

Trabajo Fin de Grado de Ciencia de Datos

Estructura y Contenidos Recomendados

Propósito del Documento.

Este documento establece pautas estructurales y de formato que deberían considerarse al elaborar la memoria de un Trabajo Fin de Grado de Ciencia de Datos en la Escuela Técnica Superior en Ingeniería Informática (ETSINF). Todos los TFG realizados en la ETSINF deben seguir la plantilla <http://www.inf.upv.es/www/etsinf/es/plantilla-tfg/> y deberían tener una extensión entre 50 y 100 páginas sin contar los anexos.



UNIVERSITAT
POLITÀCNICA
DE VALÈNCIA



Escola Tècnica Superior d'Enginyeria Informàtica
Universitat Politècnica de València

Simulating and Predicting Vera
Campus Mobility from WiFi Data

TREBALL FI DE GRAU
Grau en Ciència de Dades

Autor: Rogelio Conejo Lapin
Tutor: Robert Fuster
Xavier Molero

Curs 2014-2015

Ilustración 1 Portada

A continuación, se presentan las secciones o apartados en los que se recomienda estructurar la memoria.

Portada

Obligatorio. En la portada se incluye toda aquella información obligatoria o formal del TFG.

Se recomienda que estas páginas no tengan numeración o, en todo caso, que estén numeradas con números romanos al margen del resto de numeraciones del resto de partes de la obra.

Título

Obligatorio. Identifica de forma única el TFG. Es el nombre del TFG. Conviene que no sea artificialmente complejo. Un título no es el resumen de una obra. Es una referencia que ayuda a entender su contenido, sin desvelar todos sus detalles. Los más efectivos tienen menos de 10 palabras. No deben incluir abreviaturas ni acrónimos, a no ser que sea el nombre del sistema o herramienta. Por ejemplo: "e-Brain: Sistema de simulación cerebral basado en redes neuronales". El título deberá ser el mismo que el título aprobado por la Comisión Académica del Grado, si fuese diferente el alumno deberá solicitar una modificación del título.

Subtítulo

Opcional. A veces el TFG forma parte de una obra más grande o requiere una aclaración. Por ejemplo, un sistema de inteligencia artificial más amplio dentro del cual el alumno ha desarrollado un aspecto o parte concreta. El título puede repetirse entre varios TFG y cada TFG aclarar en el subtítulo qué parte ha desarrollado. En el ejemplo del título anterior, un subtítulo podría ser: "Subsistema de reconocimiento visual de bordes".

Autor/a y año de creación del TFG

Obligatorio. Incluye al autor con su nombre completo. Normalmente este será el nombre con el que se registrará en las bases de datos públicas de la biblioteca de la universidad. Se pone el curso académico de la convocatoria de defensa del TFG.

Tutor y Cotutores (si es el caso) del TFG

Obligatorio. De carácter obligatorio. Incluye el nombre y apellidos del tutor del TFG. Si existen dos tutores ambos deben aparecer como co-tutores. También puede incluirse el director/a experimental si se da el caso.

Dedicatoria

Opcional. A criterio del alumno. Típicamente a personas muy allegadas, familiares, etc. Es importante pensar bien el orden y los términos en los que se realizará esta dedicatoria.

Agradecimientos

Opcional. Puede colocarse al principio o al final de la obra justo antes de los puntos complementarios finales como anexos o bibliografía.

Normalmente los agradecimientos son a compañeros que pueden haber ayudado en la redacción o desarrollo del TFG, al tutor del TFG, el tutor en la empresa en la que se ha realizado el TFG, a la propia empresa en la que se han realizado las prácticas. Si procede, a compañeros que hayan desarrollado contenidos que no forman parte del TFG pero que le ayudan en su

aspecto profesional, etc. Si el trabajo ha sido financiado bajo algún proyecto, beca o similar deberá incluirse de forma obligatoria.

Resumen

Obligatorio. Esta parte será pública y seguramente la más leída de todo el TFG. Es mejor escribirlo cuando el TFG esté prácticamente finalizado, pero no en el último minuto. Debe ser una síntesis del trabajo y, por lo tanto, deben aparecer aquí todos los elementos fundamentales del TFG. Se precisa de una descripción concisa del problema que se pretende resolver, la metodología seguida para resolverlo, las herramientas utilizadas, los resultados y las conclusiones obtenidas de todo el proceso. Se deben presentar de forma clara, coherente y correctamente estructuradas. Un resumen debe ser autocontenido, sin hacer referencia a páginas o capítulos de la memoria. Normalmente no contiene referencias. La extensión máxima recomendada suele ser de 200 a 500 palabras. El alumno redactará el resumen al menos en inglés y en español o valenciano, pudiéndose presentar en los tres idiomas.

Palabras clave

Muy recomendable. Las palabras clave ayudan a tipificar y clasificar el TFG en las bases de datos, facilitando su búsqueda. Las palabras clave generalmente serán de 3 a 10 y podrán ser compuestas.

