



Actuaciones de la Oficina Técnica de Sanidad

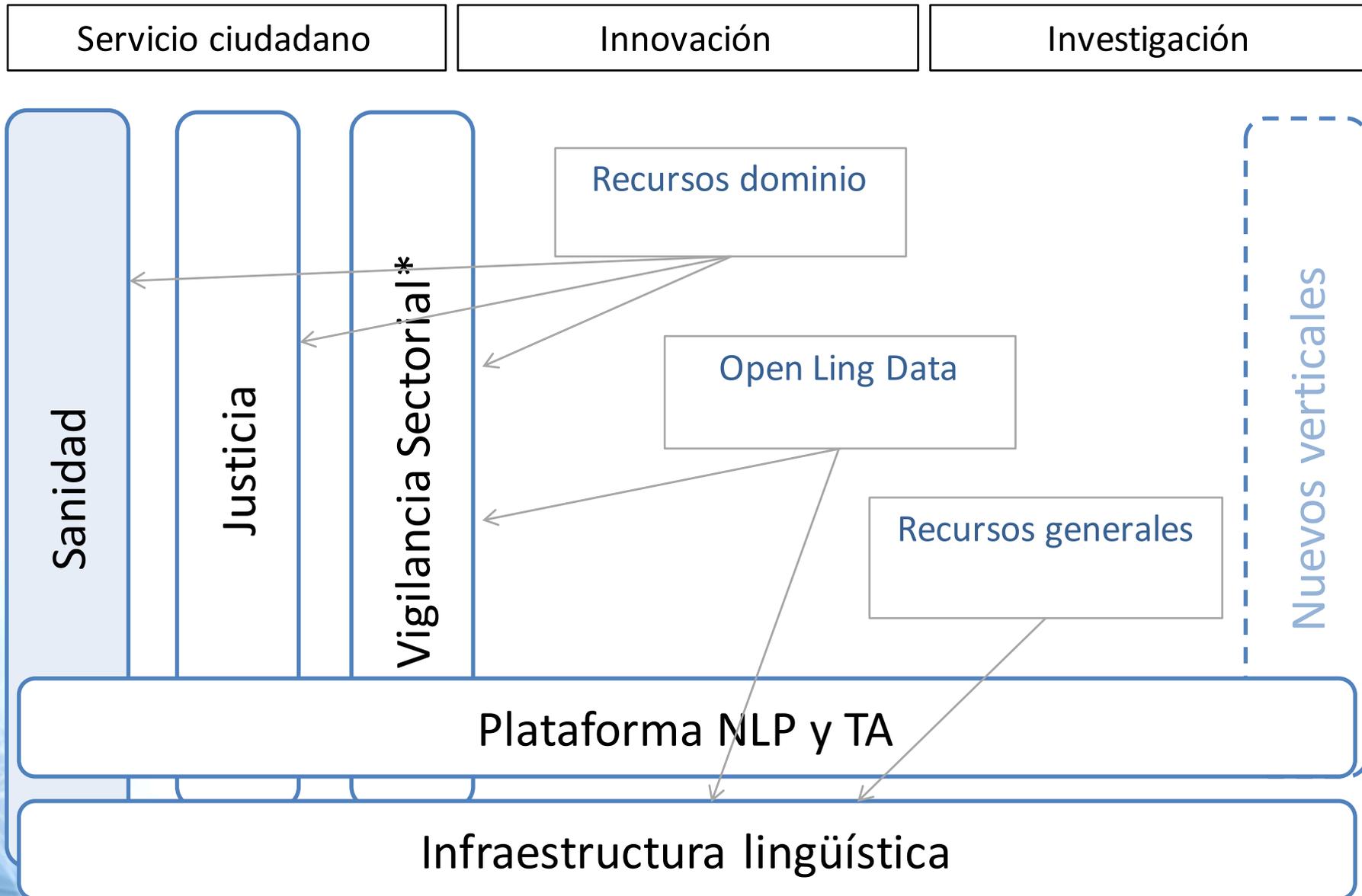
Martin Krallinger

Head of Biological Text Mining Unit
Spanish National Cancer Research Centre



[***martin.krallinger@bsc.es***](mailto:martin.krallinger@bsc.es)

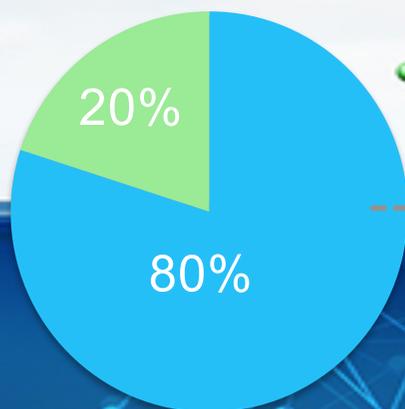
Mapa del Plan TL: proyectos faro - Sanidad



*Inteligencia competitiva

Datos clínicos

80% de datos sanitarios
"bloqueada" en texto
clínico no estructurado



Datos clínicos estructurados



Fuente de información para:

- Apoyo a la decisión clínica
- Estratificación de cohorte de pacientes
- Efectos adverso de medicamentos
- Vigilancia sanitaria y farmacovigilancia
- Toma de decisión para gestión
- Reposicionamiento de fármacos etc.

