



Campañas de evaluación como motor para la generación de recursos en el dominio de biomedicina y sanidad

Martin Krallinger

PI - Text Mining Unit
(BSC-CNS)

INFODAY SOBRE TECNOLOGÍAS DEL LENGUAJE EN SANIDAD YBIOMEDICINA

Fecha: Organizado por BSC,
Barcelona Miércoles 4 de diciembre de 2019

Sesión 1: Datos en Salud para tecnologías del lenguaje

Sesión 2: Componentes y infraestructuras tipo software de PLN para sanidad

Sesión 3: Terminologías, ontologías, integración e interoperabilidad de recursos en Sanidad

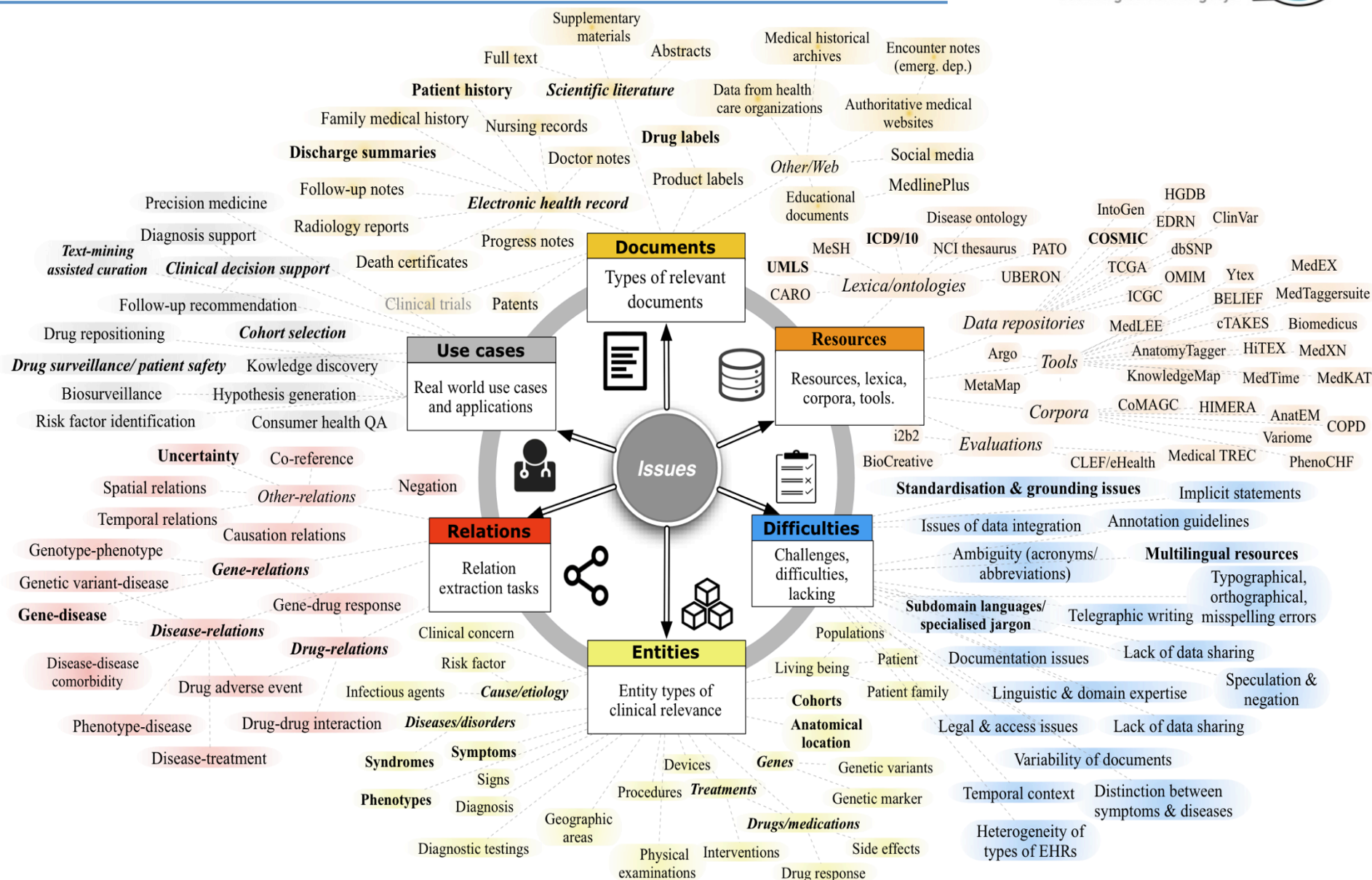
Sesión 4: Plataformas para PLN en Sanidad

Mesa redonda: Sanidad y escenarios de uso de PLN

Mesa redonda: Industria de tecnologías del lenguaje e informática de la salud

- **80% datos** en salud son no estructurados, con un valor clave para la toma de decisiones clínicas y seguridad del paciente.
- Los datos que generará el sistema sanitario en 2020 se duplicarán cada 73 días.
- A lo largo de su vida cada persona pueda generar más de un **millón de gigabytes** de información relacionada con su salud.
- Sólo en Galicia (con 2.7 millones de habitantes) se **generan 200.000 notas clínicas al día**.
- **1 de cada tres búsquedas** en internet tiene que ver sobre salud
- 67% de los usuarios de Internet españoles busca información sobre temas relacionados con su salud (www.red.es)

PLN en Salud: Escenario para Inglés



Datos: Shared corpora (& Annotations)

- Corpora from i2b2, n2c2, TREC, MIMIC III, BioCreative, GENIA,...

Software/código/repositorios abiertos o compartidos

- Mallet, WEKA, scikit-learn, Theano, TensorFlow,...
- cTakes, MedLEE, SemRep, MetaMap, GATE,...

Campañas de evaluation/hared tasks

I2b2, CLEF eHealth, TREC CDS (Precision Medicine),...

Marco legal de acceso a datos

- Protected Health Information (18 PHI items)
- De-identification based on PHI approved by clinical ethical board
- Distribution via Data Use Agreement (DUA)

Plan de Impulso de las Tecnologías del Lenguaje

Plan TL

Plan de Impulso de las
Tecnologías del Lenguaje



The Plan TL aims to promote the development of NLP and machine translation in Spanish and Spain's co-official languages.



The **Plan for the Advancement of Language Technology (Plan TL)** aims to promote development of NLP and machine translation in Spanish and Spain's co-official languages.

One of the flagship projects of the Plan TL is related to Healthcare and Biomedical domain.

- Increase the amount, quality and availability of **linguistic infrastructure**.
- **Transfer knowledge from the research field to the industry.**
- Improving the quality and capacity of public services, integrating NLP and machine translation technologies. Identify **use cases in public administration**

Corpora

- **Bilingual corpora (MeSpEN)**
- **Clinical cases corpus (SPACC)**: chemicals, drugs, genes, diseases, treatments, symptoms, abbreviations, procedures
- **Medical literature corpus**

Terminological resources

- **CUTEXT**: medical term recognition (incl. negation & certainty, UMLS mapping)
- **Medical Bilingual glossary, abbreviation-definition pairs**

Components

- **Medical tokenizer, sentence splitter, lemmatizer, PoS tagger, clinical sectionizer, HeidelTime adapt. NegEx adapt.**

Evaluation campaigns

- **Spanish medical abbreviations (IberEval2017, 2018)**: BARR (abstracts), BARR2 (clinical cases)
- **De-identification Spanish (IberLEF 2019)**: MEDOCCAN



Hospitals

- Secondary use of EHR:
- ICD 10 codification
- Anonymization



Libraries

- Indexing articles using DeCS-MeSH
- Apply NLP to allow for new services



R&D funding agencies

- Technological supervision and competitive intelligence: tools for strategic decision-making



AEMPS

- Access to resources and services to be integrated in their portfolio
- Pharmacoepidemiology studies

Infraestructuras y recursos Lingüísticos (I)

<https://github.com/PlanTL-SANIDAD> <https://zenodo.org/communities/medicalnlp>

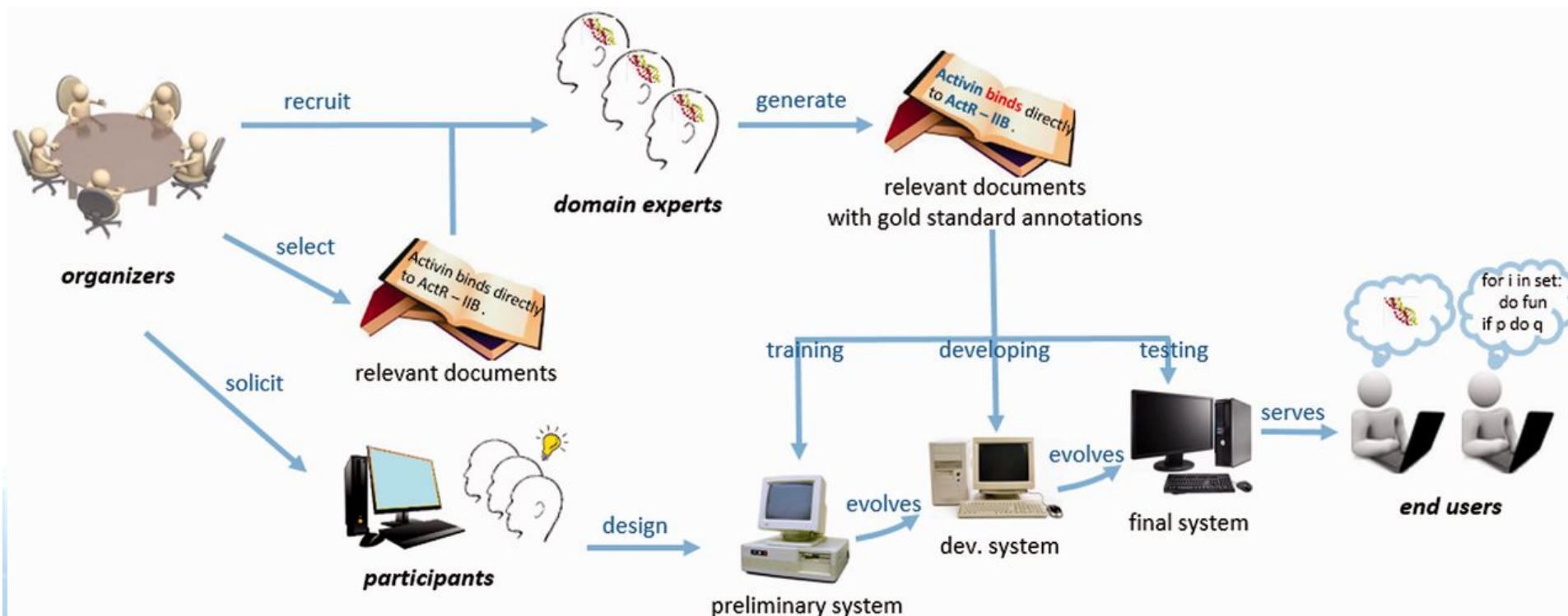
Recurso	Descripción / resumen
Terminological resources	• Registry of linguistic / terminological resources (368 in total, 103 of medical domain) • Spanish medical abbreviation definition database (> 34,064 , AbreMES) • Spanish medical term database generated by CUTEXT – automatic term recognition of the medical domain in Spanish from the medical literature, EHRs, etc. • Bilingual medical glossary for multiple language pairs (MeSpEN 400 thousand) • Spanish medical vocabularies by deep learning based medical machine translation • Medical Word embeddings for Spanish
Resources and components of preprocessing & linguistic annotation	• Document standardization software • Spanish clinical text automatic sectionizer • Spanish medical spellchecker component • Spanish medical text tokenization component • Spanish medical sentence splitting component • Spanish medical Part-of-speech labelling (grammatical label) component • Spanish medical word lemmatizer component
Medical concept recognition and semantic tagging	• Gazetteer-lookup: System of recognition of terms from a terminology (tested with SNOMED CT) • Automatic Term recognition (ATR): a system called CUTEXT for the recognition of terms using linguistic-statistical standards • PharmaCoNER Tagger: deep learning based system for detecting drugs, chemical compounds and genes in clinical texts • Time expression entity recognition system HeidelTime grammar for temporal tagging of Spanish EHRs