Prefacio o prólogo

Opcional. El prefacio o prólogo suele ser una introducción que acostumbra a destacar los méritos, el valor del trabajo, o también a situar la obra en un contexto y unas circunstancias determinantes. Esta parte es opcional y se incluirá en el caso de que se desee ampliar el resumen.

Índices y glosarios

Son colecciones de entradas que aparecen a lo largo de la memoria y que tienen como misión facilitar la localización de información relevante para el lector, así como poner de manifiesto la estructura general de la obra. Normalmente los índices suelen ir al principio de la obra, justo antes del primer capítulo de introducción y después de los agradecimientos (si estos se colocan al principio de la memoria).

Índice de contenido

Obligatorio. Tabla de contenido donde se pueden observar las partes en que está dividido el TFG. También se denomina índice de títulos. Como regla general, se recomienda mantener el contenido de los índices de títulos en un nivel 3 y excepcionalmente ampliar al 4 si la extensión y estructura de la obra lo requiere. Más allá del nivel 4 no tiene mucho sentido porque los índices se hacen muy extensos y el nivel de detalle hace perder al lector la estructura del TFG.

Índices generales, tablas e imágenes

Opcionales. Hay de varios tipos.

Índice o glosario de términos. Lista ordenada de conceptos, nombres propios, etc.; que aparecen en la obra, con las indicaciones necesarias para su localización. Ejemplo: Índice alfabético. Índice onomástico o un índice de materias. Se pueden incluir también índices de referencias a estándares. También es conveniente que aparezca al final del documento, como un anexo más, un capítulo a modo de diccionario breve, que contenga las principales palabras específicas o tecnicismos que se emplean en el trabajo para facilitar la comprensión por parte de

aquellos lectores no especializados en el área del TFG. Conviene avisar de la existencia de este diccionario en el capítulo de introducción.

Índice de tablas, imágenes, esquemas, etc. Suele ser un conjunto de listas ordenadas que incluyen las tablas que aparecen diseminadas por toda la memoria, imágenes, ilustraciones o esquemas. Se hace referencia al título que identifica el objeto incluido en la memoria, así como la página en la que aparece. Se pueden incluir tantas listas como se desee, si bien, lo habitual es colocar una de tablas y otra de ilustraciones.

Definiciones, abreviaturas y acrónimos. Recomendable. Dependiendo del área de conocimiento en el que se desarrolle el trabajo, puede hacerse uso de un gran número de éstas. Conviene poner un punto común en el que estén todas registradas a modo de diccionario. Suele aparecer conjuntamente con el índice o glosario si existen.

Autenticidad de la Información

Obligatorio. Cualquier texto, por pequeño que sea, extraído literalmente o con pequeñas modificaciones que no esté entrecomillado y convenientemente referenciado se considera un plagio inadmisibles para un TFG. Aun así, existen maneras de adaptar contenido sin aportación original por parte del autor. Por tanto, los alumnos deberán pasar el sistema TURNITIN (<https://www.upv.es/entidades/ASIC/catalogo/FAQS-Antiplagio.pdf>) a la memoria del TFG, y deberán subir el resultado del análisis en la aplicación EBRÓN cuando depositen su TFG. Actualmente se realiza a través de una tarea creada a propósito en el sitio PoliformaT asociado a los TFG. En un futuro esta funcionalidad se integrará en la aplicación Ebrón.

El sistema TURNITIN no examina las imágenes, ni otros elementos volantes (tablas y diagramas) pero los requisitos de autenticidad son los mismos. Cualquier imagen tomada de un sitio externo y no referenciada propiamente será considerada un plagio inadmisibles. Para evitar problemas se recomienda redibujar las gráficas e, incluso en ese caso, se ha de citar la fuente si la ilustración está basada en parte o totalmente en otras.

Narrativa y Visualización

Obligatorio. Un TFG de ciencia de datos debe contar una “historia” sobre los datos, y enfatizar qué se ha encontrado en ellos, qué se ha descubierto y por qué es novedoso o inesperado. Un TFG de ciencia de datos debe contener un número importante de tablas y de gráficas, elegidas convenientemente para representar la información y el conocimiento de la manera más adecuada posible según el momento de la memoria y lo que se quiere transmitir. Todas las tablas y gráficas deben ir numeradas, con un pie que las describa brevemente para poderlas entender básicamente sin necesidad de ir al texto principal. Se debe evitar pixelaciones. Se permiten visualizaciones interactivas siempre que se incluya una imagen estática y un enlace desde la memoria a la versión dinámica de la gráfica.

CAPÍTULOS O SECCIONES DEL TFG

Aparte de las secciones iniciales presentadas, el TFG presentará un cuerpo principal donde aparece el contenido fundamental del TFG desarrollado y posteriormente todos los anexos y apéndices complementarios. A continuación, se presentan una serie de capítulos o secciones

que se recomienda incluir. En la ETSINF se presentan TFG de distinta naturaleza (desarrollo de software, arquitectura de sistemas, emprendimiento, consultoría, etc.) y por ello, es importante destacar que estamos ante una recomendación, y pueden existir TFG donde se pueda prescindir de ciertas secciones o sea necesario incluir nuevas que no aparecen en este documento.

