



Deliverable 4 : Quality Assurance and Risk Management Plan

Última modificación: 25/11/2024

Fuentes: El proyecto latex editable de este documento se encuentra en overleaf, en: <https://www.overleaf.com/project/672b71e152b8be8c61e1bd93>

1. Introducción

Este entregable del proyecto **ANIMaLICOs**: *Advanced Networkmetrics: Interpretable Machine Learning for Intelligent Communication Systems*¹, se centra en describir las pautas que se han seguido para garantizar la calidad de los resultados del proyecto, así como enumerar los riesgos encontrados a lo largo del desarrollo del mismo (de entre los planificados en la propuesta) y las soluciones que se han adoptado para resolverlos satisfactoriamente.

2. Medidas de calidad adoptadas

Las medidas de calidad en un proyecto de investigación son fundamentales para garantizar la validez, confiabilidad y relevancia de los resultados. En este sentido, el proyecto ANIMALICOS, pretendía ofrecer múltiples resultados de diversa naturaleza: datasets, publicaciones científicas, reuniones científicas, DAaS, documentos entregables, web, multimedia (tutoriales, promoción de las tecnologías).

Enumeraremos a continuación cada uno de estos resultados y las medidas de calidad consideradas.

2.1. Publicaciones científicas

Se ha buscado que las publicaciones sean relevantes, precisas y de impacto, así como de ámbito internacional.

Se han elegido conferencias y revistas dentro de las líneas principales de trabajo del proyecto, en su mayoría, si bien algunas se han enviado a congresos relacionados con alguna de las demás

¹<https://www.codas.ugr.es/animalicos/en>



ramas que se estudian en el mismo (ya que es multidisciplinar).

En relación a los congresos, se ha optado por conferencias internacionales de calidad contrastada, con cierta antigüedad y conocidas por los investigadores del equipo. Por supuesto, todos con procesos de revisión ciega por pares. Se ha buscado principalmente aquellas que tengan cierto impacto reconocido, como por ejemplo, bien situadas en rankings como CORE (A o B).

En cuanto a las publicaciones en revistas, en su mayoría han sido enviadas a revistas recogidas en el Journal Citation Reports, es decir, con un índice de impacto reconocido y, a ser posible, ubicadas dentro de los cuartiles más relevantes.

En pos de promover la *ciencia abierta*, todas las publicaciones en revistas y congresos se han puesto a disposición de la comunidad científica por medio de plataformas de acceso libre, como Digibug, Researchgate, ArXiv o publicaciones e revistas *open access*.

2.2. Datasets

Uno de los resultados principales del proyecto, han sido los conjuntos de datos sobre tráfico de red, que se podrían utilizar por investigadores que apliquen métodos de análisis de datos o de aprendizaje automático para detección de patrones, modelado de tráfico de red, detección de ataques, y muchos otros problemas relacionados con las redes de computadores y la ciberseguridad.

En concreto, se han producido dos datasets: *UGR Feature Data* y *Dartmouth Feature Data*. Ambos han sido contruidos a partir de otros datasets existentes, los cuales han sido limpiados/refinados (corrigiendo valores atípicos o inconsistencias) o “reinterpretados” de alguna forma (calculando variables sumativas siguiendo enfoques más novedosos o precisos, por ejemplo), para que así resulten más útiles para los investigadores.

Este procesamiento en sí mismo ha mejorado la calidad de los datos iniciales, que fueron escogidos por ser dos dataset muy utilizados en diversas investigaciones desde hace unos años.

2.3. Organización de reuniones científicas

Se ha organizado un Workshop titulado “Quality of Data in Network Telemetry” (QoDaNeT 2024), al amparo del 2024 IEEE/IFIP Network Operations and Management Symposium (NOMS 2024)², cuya calidad está garantizada por la entidad que propociona el evento, es decir, el IEEE.

En dicho workshop participaron expertos de la academia y la industria como miembros del

²<https://noms2024.ieee-noms.org/authors/qodanet-2024>



comité de programa, incluyendo investigadores relevantes y empresas punteras en el entorno de las redes de comunicaciones (Nokia, Huawei, Swisscom, Telefónica...).

2.4. Herramientas

En el proyecto se han desarrollado varias herramientas software, entre las que destaca el despliegue de un DAaaS (*Data Analysis as a Service*)³.

Dicho servicio se ha desarrollado usando la tecnología más novedosa y que ofrece las mejores garantías, como es el caso de Docker para crear contenedores con software y datos necesarios para hacer las pruebas deseadas por los usuarios de manera fiable y segura.

Otros aspectos que se han cuidado para garantizar su calidad, es la creación de ejemplos y tutoriales de uso (en inglés), para guiar a los usuarios potenciales.

Este tipo de propuesta software como servicio (y más aún, análisis de datos como servicio) es un resultado de calidad en sí mismo dentro de un proyecto de investigación, ya que permite ofrecer un mecanismo para probar en primera persona las tecnologías y técnicas desarrolladas en el mismo, pudiendo tener un impacto mayor en la comunidad científica interesada en las líneas de estudio de dicho proyecto.

Por otra parte, también se ha desplegado un laboratorio de SDN, para su uso en determinados experimentos, aunque éste está circunscrito a su uso dentro de la Escuela Técnica Superior de Ingenierías Informática y de Telecomunicación (ETSIT) de la Universidad de Granada.