Corpus anotados tipo Gold Standard

<https://github.com/PlanTL-SANIDAD> <https://zenodo.org/communities/medicalnlp>

Corpus	Descripción / resumen
Annotated corpora & annotation schema	• Corpus and annotation schema for drugs, chemical compounds, genes/proteins (PharmaCoNER) • Corpus and annotation schema for abbreviations and definition pair annotations (BARR1) • Corpus and annotation schema for abbreviations resolution and normalization to SNOMED-CT (BARR2) • Corpus and annotation schema for sensitive protected health information entity mentions (MEDDOCAN): de-identification • Corpus for manually annotated clinical coding with CIE10 together with annotation evidence text labels • Corpus for manually annotated clinical entity mentions and their mapping to SNOMED-CT for: diseases, treatments, drugs, symptoms and procedures • Corpus for manually annotated entity mentions and their relations: drug-drug target relations (English – ChemProt)

Escenario típico de tareas/campañas de evaluación



Tareas en biomedicina/clínica desde perspectiva PLN



TASS
Taller de Análisis Semántico
en la SEPLN

BioNLP

BARR

BARR2

DIANN

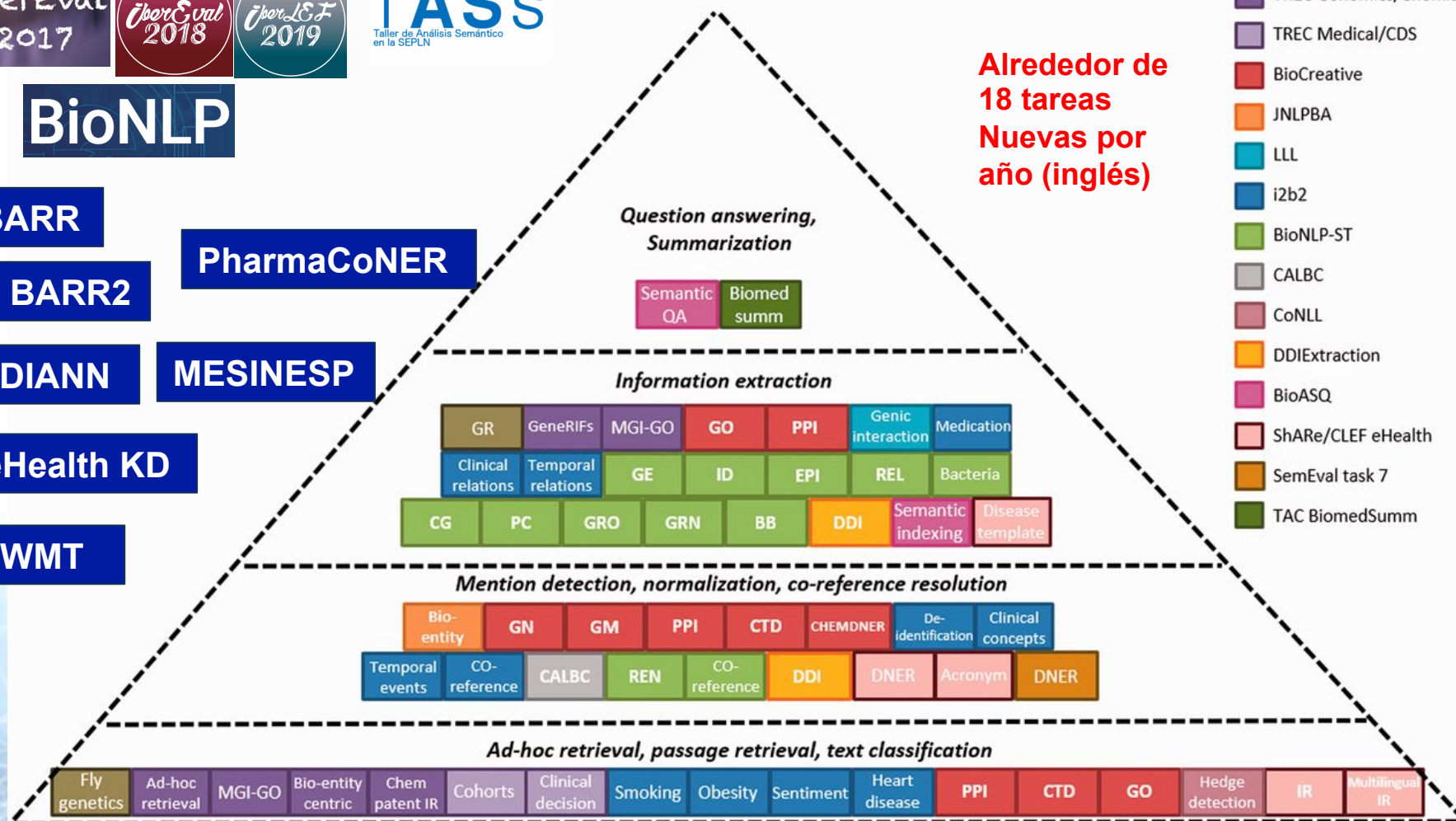
MESINESP

eHealth KD

WMT

PharmaCoNER

Alrededor de
18 tareas
Nuevas por
año (inglés)



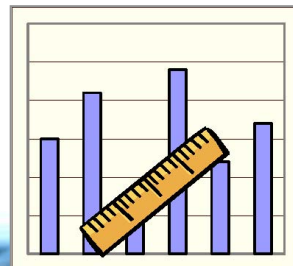
- KDD Cup
- TREC Genomics/Chemistry
- TREC Medical/CDS
- BioCreative
- JNLPBA
- LLL
- i2b2
- BioNLP-ST
- CALBC
- CoNLL
- DDIExtraction
- BioASQ
- ShARe/CLEF eHealth
- SemEval task 7
- TAC BiomedSumm

Chung-Chi Huang, and Zhiyong Lu Brief Bioinform
2015;bib.bbv024

Briefings in
Bioinformatics

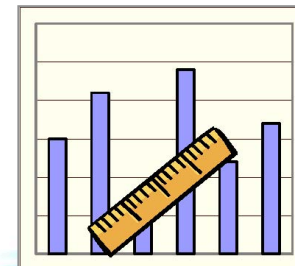


- Evaluaciones comparativas objetivas y independientes
- Interoperabilidad, armonización de recursos y sistemas combinados
- Generación de nuevos recursos y tecnologías
- Creación de confianza y soluciones de calidad Industria TL
- Promover el desarrollo de la industrial de TL y PYMEs
- Ayuda a la toma de decisiones (que funciona)
- Reproducibilidad y replicabilidad
- Generalizabilidad y adaptabilidad de metodologías y sistemas





- Identificación de nuevos retos de relevancia
- Promover esfuerzos colaborativos y de cooperación
- Aspectos de educación y formación, adquisición de nuevos conocimientos y aptitudes
- Promover investigación y desarrollo entre distintas disciplinas científicas y técnicas
- Mejoras en perspectivas laborales, de empleo y perspectivas académicas
- Facilitar el acceso a conjuntos de datos
- Impacto económico (de los datos, los recursos, conocimiento generado)



Impacto económico de campañas de evaluación: caso de campañas TREC



En el 2010 se publicó el resultado del análisis sobre la evaluación de **impacto económico del programa TREC** (patrocinado por el **NIST y DARPA de EEUU**) desde 1999 al 2009.

Coste total	Coste uso	Beneficios	Beneficios extrapolados	Beneficios netos	Beneficios netos extrapolados
-\$29,046,000	-\$970,000	\$105,084,000	\$152,994,000	\$75,068,000	\$122,978,000