1. Introducción

Muy recomendable. Es el primer capítulo de la memoria. Se debe exponer el problema global de manera sencilla. Presentar una visión amplia y holística del problema. No se debe sobrestimar la familiaridad del lector con el tema del trabajo. No todos los lectores son especialistas en la materia ni la memoria debe estar redactada solo para los miembros del tribunal.

Ayudaría imaginar a una persona, profesor, profesional, etc. que se desenvuelve en un área diferente, aunque tenga conocimientos generales técnicos propios del Grado. Esta persona tiene vuestro mismo nivel de conocimiento general, pero sabe poco de la literatura, jerga o los trucos que se refieren a su tema particular.

Si aparece al final del documento un pequeño glosario/anexo de términos técnicos empleados comúnmente en la memoria pero que el lector medio pueda tener dificultad para entenderlos, conviene avisar de la existencia de este diccionario en este capítulo.

Se ha de escribir de manera que motive al lector a continuar leyendo. Para los primeros párrafos, la tradición permite la prosa, que es menos dura que el rigor exigido por la escritura científica. Es una buena idea preguntar a alguien que no sea un especialista sobre lo que opina tras leer la memoria. ¿Es una introducción adecuada? ¿Es fácil de seguir? ¿Es interesante?

Los siguientes tres apartados deberían ser subtítulos dentro del primer capítulo del TFG.

Motivación

Recomendable. Se ha de indicar los motivos que han llevado a realizar el TFG y en concreto uno de la temática elegida. Si había alternativas, hay que valorarlas y justificar las razones de esta decisión. Se debe dejar claro cuál es el tema y por qué es importante para el alumno (la empresa, asociación, etc.). Aquí se puede incluir las motivaciones personales que justifican el contenido y orientación del trabajo, motivaciones técnicas, objetivos de aprendizaje que se pretendían, motivaciones profesionales posteriores, etc. Tras leer este punto, el tribunal debe tener claro por qué le interesa esta temática al alumno. Se puede subdividir este apartado de motivación en una motivación personal en la que estaría justificado hablar en primera persona del singular y otra profesional más aséptica en tercera persona del singular reflexiva. Esta motivación debe ser la base para el establecimiento de los objetivos.

Objetivos

Muy recomendable. Qué se pretende conseguir con este trabajo. Hay que ser claros y concisos. El TFG posee un propósito que constituirá el punto de partida y orientará todo el proceso llevado a cabo durante su desarrollo o concepción. Ese propósito puede ser desarrollar un producto o un servicio, con sus posibles dependencias o relaciones con otros proyectos/trabajos.

En el caso de ser un trabajo realizado para una empresa o terceros, se ha de contextualizar el trabajo sobre la situación de partida, y describir cuáles serán las potenciales ventajas del producto o servicio tras su implantación.

El objetivo manifiesta la intencionalidad del trabajo. El principal requisito en el momento de plantear los objetivos es que estos sean alcanzables, razonables y medibles. Es decir que el interés que se persiga considere las posibilidades y limitaciones del trabajo realizado.

Pueden plantearse uno o varios objetivos primarios o generales de los cuales se derivarán una serie de objetivos secundarios o específicos; ya sean los primeros o los segundos siempre se ordenarán según su relevancia.

La valoración de todo el trabajo se da a partir de los objetivos propuestos, los cuales deben ser claramente identificables con el resultado final obtenido. Por lo tanto, el apartado Objetivos y el de Resultados están muy relacionados entre sí. En la redacción se deben considerar las siguientes recomendaciones para considerar los aspectos más relevantes en su formulación:

- Deben ser concretos; no deben redundar en frases largas.
- Deben ser factibles y medibles en su consecución.
- Deben identificar aquello que se quiere conseguir con el trabajo.
- Se deben utilizar los verbos en infinitivo, para identificar de forma clara los resultados de cada objetivo como una acción por realizar.

Impacto Esperado

Recomendable. Debe explicar las ventajas/mejoras que va a suponer el producto/servicio resultante del trabajo para el usuario.

Pueden identificarse diferentes usuarios y explicarse el impacto para cada uno de ellos. Por ejemplo: Prototipo de análisis de venta online con dos tipos de usuarios (gerente y comercial). Adicionalmente se puede relacionar el TFG con problemas contemporáneos, y en concreto con los ODS (Objetivos de Desarrollo Sostenible)¹.

Metodología

Recomendable. Debe explicar los diferentes pasos que se proponen para el cumplimiento de los objetivos. De alguna manera esta metodología guiará o determinará la estructura del trabajo a realizar y, por lo tanto, indirectamente la estructura de la memoria. Debe quedar claro que la metodología utilizada se adecua al ámbito del trabajo y a la orientación del TFG.

Estructura

Muy recomendable. La estructura del documento es un pequeño índice general comentado. Se hace un recorrido rápido de aquello que el lector puede encontrar en los sucesivos capítulos. No es necesario detallar el contenido más allá del segundo nivel de índice por no ser tedioso ni repetir aspectos que se verán más adelante. Básicamente es un párrafo grande que justifica y describe los capítulos para que el lector sepa qué va a encontrar en cada capítulo del TFG.

