emprat per arribar a la facultat, i aquest ja és molt sostenible ja que es tracta del transport públic, dels més respectuosos que hi ha amb el medi ambient.

Per altre banda, també cal recordar que aquest projecte no parteix de zero en la búsqueda de la solució del problema plantejat, sinó que ja parteix d'una solució parcial. En aquest aspecte es reutilitzen recursos prèviament utilitzats. També és important recordar que només està planejat l'ús de còpies digitals tant de la memòria com d'altres documents per evitar al màxim l'ús de papers físics, així reduint l'impacte ecològic.

Per últim, cal evaluar el possible impacte ambiental de la posada en marxa de la solució buscada en aquest projecte. Aquest software, segurament reduiria el temps de feina dels actors implicats (veure punt 1.3) en el sector empresarial, i en general els facilitaria la vida. Però parlant de l'impacte ambiental, dubtablement aquest tindria un impacte significatiu, ja que no és un camp on s'hi destinin molts recursos. Estaríem parlant de l'estalvi de folis a l'hora de manipular diagrames BPMN i d'un hipotètic estalvi energètic degut a l'agilització a l'hora de manipular aquests diagrames en les seves versions digitals.

5.4 Dimensió social

L'impacte social més significatiu es produirà sobre la meva persona. Al acabar aquest projecte hauré acabat la carrera en Enginyeria Informàtica i això significarà una gran fita a la meva vida com també un gran canvi.

Per als integrants de PADS-UPC, si aquest projecte té èxit, aquest projecte els ajudarà a preparar un paquet de software per vendre a empreses i organitzacions, dotant així de més recursos i visibilitat a la organització associada a la UPC. De fet, encara que el projecte no acabi amb un software funcional, de ben segur que s'haurà avançat significativament en aquest camp de recerca. Avenços dels quals també en podran treure profit i es podran continuar posteriorment per acabar finalment en un software enfocat a resoldre problemes de la vida real.

Per últim, quant aquest projecte estigui acabat, l'impacte social que tindrà sobre els usuaris finals (veure punt 1.3) serà de facilitar-los la vida i reduir les seves hores de feina, sobretot pel que fa la manipulació de diagrames BPMN, però també facilitant la vida als tècnics que hagin d'implementar el contingut dels diagrames. Colateralment, també es beneficiaran les empreses que usin aquest software com a entitat, ja que seran més eficients degut a aquesta millora, de la mateixa manera que es beneficiaran els clients d'aquestes empreses ja que suposadament els costos globals s'haurien d'haver vist abaratits.

6. Integració de coneixements

En aquest apartat es parlarà de les diferents disciplines derivades de l'enginyeria informàtica que s'han vist implicades en l'elaboració d'aquest projecte

En aquest projecte s'ha fet servir una combinació de dues disciplines concretes de la informàtica. Estem parlant de la visió per computador (Computer Vision) i el Business Process Management. Tot i que la majoria de tècniques que s'han fet servir per implementar aquest projecte s'emmarquen en aquests dos camps, això no vol dir que s'hagin fet servir altres tècniques d'altres camps totalment diferents. A més, com es tracta d'un projecte amb una forta càrrega d'investigació, moltes de les tècniques que s'han fet servir són híbrids entre un disseny propi i algoritmes ja existents.

Per a l'elaboració del projecte s'utilitzen o s'ha plantejat utilitzar les següents tècniques i algoritmes relacionats amb els continguts de la carrera:

- Transformacions geomètriques i registre d'imatges.
- Filtrat d'imatges.
- Operacions morfològiques.
- Aritmètica d'imatge.
- Operacions lògiques sobre imatges
- Segmentació d'imatge i anàlisi d'objectes.
- Propietats de regió i imatge (detecció i extracció de característiques).
- Aprenentatge profund per al processament d'imatges.
- Algorismes de grafs de flux.
- Xarxes neuronals convolucionals pre-entrenades.
- Computació paral·lela
- OCR (*Optical Character Recognition*)
- Segmentació semàntica
- Detecció d'objectes usant característiques
- Detecció d'objectes usant aprenentatge profund