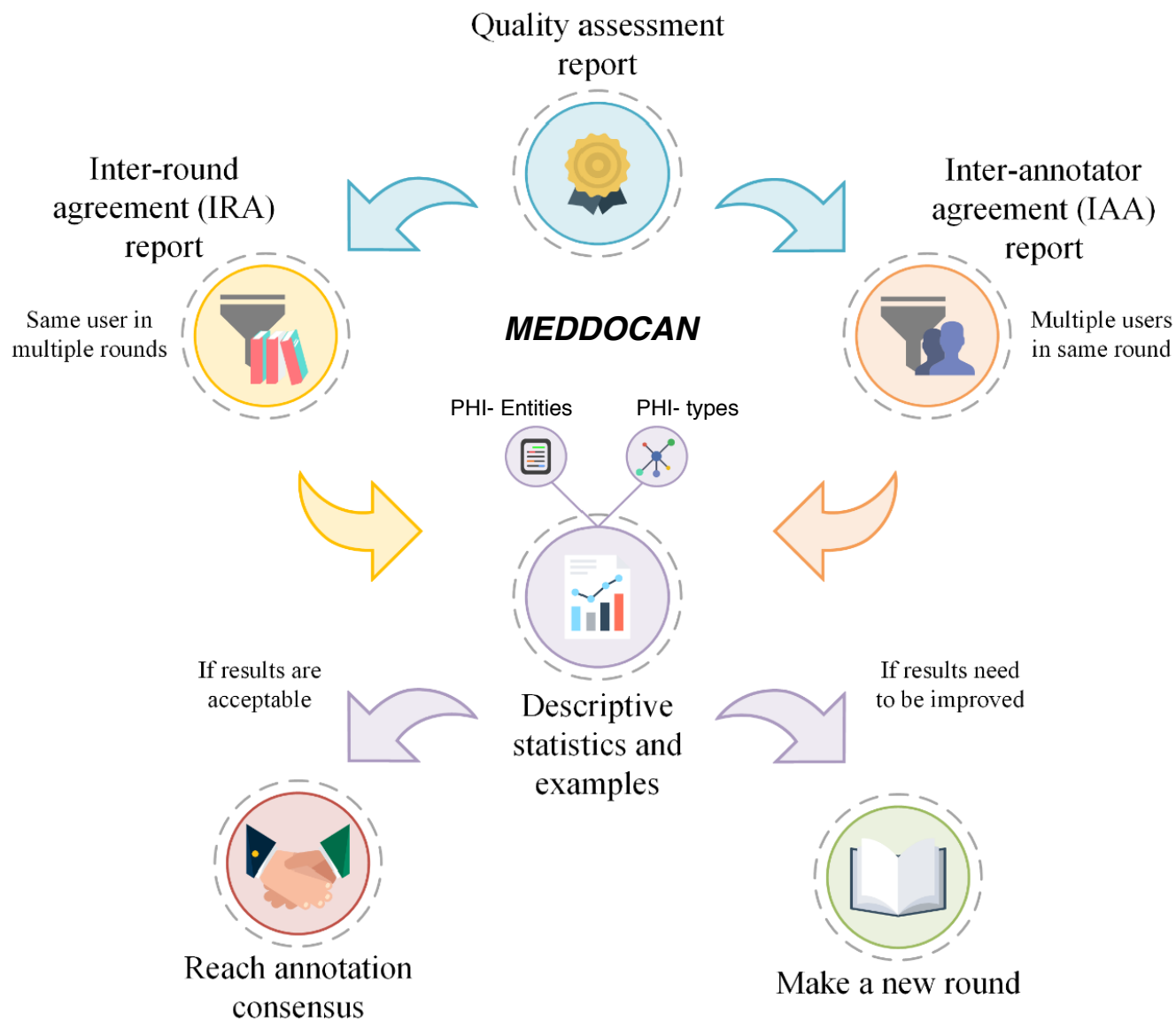
La tasa interna de retorno (IRR) se estimó en más del 250% para beneficios extrapolados y más del 130% para beneficios no extrapolados. (<http://trec.nist.gov/pubs/2010.economic.impact.pdf>)

Campaña/tarea	Resumen	Características	Enlace
MEDDOCAN (IberLEF) 	Task related to detection of sensitive entity mentions (de-identification of protected health information), based on adaption to GDPR of HIPPA-PHI (held at SEPLN/ IberEval 2019)	• Annotation guidelines • Manual Gold Standard corpus • IAA (Consistency/quality) • Format BRAT y i2b2 • AnnotateIt, BRAT • Synthetic corpus clinical cases	http://temu.bsc.es/meddocan
PHARMACONER (BioNLP-OST) 	Task related to recognition of medication/drugs, chemical compounds and genes. Subtask: • Mention recognition (NER) • SNOMED CT concept indexing	• Annotation guidelines • Manual Gold Standard corpus • IAA (Consistency/quality) • Format BRAT y i2b2 • AnnotateIt, BRAT • Clinical cases	http://temu.bsc.es/pharmaconer/
Mesinesp/ BioASQ Español 	Automatic indexing with structured medical terminologies in Spanish using DeCS (Descriptores en Ciencias de la Salud): medical literature (LILACS, IBECs)	• Indexing guidelines • Consistency assessment • Format BioASQ JSON • Indexing Experts • Medical Literature	http://temu.bsc.es/mesinesp/

Campañas de evaluación – Plan TL: futuras

Campaña/tarea	Resumen	Características	Enlace
CodiEsp	Automatic clinical coding and explainable AI task	• CIE10 guidelines • Consistency assessment • Format TSV • Clinical coding experts experts	TBA
MedTermMap Sub-tasks (CIE10, HPO, SNOMED CT)	Automatic mapping/ normalization of medical and clinical terms (from EHRs) to HPO (Human Phenotype Ontology)	• Mapping guidelines • Quality evaluation • Format TSV y JSON • Medical terminology Experts (clinical phenotypes)	TBA
MedTrans	Automatic machine translation of medical terms from English to Spanish (focus on diseases and UMLS-related terms)	• Translation guidelines • Manual translation by professionals • Format TSV	TBA

Control de calidad y consistencia



Guías/esquemas de anotación y control de calidad



<http://temu.bsc.es/pharmaconer>

Varón de 38 años de edad alérgico a Penicilina, bebedor de 80 gramos de alcohol/día y obeso que acude al Servicio de Urgencias de nuestro Hospital por presentar un cuadro de edemas en las extremidades inferiores, distensión abdominal y febrícula de dos días de evolución. Refiere además astenia importante de varias semanas de evolución acompañada de náuseas, vómitos y diarrea en los últimos 7 días. A la exploración física destaca la presencia de 37,5° C de temperatura, datos de ascitis abdominal y edemas en ambas extremidades inferiores, principalmente en la derecha, asociados en este miembro a eritema, petequias y equimosis. No se aprecian otros datos patológicos a la exploración.

En la analítica de ingreso se obtuvieron los siguientes resultados: Hemoglobina 8,3 gr/dl; Hematocrito:23,3%;Leucocitos 20.420 por µl (neutrófilos 91,5%); Plaquetas 119.000 por µl; Dimeros D 14.080 ng/dl; Urea: 178 mg/dl; Creatinina 9 mg/dl; Na 124 mEq/l; K 3,9 mEq/l; Proteínas totales 5,6 gr/dl, LDH 559 UI/l; CPK 239 UI/l; GPT 35 UI/l; GOT 77 UI/l. Se realizó una radiografía de tórax que era normal y una Ecografía y TAC abdominales que reflejaban una ascitis masiva, datos de hepatopatía crónica y esplenomegalia. Con el juicio clínico de insuficiencia renal aguda, en el contexto de un hepatópata crónico de origen enólico y celulitis en extremidad inferior ingresa en el Servicio de Nefrología. Se instaura un tratamiento con diuréticos (Furosemida) y antibioterapia empírica con Ciprofloxacino (1gr/ 24 horas) tras extracción de hemocultivos. A las 24 horas del ingreso el paciente presenta fiebre (38,4° C) y empeoramiento de las lesiones en miembro inferior derecho (MID), con aumento del dolor, extensión de la celulitis y presencia de ampollas. En la analítica se objetiva un empeoramiento en la función renal con valores de creatinina plasmática de 10,60 mg/dl y urea 181 mg/dl, un aumento de la leucocitosis (35.340 por µl, neutrófilos 96,8 %) y alteraciones en la coagulación (tiempo de protrombina de 28,8 segundos y tiempo de tromboplastina parcial activada de 61,4 segundos). En el hemocultivo realizado al ingreso se aísla *Streptococo Pyogenes*, por lo que se inicia antibioterapia intravenosa con Clindamicina y Gentamicina y es ingresado en la Unidad de Cuidados Intensivos (UCI) por presentar inestabilidad hemodinámica y progresión rápida de las lesiones en extremidad inferior visible en pocas horas, con anestesia cutánea, grandes ampollas hasta el tercio medio de muslo y afectación escrotal. Precisa ventilación mecánica invasiva, aminas vasoactivas y hemofiltración veno-venosa continua y se indica intervención quirúrgica urgente en la que se realiza desbridamiento escrotal, desbridamiento de fascia hasta raíz de muslo y amputación abierta supracondílea.

Presenta una evolución desfavorable con fracaso multiorgánico (fracaso renal agudo, coagulopatía y síndrome de distrés respiratorio agudo) no respondiendo a medidas de soporte hemodinámico ni a antibioterapia y fallece finalmente a las 24h de la cirugía.

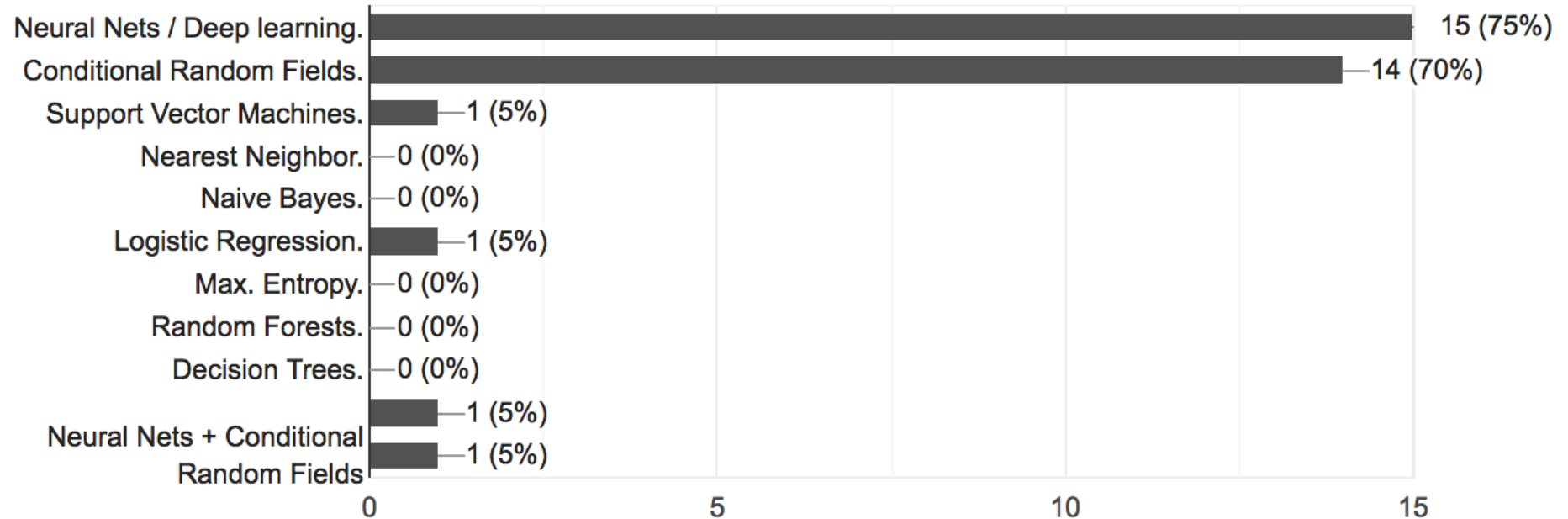
Corpus Gold Standard: Expertos del dominio y expertos en PLN y lingüistas