En esta sección se debe incluir información del contenido de los anexos con el fin de que el lector conozca de primera mano con qué contenidos adicionales puede contar.

Dado que los miembros del tribunal pueden estar formados por profesores que pueden provenir de áreas diversas ajenas a la temática del trabajo, es muy conveniente indicar que existe un glosario o diccionario de términos al final del TFG, que se puede consultar cuando se tengan dudas acerca del significado concreto de alguna palabra específica del área que toca la memoria.

¹ <https://www.un.org/sustainabledevelopment/es/objetivos-de-desarrollo-sostenible/>

Colaboraciones

Obligatorio (si se da el caso). En el caso de que el trabajo sea realizado junto con otros compañeros, hay que mencionar al equipo completo del trabajo indicando cada uno lo que ha realizado y en concreto, destacar qué parte ha realizado el alumno. En el subtítulo del TFG debería figurar claramente la aportación. Puede ser conveniente incluir el título del TFG con el que se ha colaborado.

Esta información es interesante para que, si alguien se descarga en el futuro una de las memorias de los integrantes del grupo, pueda localizar las complementarias. Por lo tanto, hay que incluir referencias a los otros autores, trabajos, explicar los contenidos desarrollados en esos trabajos que no forman parte del propio, así como la forma en la que se ha llevado a cabo la colaboración entre ambos. Esto deberían incluirlo en su TFG todos los miembros del equipo.

Por ejemplo, en el caso de un TFG de desarrollo de software de forma colaborativa, indicar los grandes bloques en que se dividirá la aplicación, cómo interactúan unos con otros, qué funcionalidad concreta va a tener cada uno, requisitos que deberá presentar el API de interacción, problemas a solucionar y el orden en que se abordan dichos problemas, etc.

Convenciones

Recomendable. Si el texto de la memoria va a incluir alguna normativa de marcado que lleve asociado un significado adicional, hay que establecer aquí los convencionalismos. Por ejemplo:

- El código fuente se muestra en letra courier cursiva. Y sólo se empleará esta tipología para este tipo de contenido.
- Las palabras extranjeras se remarcarán en cursiva.
- Se entrecomillarán las citas textuales externas a la obra.

2. Estado del arte

Muy recomendable. Esta sección también se podría denominar "Situación actual de la tecnología", "Contexto Tecnológico", "Estudio Estratégico" o términos similares según el idioma empleado (p.ej., en inglés "Background"). Se documentarán otras aproximaciones que existen actualmente o han existido en el mercado (historia de la evolución tecnológica) que aporten conocimiento o funcionalidades parecidos a las que se propone desarrollar en el TFG. Se expondrán posibles alternativas justificando la opción o el camino elegido para llevar a cabo este TFG.

En caso de considerar conveniente el uso de referencias a vídeos, las referencias deberían ser tratadas como notas al pie de página en la que aparezca la URL de acceso al vídeo y no como referencias "bibliográficas" o como hiperenlaces superpuestos al texto.

Es en este capítulo donde suelen concentrarse un porcentaje elevado de las referencias bibliográficas que aparezcan al final de la memoria.

Disponer de abundantes referencias bibliográficas en forma de revistas científicas o de divulgación profesional tiene como objetivo refrendar y dar autoridad a las afirmaciones que se realicen. La validación de las afirmaciones que vierte el alumno en esta parte mediante citas externas garantiza que el alumno no se está inventando nada de lo que está diciendo y que, en la realidad, las cosas funcionan como se está indicando. Es importante dar indicios no sólo propios sino fundamentalmente externos de que lo que se dice está validado y es real.

Si se ha realizado el TFG en el ámbito de una **práctica en empresa**, y con el consentimiento del tutor de empresa, es esta sección o en otra similar (con un título más adecuado) se puede presentar la empresa, su trayectoria, historia, productos que desarrolla, trabajos más representativos, marco en el que se engloban las prácticas realizadas, resaltar la importancia del trabajo realizado para la estrategia de la empresa, del proyecto en el que se enmarca, etc.

Si la empresa ya comercializa un producto software o hardware, datos o conocimiento al cual complementa el TFG, habría que describirlo ampliamente dado que dicho producto es el que determinará posteriormente los requisitos a cumplir por el TFG.

Tanto la empresa como el producto podrían aparecer en un par de apéndices si no se quiere hacer la memoria muy farragosa. En este último caso, podemos referenciar al producto/empresa sucintamente y referir a los apéndices para ampliar la información.

Crítica al estado del arte

Muy recomendable. En esta sección se deben analizar los trabajos anteriormente presentados en la ETSINF e identificar cualquier fallo, ineficacia, laguna, aspecto no abordado que justifique el desarrollo del TFG. Cuando uno acabe de leer este punto debería tener claro el espacio del conocimiento o de la tecnología que el autor quiere rellenar, objetivo a cumplir, etc.