2.5. Página web

La propia web del proyecto se ha realizado siguiendo un estándar de calidad, siendo eficiente, accesible, completa y ofreciendo una versión íntegra en inglés para llegar a un mayor público objetivo.

2.6. Multimedia

Este proyecto tenía como uno de sus objetivos promover el conocimiento del término “Networkmetrics” entre la comunidad científica que investiga en el ámbito de las redes de comunicaciones o redes de computadores usando técnicas de *Data Mining* o *Machine Learning*.

³<https://codas.ugr.es/animalicos/en/daaas>



Por ello, una de las medidas que se adoptó fue la de la creación por parte de una empresa de contrastada experiencia de un conjunto de vídeos, publicados en Youtube, que describirían el proyecto, el concepto de “Networkmetrics” y el uso de las principales herramientas desarrolladas en el mismo: el laboratorio de SDN y el DAaaS⁴.

3. Riesgos abordados

En la propuesta del proyecto se incluyeron una serie de riesgos posibles, así como un plan de contingencia para cada uno de ellos (ver Figura 1).

Finalmente, sólo dos de los riesgos previstos se han cumplido, si bien son dos de los más importantes que podían ocurrir, aunque ajenos a los participantes en el proyecto. En concreto: 1) *No se concedió el becarío FPI que se solicitó para el proyecto* y 2) *El presupuesto concedido fue un 60 % menor del solicitado originalmente*.

Esta situación requirió una replanificación de los objetivos del proyecto, planteados como un conjunto de KPIs (*Key Performance Indicators*) a realizar o cumplir durante el desarrollo del mismo.

El resumen de las reducciones aplicadas por cada KPI se puede ver en la Figura 2.

Como se puede observar, la reducción en personal y en presupuesto implicó una reducción en el número de resultados esperados del proyecto en varias de las dimensiones propuestas. A su vez, la limitación presupuestaria ha tenido también repercusión en relación con la infraestructura que se ha podido desplegar en el proyecto (en relación al DAaaS), y que será un aspecto limitante en parte en cuanto a su impacto en la comunidad científica.

Aún así, el proyecto ha conseguido gran cantidad de resultados de calidad y ha resultado muy fructífero, como se puede observar en la web del mismo.

⁴<https://codas.ugr.es/animalicos/en/results>, Sección Vídeos



R&C	Risks & Cont. Description
R1	The SDN Lab does not provide the adequate experimental setup to test a specific research idea (e.g., reviewer of a paper demanding additional experimentation)
C1.a	Use simulation means (OPNET++, Mininet)
C1.b	Application to European test-beds (http://www.ict-fire.eu) for large-scale experimentation
R2	Delay or problems in the data capture
C2	Resort to publicly available data sets in CRAWDAD or CAIDA
R3	KPIs under quality specifications
C3	Re-assign work force and/or budget
R4	Team member leaving the project
C4.a	Re-assign tasks to work force
C4.b	Search and invite researcher with no compatibility restrictions
C4.c	Apply for grants (e.g., " <u>Programa Estatal de Promoción del Talento y su Empleabilidad</u> ")
R5	Privacy or regulatory problem (e.g., in data management)
C5	Investigate the problem and propose a solution to the General Assembly
R6	Limitations due to COVID-19 constrains
C6.a	Re-assign work force and/or budget
C6.b	Re-plan Specific Objectives and KPIs taking into account the specific situation
R7	Funding does not include the cost of hiring a technician (TOHIRE)
C7	Implementation Tasks T2.1 and T2.2 would be severely affected, and KPIs 1.2, 1.3 and 2.2 would be discarded. There may be additional consequences in the quality of publications
R8	FPI not awarded
C8	Project research tasks T3.1, T3.2, T3.4 would be severely affected, cutting KPIs 4.1-4.4 to half.

Figura 1: Lista de riesgos posibles y planes de contingencia recogidos en la propuesta original del proyecto.



KPI	Original Proposal	Actual Proposal
KPI1.1	SDN Network in lab	SDN Network in lab
KPI1.2	5 Apps traffic emul.	3 Apps traffic emul.
KPI1.3	2 Datasets	1 Dataset
KPI2.1	DAaaS	DAaaS
KPI2.2	2 examples	1 example
KPI3.1	3 (1 international col.) papers SDN traffic monitoring & forecasting	2 (1) papers
KPI3.2	3 (1) papers SDN cybersec	2 (1) papers
KPI3.3	3 (1) papers SDN opt	2 (1) papers
KPI4.1	3 (1) papers sparse MBDA	1 (1) paper
KPI4.2	3 (1) papers N-way MBDA	1 (1) paper
KPI4.3	3 (1) papers fusion MBDA	1 (1) paper
KPI4.4	3 (1) papers supervised MBDA	1 (1) paper
KPI4.5	3 (1) papers variable selection MBDA	1 (1) paper
KPI5.1	2 workshops	1 workshop
KPI5.2	5 papers / WS	5 papers / WS
KPI5.3	5 industry participants / WS	5 industry participants / WS
KPI5.4	2 special issues	1 special issue
KPI5.5	2 transfer contract	1 transfer contract
KPI5.6	2 HE proposals	1 HE proposal

Figura 2: KPIs en el propuesta original y KPIs realizados tras la replanificación. Los números entre paréntesis indican el número de contribuciones con colaboración internacional.