T1	NORMALIZABLES	2548	2554	aminas
#1	AnnotatorNotes	T1		43201005
T2	NORMALIZABLES	2223	2234	Gentamicina
#2	AnnotatorNotes	T2		387321007
T3	NORMALIZABLES	2208	2220	Clindamicina
#3	AnnotatorNotes	T3		372786004
T4	PROTEINAS	2034	2048	tromboplastina
#4	AnnotatorNotes	T4		387124009
T5	PROTEINAS	1993	2004	protrombina
#5	AnnotatorNotes	T5		7348004
T6	NORMALIZABLES	1866	1870	urea
#6	AnnotatorNotes	T6		387092000
T7	NORMALIZABLES	1827	1837	creatinina
#7	AnnotatorNotes	T7		15373003
T8	NORMALIZABLES	1477	1491	Ciprofloxacino
#8	AnnotatorNotes	T8		372840008
T9	NORMALIZABLES	1435	1445	Furosemida
#9	AnnotatorNotes	T9		387475002
T11	PROTEINAS	1029	1032	GOT
#10	AnnotatorNotes	T11		26091008
T12	PROTEINAS	1016	1019	GPT
#11	AnnotatorNotes	T12		56935002
T13	PROTEINAS	1003	1006	CPK
#12	AnnotatorNotes	T13		75828004
T14	PROTEINAS	989	992	LDH
#13	AnnotatorNotes	T14		259319003
T15	PROTEINAS	960	977	Proteinas totales
#14	AnnotatorNotes	T15		395835001
T16	NORMALIZABLES	948	949	K
#15	AnnotatorNotes	T16		88480006
T17	NORMALIZABLES	934	936	Na
#16	AnnotatorNotes	T17		39972003
T18	NORMALIZABLES	914	924	Creatinina
#17	AnnotatorNotes	T18		15373003
T19	NORMALIZABLES	897	901	Urea
#18	AnnotatorNotes	T19		387092000
T20	PROTEINAS	873	882	Dimeros D
#19	AnnotatorNotes	T20		25607008
T21	PROTEINAS	758	769	Hemoglobina
#20	AnnotatorNotes	T21		38082009
T22	UNCLEAR	72	79	alcohol
#21	AnnotatorNotes	T22		53041004
T23	NORMALIZABLES	36	46	Penicilina
#22	AnnotatorNotes	T23		323389000

ciprofloxacino (sustancia)
SCTID: 372840008
372840008 | ciprofloxacino (sustancia) |
en Ciprofloxacino (substance)
en Ciprofloxacino
es ciprofloxacina
es ciprofloxacino (sustancia)
es ciprofloxacino

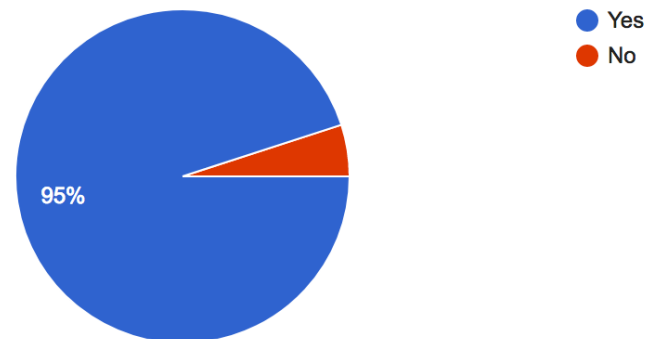
Maapeo (normalizacion/ grounding) a terminologias normativas

Cuestionario participantes (MEDDOCAN): PLN & IA

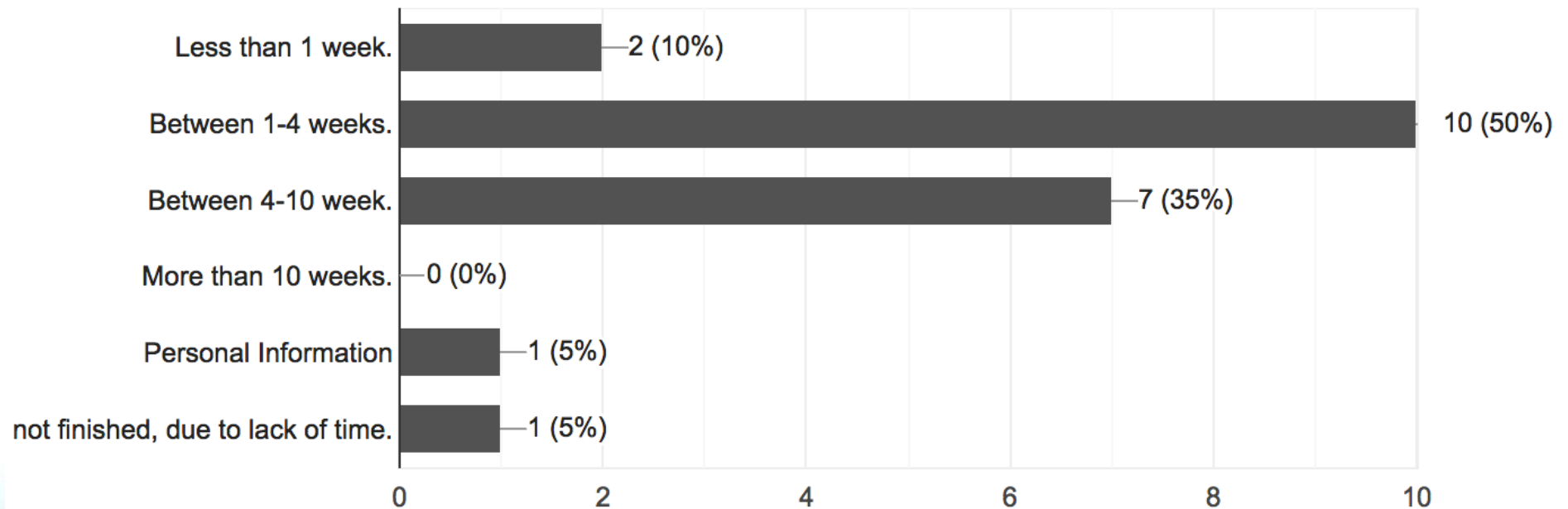


Did you use machine learning/AI methods?

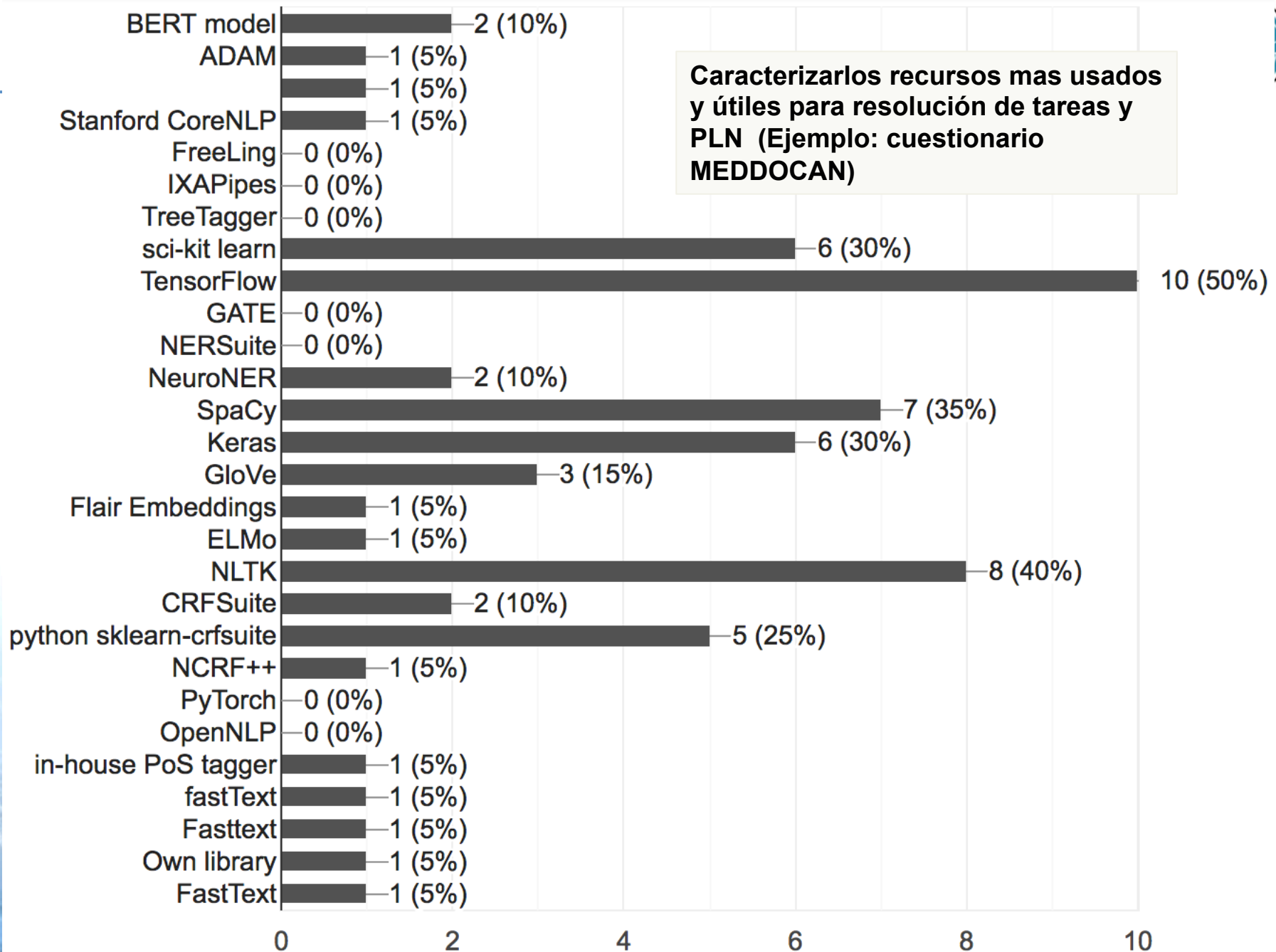
20 responses



Campañas/tareas: reducción en tiempo

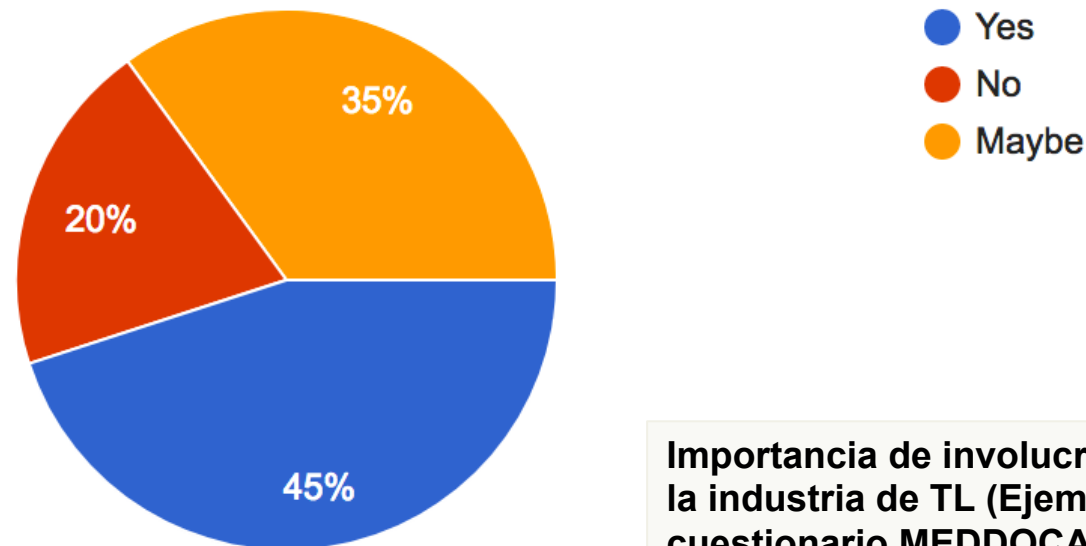


**Resultado basado en cuestionario MEDDOCAN:
Campañas de evaluación como mecanismo de
reducción en tiempo/coste y optimización en el
desarrollo de soluciones PLN**



Would you be interested in support for turning your system into a software product /start up?