En particular, en los proyectos de ciencia de datos, debe estar claro que nuevo valor se va a extraer a partir de los datos que no existiera ya, tanto desde el punto de vista de una mejor comprensión o predicción del comportamiento de un cierto ámbito social, empresarial o científico, o del descubrimiento de patrones o tendencias que se desconocían previamente, o que se habían ajustado o entendido mal hasta la fecha.

Propuesta

Muy recomendable. Resume, aclara y justifica cual es el espacio de conocimiento y tecnológico que va a llenar este trabajo. Incide en lo que diferencia, mejora o aporta tu trabajo a lo ya existente. También se puede indicar en qué se parece a lo ya existente. Muchas veces, más que una aportación genial, puede ser una prueba de concepto que combine de una manera original o novedosa soluciones ya existentes.

3. Análisis del problema

Muy recomendable. Una vez finalizado el estudio sobre el estado del arte sería adecuado un buen análisis del problema o la identificación de oportunidades de innovación o de negocio en el TFG. Este análisis se debe llevar a cabo de forma sistemática y utilizando las técnicas y métodos adecuados según la naturaleza del problema a resolver y el criterio del alumno y el tutor. Se pueden seguir las recomendaciones de las asignaturas de proyectos I, II y III para ellos, dependiendo del tipo de TFG.

Existen otros aspectos, quizás menos incluidos en las memorias actuales, que deberían estar presentes en muchos de los casos. Estos aspectos serán opcionales. A continuación, se comentan algunos de ellos:

Análisis de la seguridad

Recomendable. No todos los trabajos tienen este tipo de requisitos, no está de más que se especifique explícitamente cuando sea un aspecto de interés.

Análisis energético o de eficiencia algorítmica

Recomendable. Cada vez más modelos descriptivos o predictivos se ejecutan en dispositivos móviles cuya vida de la batería hay que alargar todo lo que se pueda, tanto en el día a día como

a lo largo de la vida útil del dispositivo. En cualquier caso, sea el entrenamiento o despliegue de modelos local o remoto, hay que contemplar tanto el coste del entrenamiento como del despliegue, el coste computacional y el consumo eléctrico asociados, incluyendo la adquisición de datos.

Análisis del marco legal y ético

Obligatorio. El marco legal es de obligado cumplimiento por todo profesional, y debe ser considerado desde diversos aspectos tales como:

Análisis de la protección de datos

Todos los trabajos deberían acceder y procesar datos, y extraer conocimiento a partir de ellos, por lo que es importante trazar como se generan los datos, cómo se almacenan y en el caso de cesiones seguir las normativas a tal efecto.

Propiedad intelectual

Considerar aspectos como el tipo de licencia que se va a utilizar para el desarrollo de la aplicación en caso de desarrollo de software y especialmente cuando se realiza en el seno de empresas o administraciones públicas. Y otros aspectos convencionales de la propiedad intelectual (p.ej. derechos de imágenes utilizadas en el producto final).

Otros aspectos legales

Para algunos trabajos se deben considerar elementos como el procedimiento electrónico (administración pública), esquema nacional de interoperabilidad, esquema nacional de seguridad, etc.

Ética

Además, el alumno se puede encontrar ante dilemas morales que deben ser analizados, especialmente en modelos que puedan discriminar por razones de sexo, raza, creencias o ausencia de ellas, ideología, y cualquier otro atributo protegido. No es suficiente con eliminar estos atributos del análisis, sino que hay que evaluar si existen 'proxies' para estos atributos o existen efectos colaterales sobre el uso de esos modelos en aplicaciones de marketing, recomendaciones o cualquier otro tipo de modelado que afecte al comportamiento humano.

Análisis de riesgos

Recomendable. En el caso de desarrollos orientados a un usuario, debería analizarse el riesgo de aceptación y los indicadores de satisfacción que podremos medir tras la implantación.

En caso de desarrollos orientados a la interoperabilidad, se deberán analizar riesgos tecnológicos y de integración. Si además, el resultado del TFG va a ser integrado en un proyecto de mayor alcance, también deberán considerarse los riesgos potenciales de dicha integración.

Cada riesgo se definirá determinando la siguiente información:

- Tipo de riesgo.
- Impacto que producirá en el proyecto o sobre el producto/servicio resultado del trabajo.
- Medidas que se han tomado o se pueden tomar par reducir dicho impacto.

Identificación y análisis de soluciones posibles

Muy recomendable. En la mayoría de TFG, a partir de los requisitos analizados, lo habitual es que no exista una solución única para la resolución del problema planteado. Habrá que presentar el abanico de soluciones posibles, analizarlas una por una, determinar pros y contras, establecer un criterio de selección y aplicarlo para encontrar cual va a ser la solución que finalmente se acabe desarrollando en el TFG. En algunos casos, es el propio problema el que se descubre a partir de los datos. El mensaje final que se debe transmitir en esta sección es que estamos ante un trabajo de un científico con mentalidad ingenieril, en el cual se han evaluado

alternativas y se ha elegido la solución adecuada o al menos la que cumple una serie de restricciones.