7. Lleis i regulacions

Actualment, a Espanya, no existeix cap legislació que reguli com una empresa ha de documentar els seus processos de negoci. Tampoc hi ha cap regulació pel que fa les tècniques de visió de computador per tractar i manipular les imatges que s'han fet servir en aquest projecte. En aquest aspecte, les regulacions estarien més enfocades en quines imatges es fan servir i no pas en com es fan servir. Donat que aquest projecte es centra únicament en els mètodes que es fan servir per tractar les imatges aquest projecte també queda fora de l'abast regulatori.

El més pròxim a una regulació que tenen els processos de negoci és l'estàndard BPMN, però en cap cas aquest és una llei. Aquest només especifica els tipus de dades dins el procés i com es formalitza la seva visualització. Aquest estàndard, de fet, és el que s'ha seguit a l'hora de generar el XML contenidor dels BPMNs. Si existeixen altres regulacions vers els processos de negoci, probablement siguin d'ús intern a la empresa o institució i, per tant, estan fora de l'abast d'aquest projecte.

Tenint en compte que el projecte desenvolupat en aquest treball de fi de grau pretén ser un producte comercial, s'han de tenir en compte les llicències de les eines i llibreries que s'han fet servir abans de crear el producte comercial. Donat que la llicència de Matlab que s'ha fet servir per desenvolupar aquest projecte és una llicència per a estudiants (facilitada per la UPC) i que no es permeten fins comercials amb aquest tipus de llicència, s'hauria d'avaluar si és més convenient adquirir una llicència que permeti els fins comercials o si econòmicament és més rentable migrar el codi a una altra plataforma com per exemple pot ser de codi obert (on també s'haurien de tenir en compte les llicències del producte).

Bibliografia

- 1_ MATHIAS KIRCHMER. High Performance Through Business Process Management. *Strategy Execution in a Digital World*. SPRINGER. 2017. pàgs. 3 - 14.
- 2 <https://www.cs.upc.edu/~pads-upc/> [en línia] (consultat: 22/02/2020)
- 3 ALEC SHARP, PATRICK MCDERMOTT. Workflow Modeling. *Tools for Process Improvement and Applications Development*. ARTECH HOUSE. 2009. vol 2. pàgs 225 - 227
- 4 <https://www.evaluandosoftware.com/la-mineria-procesos/> [en línia] (consultat: 22/02/2020)
- 5 https://www.ecured.cu/Miner%C3%ADa_de_procesos#/media/File:Tipos-de-mineria-de-procesos.jpeg [en línia] (consultat: 19/02/2020)
- 6 <https://www.ibm.com/developerworks/community/files/basic/anonymous/api/library/7624eb5a-089a-41bf-9b71-b3c33739e18d/document/e908d328-7b50-40e3-8107-70af4e6bb48f/media> [en línia] (consultat: 19/02/2020)
- 7 <https://www.omg.org/bpmn/index.htm> [en línia] (consultat: 23/02/2020)
- 8 <https://www.omg.org/cgi-bin/doc?dtc/10-06-02.pdf> [en línia] (consultat: 22/02/2020)
- 9 https://ec.europa.eu/education/resources-and-tools/document-library/ects-users-guide_en [en línia] (consultat: 1/03/2020)
- 10 <https://app.smartsheet.com/> [en línia] (consultat: 1/03/2020)
- 11 <https://www.pagepersonnel.es/quienes-somos> [en línia] (consultat: 6/03/2020)
- 12 <https://www.tmb.cat/ca/tarifes-metro-bus-barcelona/senzills-i-integrats/> [en línia] (consultat: 7/03/2020)
- 13 <https://tarifaluzhora.es/> [en línia] (consultat: 8/03/2020)
- 14 goo.gl/kWLMLE [en línia] (consultat: 8/03/2020)