20 responses

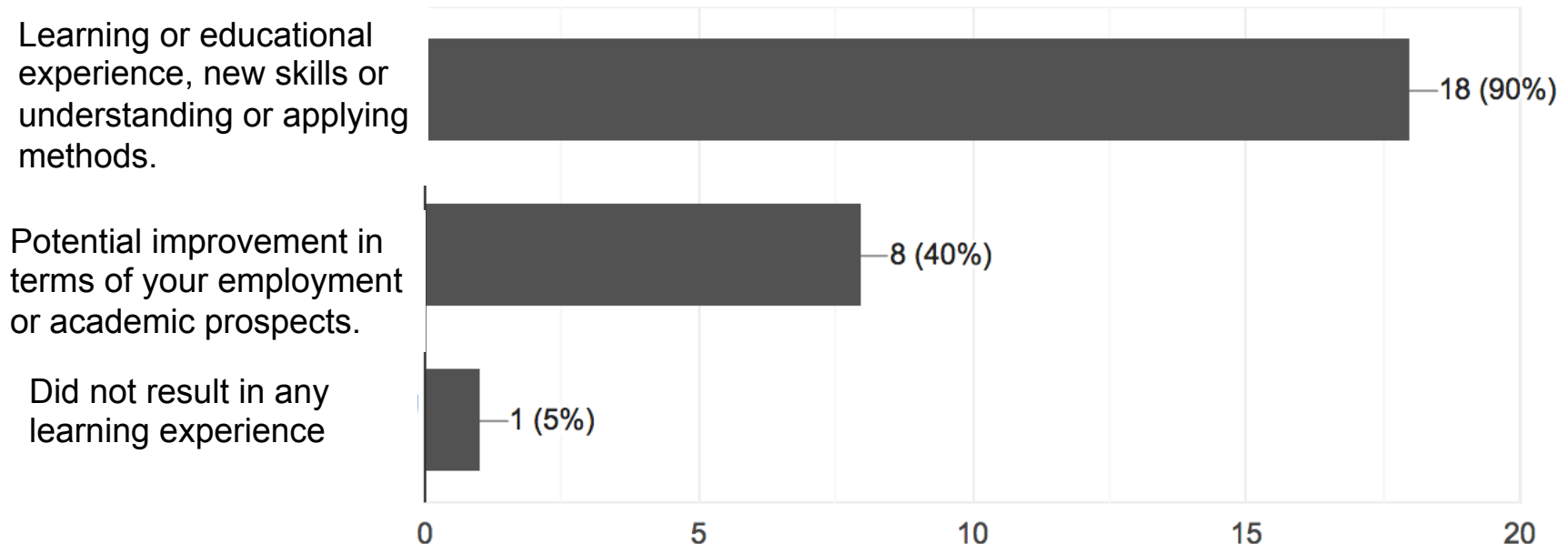


Importancia de involucrar agentes de la industria de TL (Ejemplo: cuestionario MEDDOCAN)

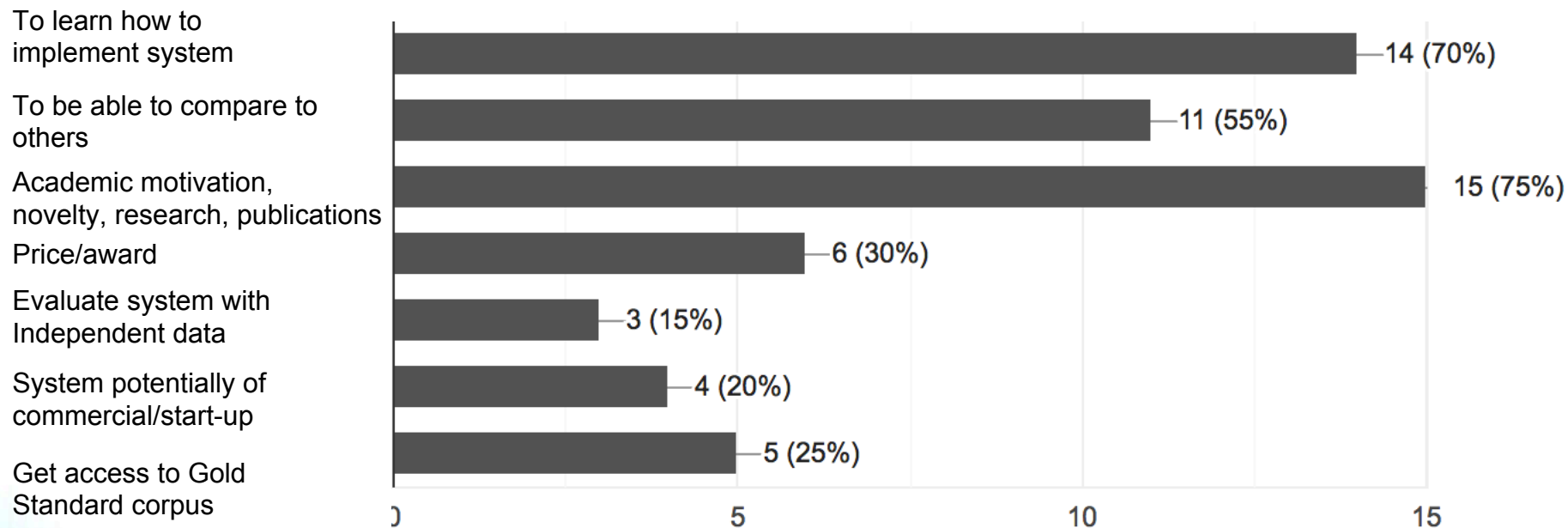
Campañas/tareas: estrategia para formación más aplicada a resolución de problemas

Do you think that your participation in the MEDDOCAN track resulted in
a:

20 responses



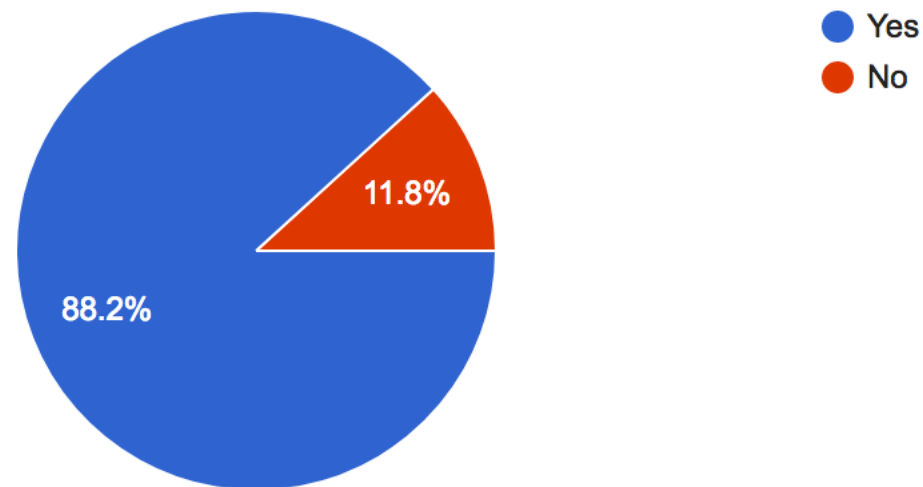
Campañas/tareas: motivación



Campañas/tareas: promover desarrollo de código abierto/accesible

Would you be able to provide a software (GitHub or similar) or webservice based on your system if there is some specific technical or financial support for this?

17 responses



Thanks!



- Martin Krallinger
- Marta Villegas
- Siamak Barzegar
- Aitor Gonzalez
- Montse Marimon
- Felipe Soares
- Alfonso Valencia (BSC Life)

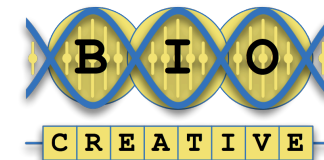


David Perez (SEAD)

- *AQuAS* (Miguel Gallofre López)
- *AEMPS-BIFAP* (Julio Bonis Sanz)
- *AEMPS-FTM* (JM Simarro)
- *FID-Salud/MSSSI* (Elena García)
- *FISEVI/Hosp. Virgen del Rocio* (Carlos Parra)
- *Hospital 12 de Octubre* (Pablo Serrano)
- *IBECS/Carlos III* (Elena Primo)
- *Informática Médica Hosp. Clínic* (Raimundo Lozano)
- *MSSSI* (Maribel García Fajardo)
- *RANM* (Cristina V. González)



- Obdulia Rabal
- Julen Oyarzabal



- BioCreative organizers
- Cecilia Arighi/Cathy Wu (Uni. Delaware)
- Lynette Hirschman (MITRE)

UniversidadeVigo

- Analia Lourenço
- Martin Perez Perez
- Gael Perez Rodriguez
- Florentino Fernández Riverola



Gold Standard annotated texts – corpora: Abbreviations (BARR2 Track at IberEval)