Solución propuesta

Muy recomendable. La solución elegida se presenta en esta sección, indicando en qué consiste, las fases por las que se pasará para su desarrollo o en las que se dividirá el TFG, cómo se va a implantar y validar (pruebas a realizar), etc.

Plan de Trabajo

Recomendable. Se recomienda definir un plan de trabajo mediante técnicas de planificación y estimación de esfuerzo que indiquen las fases en las que se dividirá el desarrollo del TFG y cuántas horas-persona necesitará para desarrollar cada fase. Todo ello permitirá calcular un presupuesto aproximado en euros y horas de trabajo del técnico. Se puede incluir fases desarrolladas por las otras personas indicando que dicho trabajo ha sido realizado por las otras personas integrantes en el grupo de desarrollo, destacando siempre la labor propia.

La forma de trabajar ideal sería que los alumnos realicen una estimación y un plan de trabajo antes de empezar el TFG y, conforme avance, anoten lo que realmente les está costando cada fase, anotando si hay desviaciones importantes, que corrijan sobre la marcha, y reasignen prioridades. El objetivo no es acertar a la primera o penalizar planes de trabajo desviados, sino que el alumno realice una reflexión crítica a posteriori de las diferencias obtenidas, que aprenda de los errores de estimación realizados, que averigüe qué detalles se dejó en la estimación para no volver a caer en ellos la próxima vez, en definitiva, que extraiga al final de todo el proceso un aprendizaje.

Presupuesto

Recomendable. Independientemente de que el alumno acabe trabajando en un centro de investigación, un servicio público o en la empresa privada, siempre tendrá que valorar si con los medios que dispone podrá ser capaz de realizar el trabajo propuesto; bien por él mismo o por un tercero. Es por ello por lo que se recomienda la realización de un presupuesto que determine cuál será la cantidad de recursos humanos y materiales necesarios para poder realizar dicho trabajo. Este punto es una oportunidad para que el alumno valore cuantitativamente el esfuerzo que él mismo ha realizado y tome conciencia del valor del producto realizado.

En el caso más sencillo (el alumno lo podrá detallar como lo estime) se deberá presentar una relación de material hw y sw necesario para poder desarrollar el TFG, el coste de los datos (si lo hubiere) y las horas-persona empleadas.

4. Preparación y Comprensión de Datos

Muy recomendable. Los proyectos de ciencia de datos deben trabajar sobre datos y estos tienen unos procesos de creación o integración, transformación, integración y preparación. En este capítulo o capítulos (cuyos nombres se pueden adaptar al contenido) se debe explicar el contexto de los datos y lo que significan, de qué fuentes provienen, que calidad tienen, cómo se transforman, etc.

Las tablas y gráficas de resumen ayudan enormemente a entender los datos, incluidos aquellos gráficos que permitan representar datos agregados o visualizarlos en su contexto (p.ej., datos geoespaciales sobre mapas).

Se debe incluir las razones de porque se incluyen o excluyen ciertos datos para el análisis o decisiones de capturar o integrar datos de distintas fuentes según los objetivos del proyecto.

5. Conocimiento Extraído y Evaluación de Modelos

Muy recomendable. Los datos se usan generalmente para extraer conocimiento de ellos mediante modelos, descriptivos o predictivos. En este capítulo o capítulos (cuyos nombres se pueden adaptar al contenido) generalmente pueden incluirse aspectos relativos al entrenamiento de los modelos, describiendo aspectos como la vista minable final para cada uno de ellos (si no se ha hecho en el capítulo anterior), qué métricas se van a utilizar para evaluarlos, tipo de evaluación (partición, validación cruzada, bootstrapping, etc.) y qué herramientas y técnicas se han utilizado en cada caso, incluyendo cómo se realiza la selección de hiperparámetros de los algoritmos.

Las tablas y gráficas deben ayudar a entender el conocimiento extraído y su rendimiento, detallando en qué partes los modelos funcionan mejor y peor, sus puntos débiles y su fiabilidad. También es importante incluir métricas sobre igualdad en el caso de que se utilicen sobre comportamiento humanos o con grupos protegidos.

En TFGs que incluyan simulación se puede incluir un capítulo específico sobre simulación, que integre los modelos extraídos en una simulación discreta o continua, p.ej., una simulación de stocks según varios modelos de compra y varias acciones de márketing, una simulación de ciudad vía *digital twins* viendo la evolución de los parámetros según distintas intervenciones, etc.

6. Validación y Despliegue

Muy recomendable. Los modelos en laboratorio y con los datos de la distribución original suelen funcionar mejor que cuando se despliegan. Por tanto, es importante incluir aquí todos los esfuerzos que van más allá de la evaluación en laboratorio realizada en el capítulo anterior. En este caso, es importante incluir pruebas pilotos, análisis de rendimiento de los modelos con diferentes datasets o distribuciones, cambios de condición de operación, encuestas sobre los usuarios sobre las decisiones de los modelos o el conocimiento extraído, siempre lo más cercano posible a la situación donde los modelos o el conocimiento van a ser utilizado y de la manera que van a ser utilizado.