Volumen de datos – Big Data:

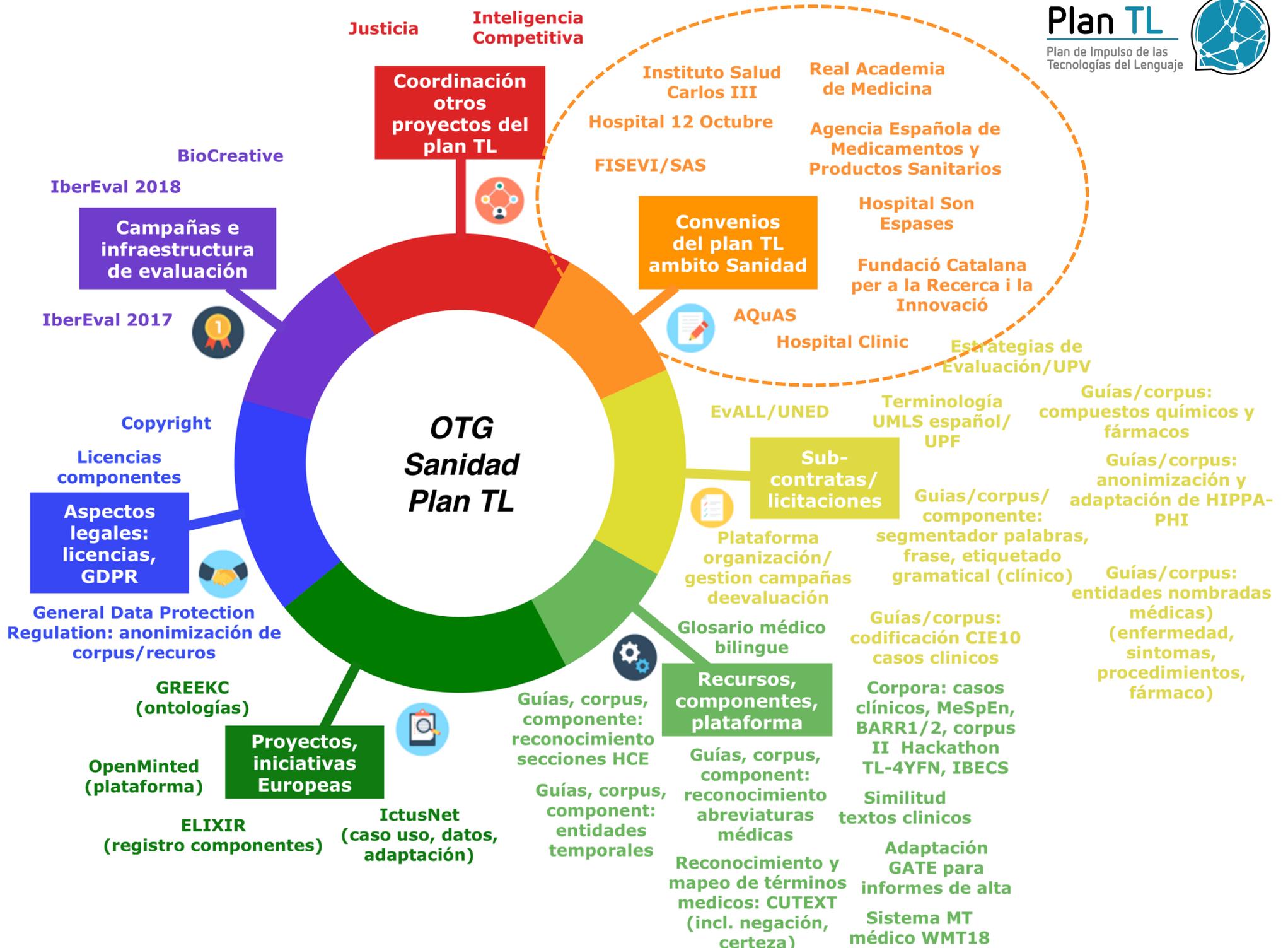
Solo en Galicia generan 200.000 notas clínicas por día (población de 2.7 millones)

Datos no estructurados: texto clínico narrativo, también imagen, video

Transformación del texto clínico generado por profesionales de la salud en representaciones estructuradas de datos clínicos -> para Big data/IA