- 1 Varón de 59 años que ingresó en septiembre de 2008 con diagnóstico de quilotórax posquirúrgico. Antecedentes de cáncer de sigma, practicándose en julio de 2008 una metastasectomía hepática, siendo necesario realizar desde entonces varias toracocentesis evacuadoras de quilotórax (volumen extraído aproximado: 9.000 mL).
- 2 Al ingreso refirió astenia, molestias en hemitórax derecho y discreta dificultad respiratoria. Se objetivó radiológicamente derrame pleural que ocupaba dos tercios del hemitórax derecho. Tras colocación de un drenaje pleural se obtuvieron 500 mL de un líquido de aspecto lechoso en las primeras 24 horas, cuyo análisis fue compatible con quilotórax: colesterol (COL) 60 mg/dL, triglicéridos (TG) 515 mg/dL, cociente COL líquido pleural/suero: 0,28; cociente TG líquido pleural/suero: 3,18. Se decidió tratamiento conservador con dieta absoluta y drenaje torácico, consultándose con el Área de Nutrición para iniciar soporte nutricional parenteral.
- 3 En la valoración nutricional destacaba una pérdida de peso de un 4% en los últimos dos meses (habitual 74 kg; actual 71 kg), albúmina 2,6 g/dL, transferrina 154 mg/dL, proteína C reactiva (PCR) 16,6 mg/dL. Presentaba ligera pérdida de grasa subcutánea y masa muscular, junto con una capacidad funcional reducida en domicilio en las últimas semanas; ausencia de edemas maleolares y signos o síntomas de déficits vitamínicos concretos. Historia dietética sin cambios recientes.
- 4 La situación nutricional fue compatible con desnutrición energético-proteica leve, planteándose el siguiente plan de cuidados nutricionales:
- 5 • Objetivos: Mantener en reposo el tracto gastrointestinal y prevenir un mayor grado de desnutrición en un paciente con elevadas pérdidas de nutrientes por el drenaje pleural
- 6 • Requerimientos nutricionales: necesidades calóricas estimadas mediante la ecuación de Harris-Benedict (factor de corrección de 1,4), de 2.200 kcal/día; necesidades proteicas 1,5-1,7 g/kg/día de aminoácidos; necesidades estándar de micronutrientes.
- 7 • Composición de la fórmula de nutrición parenteral (NP): Volumen 2.500 mL; 120 g de aminoácidos; 300 g de glucosa; 60 g de lípidos (MCT/LCT); 2280 kcal totales; vitaminas y oligoelementos según recomendaciones AMA-ASPEN.
- 8 A pesar del tratamiento, el drenaje pleural se mantuvo estable en los primeros días, persistiendo el aspecto quiloso. La ecografía mostró derrame pleural loculado, administrándose en consecuencia 200.000 UI de urokinasa durante dos días consecutivos por el tubo de toracocentesis. Ante la persistencia de drenaje quiloso abundante, se decidió iniciar tratamiento con octreótido, a dosis de 100 mcg/8 h por vía subcutánea. El drenaje se redujo progresivamente en los días siguientes, cambiando a un aspecto seroso. Paralelamente se produjo un empeoramiento del estado general del paciente, con distensión abdominal, oliguria e hiperglucemia. Ante la sospecha de reacción adversa a octreótido se suspendió al 4o día de tratamiento. La bioquímica mostró anemia (hemoglobina 9,4 g/dL; hematocrito 28,8%), leucopenia (2.390

Example of MEDDOCAN annotation



1	Datos del paciente.
2	Nombre: Pedro.
3	Apellidos: Jimenez Ramos.
4	NHC: 4763954.
5	NASS: 47 37584930 84.
6	Domicilio: Calle del pez, 28.
7	Localidad/ Provincia: Madrid.
8	CP: 28001.
9	Datos asistenciales.
10	Fecha de nacimiento: 20/05/2000.
11	País: España.
12	Edad: 16 años Sexo: H.
13	Fecha de Ingreso: 26/08/2017.
14	Servicio: Urgencias.
15	Médico: Luis Moyano Calvo N°Col: 28 31 23567.
16	Informe clínico del paciente: Adolescente Varón de diecisiete años sin antecedentes de interés que acude por
17	En la analítica de orina se aprecian 30-50 hematias por campo. Urocultivo negativo.
18	Se practica ecografia abdominal observándose pequeña lesión de medio centímetro de diámetro, sólida con refuerzo hiperecogénico anterior.
19	Realizamos cistoscopia observándose en cara lateral derecha, por fuera de orificio ureteral dos pequeñas lesiones sobrelevadas, con muco:
20	Sospechándose lesión inflamatoria se prescribe tratamiento con A.I.N.E. durante diez días sin que desaparezcan las lesiones, decidiéndose in
21	Se realiza RTU de ambas lesiones vesicales, siendo el informe anatomopatológico el de leiomioma vesical, describiendo la lesión como "pro
22	Remitido por: Dr. Luis Moyano Calvo C/ Eduardo Rivas, 3 28018 Madrid, España. e-mail: joseluis Moyano@ya.com

MEDDOCAN track corpus



PHI entity types	PHI entity types
NOMBRE_SUJETO_ASISTENCIA	CALLE
EDAD_SUJETO_ASISTENCIA	TERRITORIO
SEXO_SUJETO_ASISTENCIA	PAIS
FAMILIARES_SUJETO_ASISTENCIA	NUMERO_TELEFONO
NOMBRE_PERSONAL_SANITARIO	NUMERO_FAX
FECHAS	CORREO_ELECTRONICO
PROFESION	ID_SUJETO_ASISTENCIA
CENTRO_SALUD	ID_CONTACTO_ASISTENCIAL
HOSPITAL	NUMERO_BENEF_PLAN_SALUD
INSTITUCION	ID_ASEGURAMIENTO
ID_TITULACION_PERSONAL_SANITARIO	URL_WEB
ID_EMPLEO_PERSONAL_SANITARIO	DIREC_PROT_INTERNET
IDENTIF_VEHICULOS_NRSERIE_PLACAS	IDENTIF_BIOMETRICOS
IDENTIF_DISPOSITIVOS_NRSERIE	OTRO_NUMERO_IDENTIF
	OTROS_SUJETO_ASISTENCIA

Desde casos de uso a campañas de evaluación



FISEVI



agencia española de
medicamentos y
productos sanitarios



Use case
definition

Design of text
Processing workflow

Implementation of
executable PLN
workflow

Platform
adaptation

Functional
evaluation

Usability
Evaluation

aquas.gencat.cat



Instituto de Salud Carlos III



Son Espases
hospital universitari



Define required
component
set

Identify
Currently missing
components

**Implement
Missing
Components**

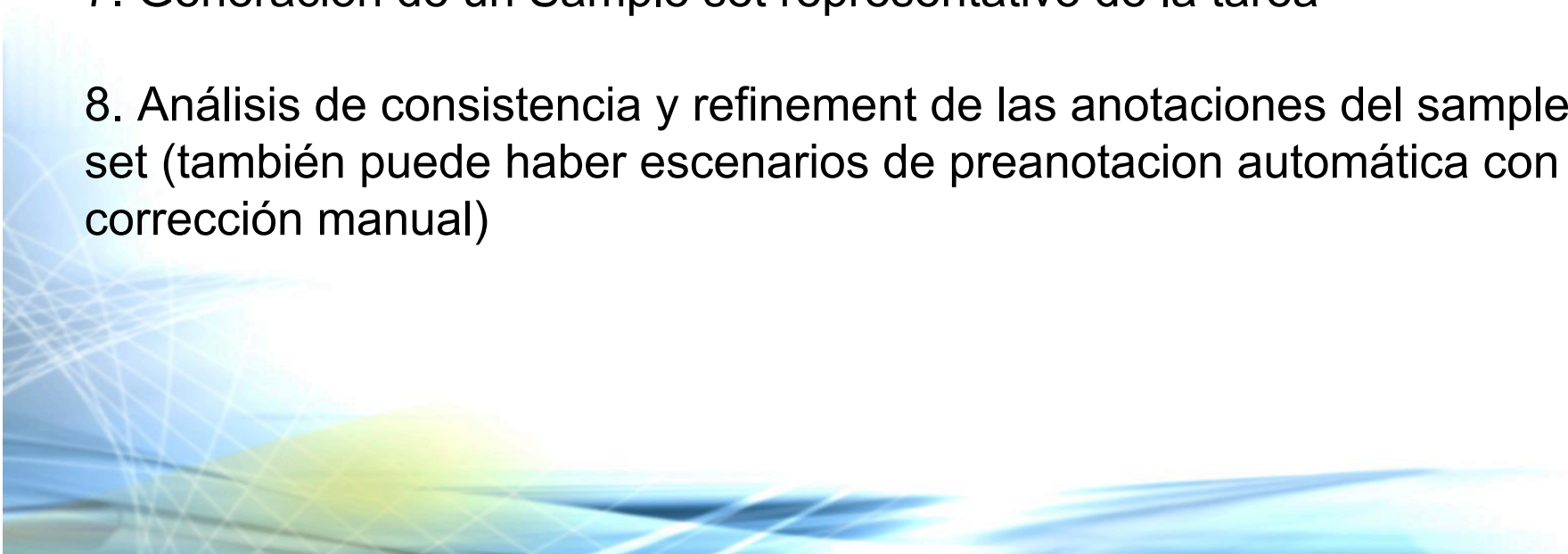
**Shared
tasks**

Real world use cases
with high impact and
which generate
resources and
linguistic
infrastructure

Aspectos relevantes

1. Definición de la tarea y estudio comparativo de la especificación de la tarea (tareas similares, complementarias, en otros idiomas o parecida pero en otro tipo de documento o dominio)
2. Identificación y desarrollo de propuestas para campañas de evaluación (aquí hay deadlines específicas)
3. Selección, preprocesamiento y generación de corpus documental (conseguir la colección de documentos a anotar, hacer un sampling representativo, análisis de aspectos legales de re-/distribución, análisis de formatos, preprocesamiento. Si se va a usar active learning puede haber una dependencia entre el proceso de anotación, el sistema y la selección de documentos)
4. Desarrollo de guías y reglas de anotación (estudio de guías de tareas similares, complementarias, en otros idiomas o parecida pero en otro tipo de documento o dominio)

5. Análisis, benchmarking y adaptación de herramienta de anotación (cual es mas eficiente, rápido genera anotaciones mas eficientes, tutorial de uso para anotador, instalación)
6. Recruitment de expertos para la definición de las guías de anotación (y para la anotación del corpus entero, pueden ser dos grupos), requiere definir el perfil del anotador y training
7. Generación de un Sample set representativo de la tarea
8. Análisis de consistencia y refinement de las anotaciones del sample set (también puede haber escenarios de preanotacion automática con corrección manual)



9. Análisis de tiempo y costes de anotación
10. Anotación de corpus (anotación del training , development, test y post-campaña set)
11. Infraestructura de evaluación (interfaz, visualización y definición de las métricas de evaluación)
12. Comparación de las predicciones (sistemas automáticos) con el gold standard (anotaciones manuales, definir por ejemplo si son matches exactos , parciales, ..)
13. Diseminación and recruitment de participantes, org. eventos



MULTIPLE FORMATS

Predictions in challenge format (TSV, BioC, JSON...)