Se debe diseñar el despliegue de los modelos, a veces integrados en aplicaciones o en mockup, en cuadros de mandos o aplicaciones móviles, en el caso de que este tipo de soluciones (o algún prototipo de las mismas) se haya incluido para mostrar el valor de datos. Este diseño del despliegue debe acompañarse de cómo se realizará la monitorización y el mantenimiento, en particular la detección de un decremento significativo en el rendimiento de los modelos o cómo se irían integrando nuevos datos.

7. Conclusiones

Muy recomendable. Esta sección es obligatoria y debe redactarse en forma sencilla, exhibiendo concordancia con las hipótesis aceptadas u objetivos formulados inicialmente. Todo lo que aparezca en esta sección debe de haber aparecido también al principio de la memoria en el apartado Objetivos de la Introducción. En esta sección se debe poner de manifiesto claramente si se han alcanzado todos los objetivos planteados y si se han desarrollado éstos satisfactoriamente, proponiendo ideas, soluciones o incluso nuevos objetivos surgidos a raíz de los anteriores.

En la memoria debe quedar claro el dominio que tiene el alumno de diferentes tecnologías y de cómo las ha empleado para poder dar solución al problema planteado en la introducción de la memoria. Debe poner de manifiesto que ha sido capaz de integrar los conocimientos de diferentes disciplinas o áreas para resolver el problema planteado haciendo un uso adecuado de la tecnología de vanguardia. Posiblemente esas tecnologías no se hayan visto durante la carrera o incluso se hayan visto de manera insuficiente para el contenido del TFG. Es necesario explicitar qué se ha aprendido que no se sabía o si ha sido muy difícil el desarrollo.

Hay que evitar cometer el error muy frecuente de repetir en la conclusión los resultados. Otro error es ir más allá de los objetivos planteados. El alumno tiene aquí la posibilidad de reflexionar sobre el trabajo realizado:

- Qué problemas se han encontrado y cómo los ha solucionado.
- Qué errores se han cometido y cómo se podrían haber evitado.
- Qué ha aprendido el alumno con este proyecto tanto profesional como personalmente. En este punto hay que indicar también qué nuevos conocimientos y tecnologías ha hecho falta que aprenda el alumno para poder realizarlo y el dominio alcanzado con ello.

Legado

Obligatorio. Es importante destacar que contribuye este proyecto y cómo los beneficiarios sacarán partido de él. Aparte de la memoria, el TFG debe permitir el acceso a datos y código en repositorios abiertos (o internos de la empresa u organización cuando haya limitaciones de derechos de propiedad intelectual o de datos personales), así como documentación complementaria al TFG, en el caso de que esta sea necesaria y no se encuentre en los apéndices.

Debe ser posible para un tercero poder *reproducir* el análisis y obtener resultados similares en el caso de que pudiera acceder a los datos y repetir los procedimientos.

En el caso de que haya algún elemento persistente (web, publicación, aplicación, etc.), debería aclararse cuál es el impacto o futuro a corto, medio y largo plazo que pueden tener. En definitiva, la memoria debe explicitar cuál es el legado que deja el TFG, tanto al alumno como al resto de la sociedad (en el contexto en el que se desarrolla el proyecto).

Relación del trabajo desarrollado con los estudios cursados

Obligatorio. Es conveniente e interesante pararse a pensar y realizar un ejercicio de introspección que incluya un análisis de la relación de los estudios realizados con el trabajo desarrollado por el alumno.

Este punto es una justificación de que el contenido del TFG es conforme a los estudios cursados. El objetivo del trabajo es poner en marcha y coordinar conocimientos recibidos a lo largo de los estudios con el fin de demostrar que se saben dar soluciones a problemas reales en el mundo laboral. Tener en cuenta que:

1. Un trabajo que demuestre el uso de muy pocas tecnologías, o muy específicas, demuestra que:
 1. Es un trabajo pobre desde el punto de vista de la demostración del dominio tecnológico amplio del alumno.
 2. Tal vez el alumno haya realizado el trabajo de aquello que se ha enterado y no sabe mucho más.
2. Un trabajo muy amplio que emplee muchas tecnologías puede ser agotador y acabar siendo muy largo o, caso de que haya tiempos máximos de entrega, acabe poco lucido por el coste de aprendizaje de las tecnologías no enseñadas en clase.

Por último, destacar que sería adecuado comentar qué competencias transversales y en qué

grado se han requerido y puesto en práctica para la elaboración del TFG.

8. Trabajos futuros

Recomendable. En este apartado se puede presentar una lista de:

- Fleclos que le hubiera gustado al alumno acabar en este trabajo pero que no ha podido ser porque se acababa el tiempo disponible para la realización del TFG.
- Líneas de desarrollo que se abren para aplicar estos resultados a otras áreas.
- Ampliaciones o mejoras tanto de eficiencia como de funcionalidades del trabajo realizado

También es importante indicar caminos por los cuales es mejor no seguir y esbozar las razones que se argumentan para evitar hacerlo.