User uploads predictions

RESULTS

Analysis report is sent to participant's email



NO INSTALLATION

User doesn't need to install any program



Markyt
Mining Evaluation system



<http://www.markyt.org/mmes>

Markyt evaluates predictions

Markyt shows different prediction statistics

FILE VALIDATION

Validates that prediction file is compliant with competition format



...ives	
...e Negatives:	2347
...alse Positives:	23474
Evaluated results:	33543
Number of annotations in gold set:	33543
...otal documents:	7000
...ated documents:	552
... in prediction results:	

RESULTS EXPLORATION

User may explore FP and FN to assess further improvement to its prediction system.



You		
LD STANDARD (MANUAL)	FALSE	
	TRUE POSITIVES	NEGATIVES
FAMILY Color	4547	7662
SYSTEMATIC Color	2786	6784
FORMULA Color	869	1370
TRIVIAL Color	1592	7106
PREVIATION Color	190	398
Color	49	

TP, FN AND FP

Statistics by annotation type, text mention and document



MICRO & MACRO AVERAGE STATISTICS

Calculation of precision, recall and F-score against gold standard

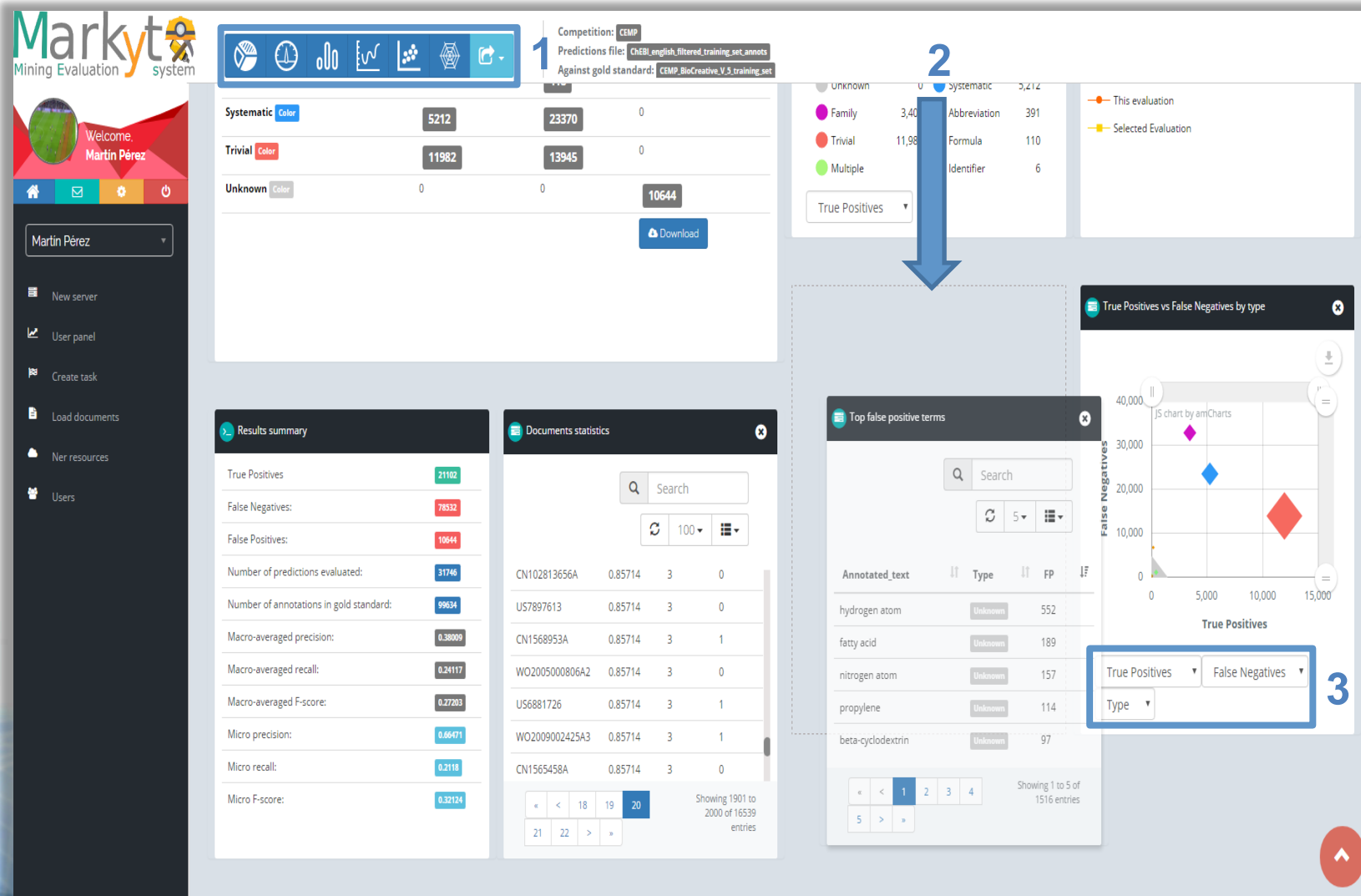


HISTORICAL EVALUATION

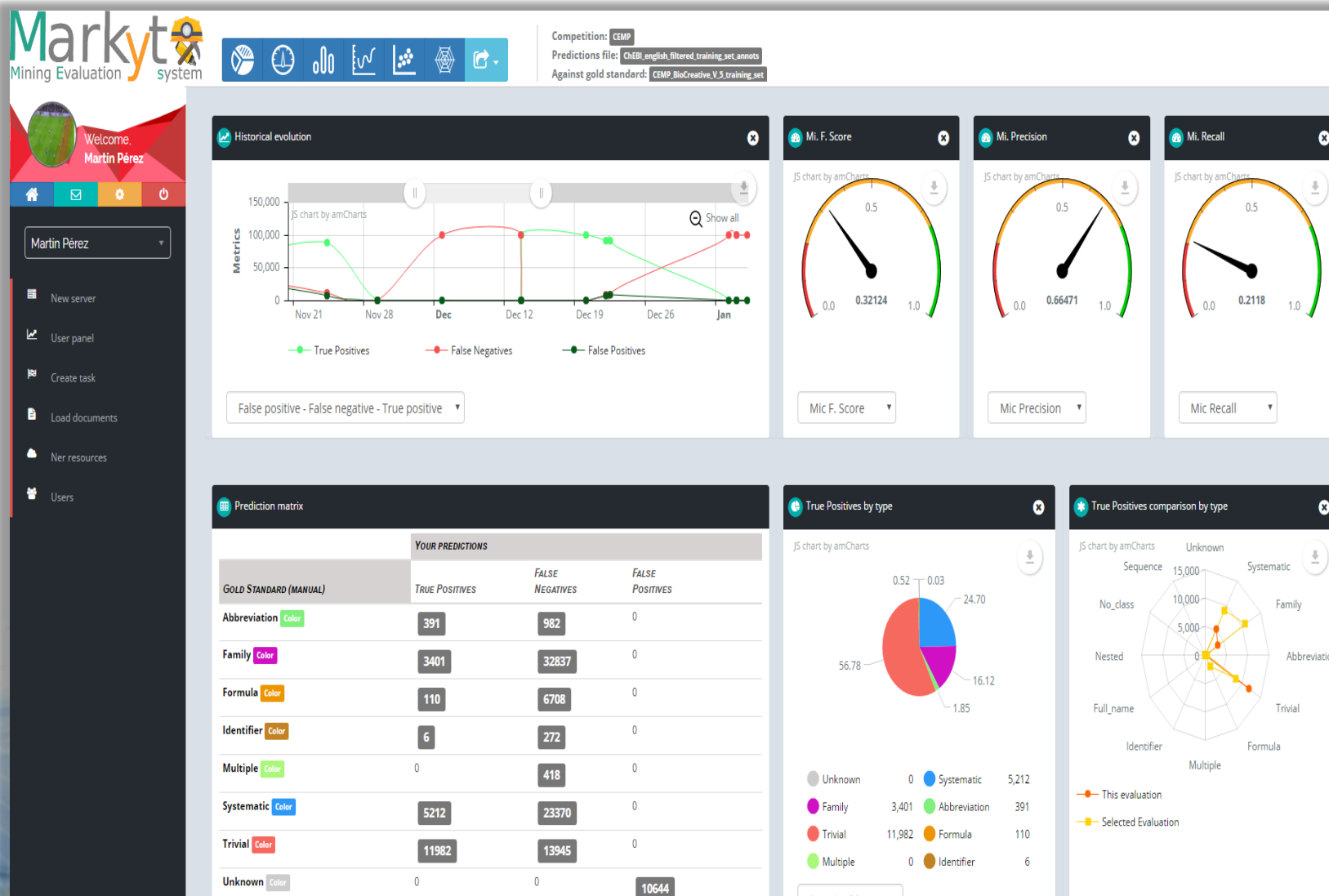
Markyt calculates the score of the prediction against CHEMDNER 2015 ranking