9. Referencias

Obligatorio. Las referencias permiten al lector comprobar varias cosas acerca del trabajo:

1. La verificación de las fuentes originales en las que se basa el trabajo.
2. Que el alumno no ha reinventado la rueda y que ha construido sobre trabajo previamente desarrollado.
3. Es un indicador directo del grado de profundidad del trabajo realizado.

En este apartado se valora la abundancia de las fuentes de información, la relevancia y la adecuación al tema trabajado en el TFG.

El nivel de detalle al que se puede llegar referenciando desde el texto principal puede ser extenuante. Como regla de sentido común, hay que pensar en el lector o miembro del tribunal como un profesional de la informática con amplios conocimientos comunes pero que no es especialista en la materia de la que versa el TFG. Por lo tanto, no es necesario documentar ni referenciar permanentemente a la Wikipedia² cada vez que se emplea una palabra por primera vez ni para aclarar conceptos básicos de primero de carrera de informática. Sí que hará falta referenciar a artículos serios o científicos cuando haya que introducir determinados algoritmos, mejoras o determinadas tecnologías específicas propias de la temática del TFG.

Actualmente todo está en internet por lo que la tentación de convertir una bibliografía en una colección de enlaces a páginas web es muy alta. A continuación, se presentan algunas recomendaciones prácticas a tener en cuenta:

- No abusar de referencias *bibliográficas* a páginas web. Dejar para esta sección, al final del todo, todas las referencias a libros, artículos y revistas, libros o artículos serios de revistas científicas, sean electrónicos o no.
- Cualquier referencia a una página web del fabricante o del producto fabricado (herramienta empleada por el alumno o programas que sirven de referencia al TFG a desarrollar), se resuelve como una nota al pie de página. Normalmente se referencia la primera vez que aparecen.
- No puede aparecer ninguna referencia bibliográfica en la bibliografía que no esté referenciada al menos una vez en alguna parte del texto.
- No puede aparecer ninguna referencia en el texto a una referencia bibliográfica inexistente en la bibliografía.
- Puede haber varias referencias en el texto a la misma fuente bibliográfica.

²<http://www.wikipedia.org/>

Se deben referenciar en el texto las figuras e imágenes incluidas en la memoria.

Se deben documentar todas las fuentes bibliográficas utilizadas en un formato de citación estándar. Existen muchísimas normativas para establecer referencias bibliográficas. Se recomienda la norma internacional ISO 690-2010 para recoger las referencias bibliográficas [Bezos14]. Algunas bibliotecas han realizado un resumen organizado de la norma como la de la Universidad de Alicante³ que, a modo de prontuario, resuelven las principales dudas que puedan aparecer durante la redacción de referencias bibliográficas. Se pueden usar otras alternativas, como por ejemplo las distintas opciones que vienen con LaTeX (numeradas o abreviadas), o las que sugiere Google Académico (APA, Chicago, Harvard). En cualquier caso, si se elige un formato de referencias, se debe ser fiel a ese formato para todas las referencias.

Referencias a pie de página

Recomendable. Cualquier referencia a una página web del fabricante o del producto fabricado (herramienta empleada por el alumno o programas que sirven de referencia al TFG a desarrollar), se resuelve como una nota al pie de página. Normalmente se referencia la primera vez que aparecen.

Si al buscar información por internet aparecen varias referencias que finalmente dirigen a una web del fabricante o del producto, hay que referenciar a la web final, a la original y a la cual se refiere en el artículo. Visitarla y verificar que lo que dice el primer autor es cierto y añadir lo que se considere por el autor.

Anexos

Opcional. Normalmente se reserva para esta parte todos aquellos aspectos del TFG que no sea necesario leer en la memoria para entender su contenido de forma completa. Suele recoger detalles técnicos, gráficas o tablas adicionales, trozos de código, ampliaciones de una determinada parte, etc. Esta parte va dedicada a aquellos lectores que buscan algo muy específico o que tienen especial predilección por un tema concreto y quieren profundizar un poco más en ese aspecto por cuestiones personales.

También se puede colocar material suplementario que se ha desarrollado pero que se descartó pero que presenta interés. Es una forma de poner en valor la cantidad de trabajo realizado, aunque no todo haya ido en la dirección final del TFG.

Glosario

Recomendable. Dado que los miembros del tribunal pueden estar formados por profesores que pueden provenir de áreas diversas ajenas a la temática del trabajo, es muy conveniente indicar dentro del texto, en el apartado de estructura de la obra, que existe este glosario.

Debe contener aquellas palabras específicas y acrónimos del área que no tienen por qué ser conocidos por una persona que posea conocimientos medios en informática y no sea especialista en la materia.

CORRECCIONES Y MEJORAS AL DOCUMENTO.

Este no es un documento cerrado, se pueden proponer mejoras y correcciones al documento que serán estudiadas y validadas para ser incorporadas en futuras versiones. Se pueden enviar cualquier propuesta de mejora a la siguiente dirección de correo: scei@inf.upv.es.

³http://rua.ua.es/dspace/bitstream/10045/33984/1/Norma_ISO_Doctorado.pdf