Markyt statistics

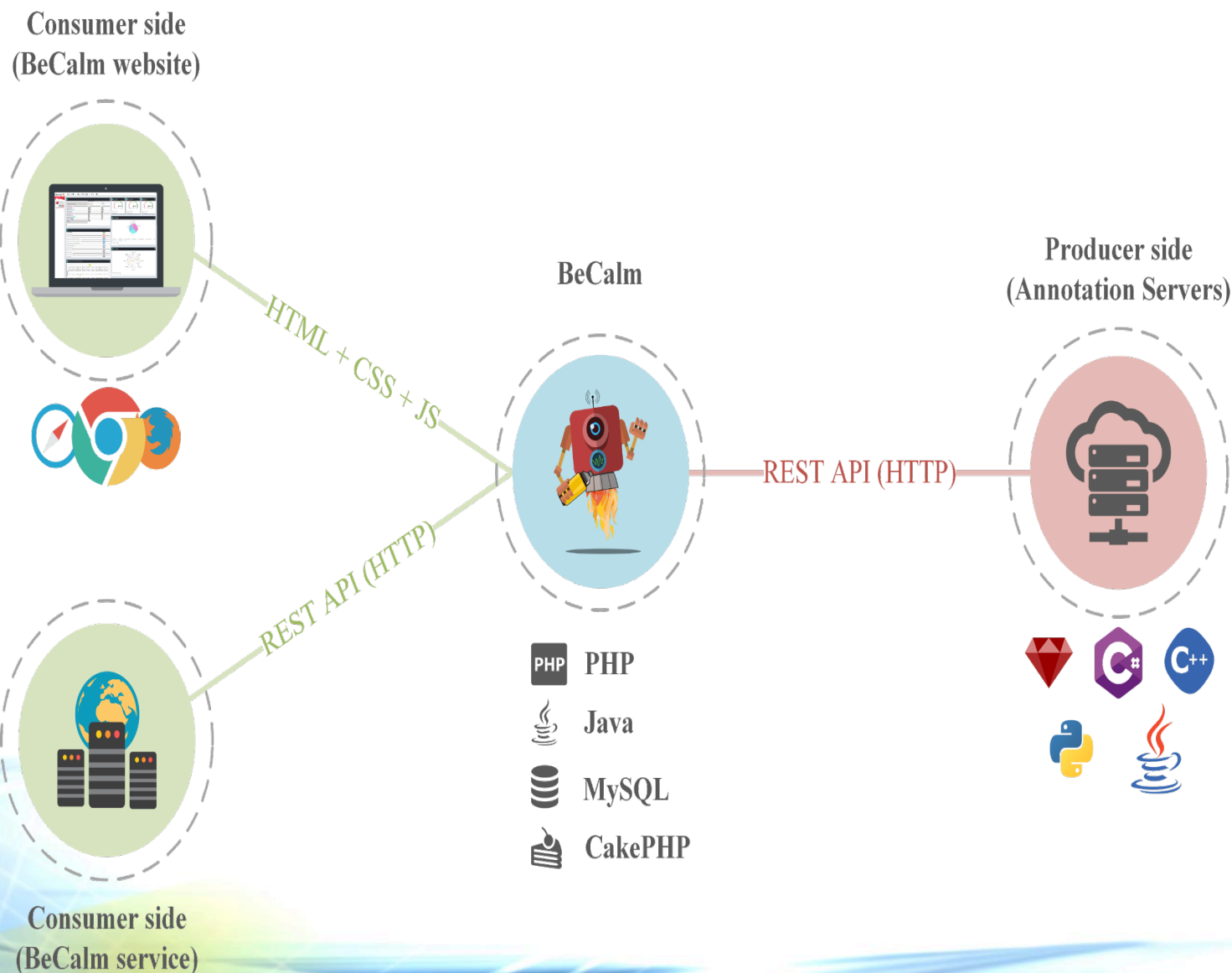
<http://www.markyt.org/mmes>



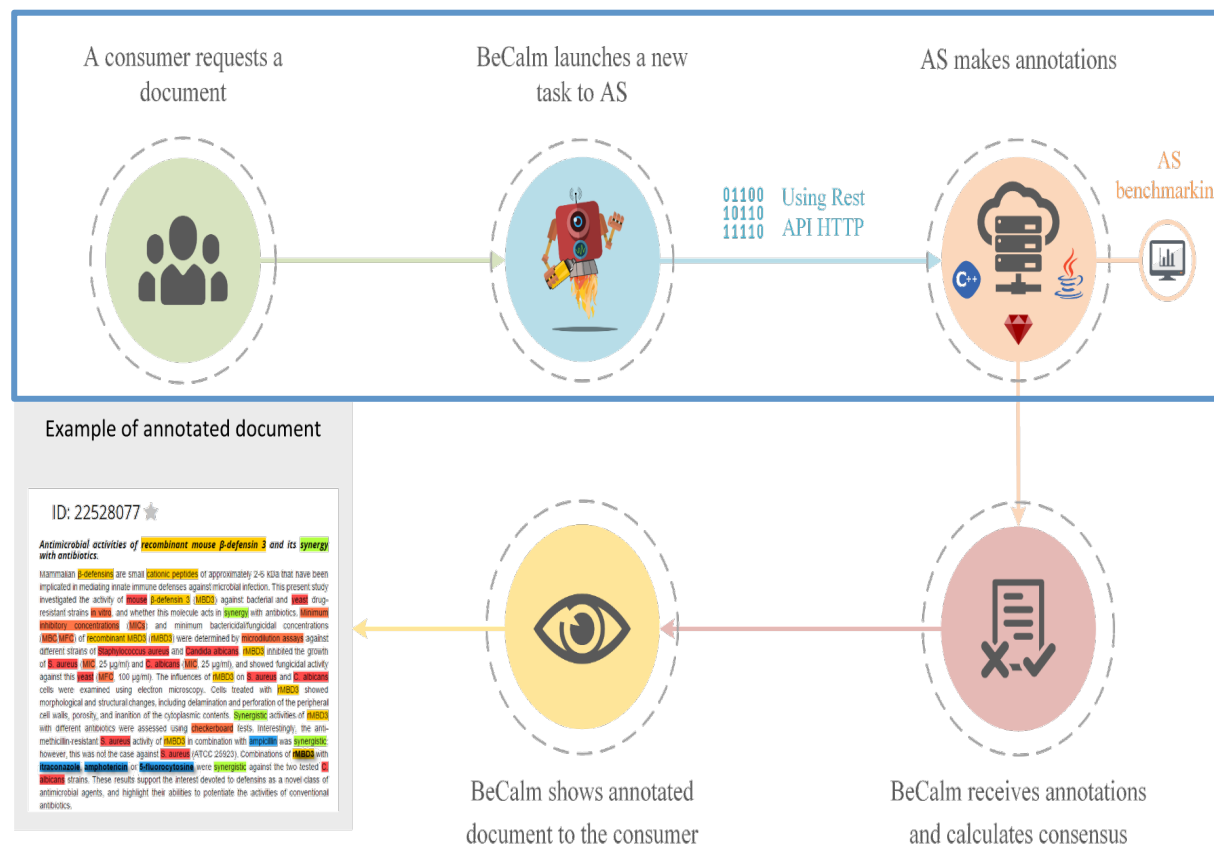
Markyt statistics



Becalm architecture



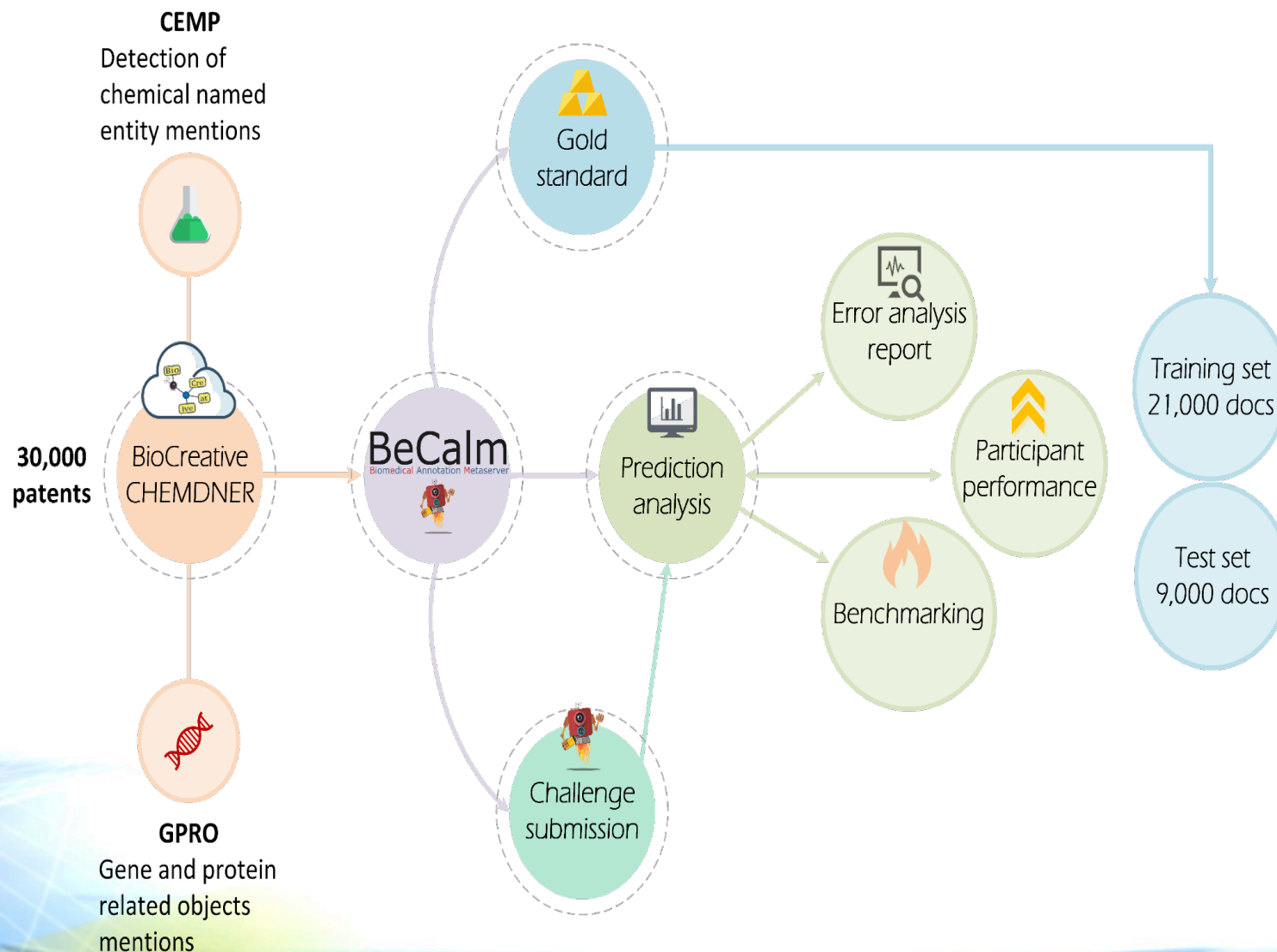
Becalm workflow



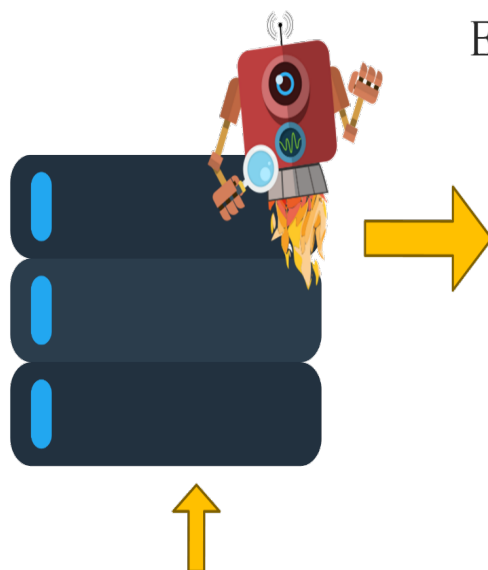
BeCalm
Biomedical Annotation Metaserver



CEMP/GPRO @ Becalm



Annotation Servers @ tips



Evaluation metrics

- Mean time between failures (MTBF)
- Mean time to repair (MTTR)
- Mean annotations per document (MAD)
- Mean time per document volumen (MTDV)
- Mean time seek annotations (MTSA)
- Average response time (ART)