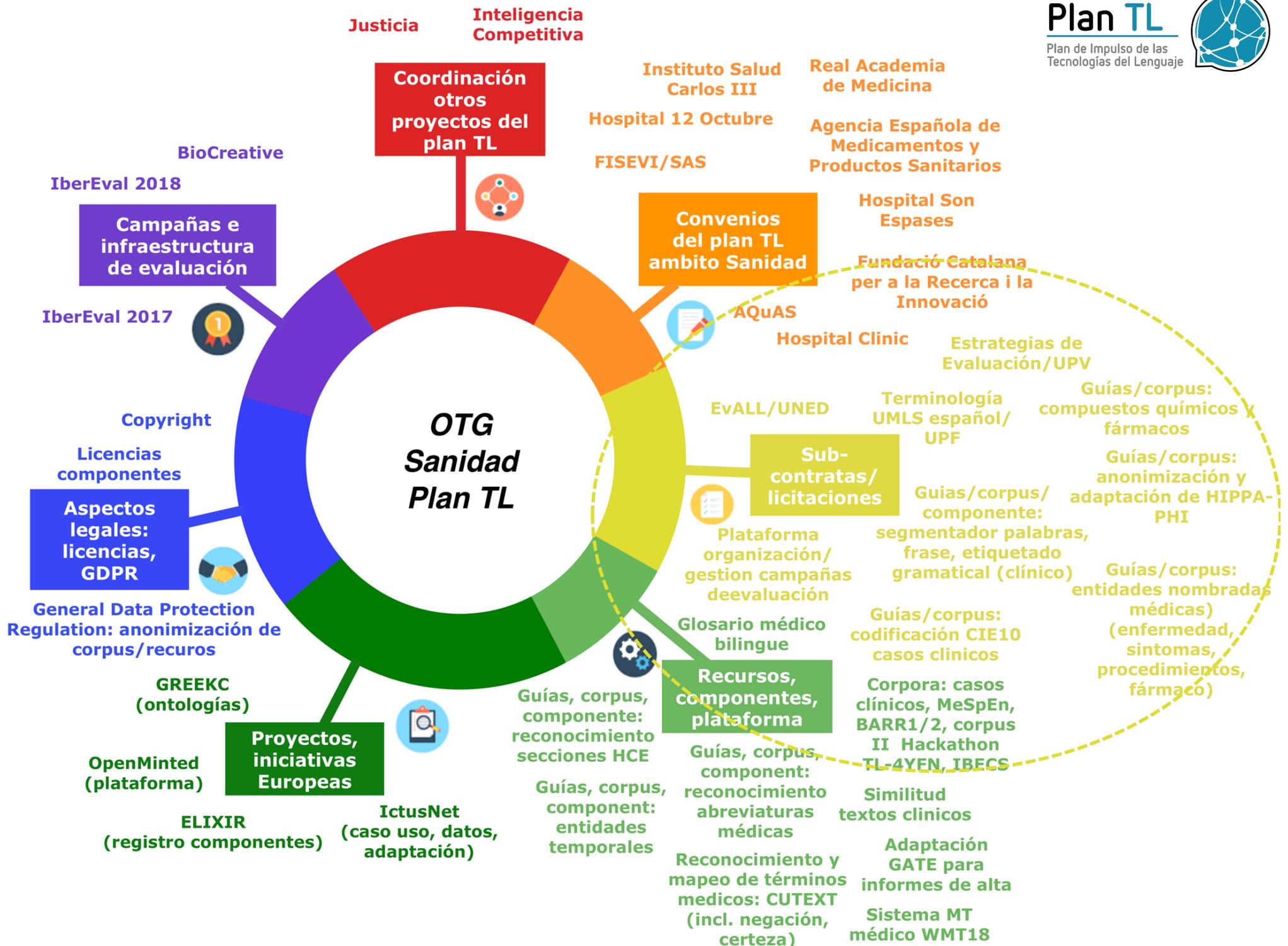
- Fármacos: Enriquecimiento y anotación semántica de fichas técnicas de medicamentos.
- Fármacos: Sistema de traducción automática de textos médicos y farmacéuticos.
- Fármacos: Sistema de PLN para farmacovigilancia y vigilancia sanitaria de textos clínicos de atención primaria y pediatría.
- Seguridad paciente: Identificación de eventos adversos asociados a procedimientos, medicación, a los cuidados, infección nosocomial.
- Privacidad/legal: Anonimización de textos clínicos (GDPR-HIPPA-PHI): protección de datos de carácter personal



- Neurología: (a) extracción de datos clínicos de informes de alta para mejora del proceso asistencial de ictus y (b) riesgo de ictus.
- Enfermedad infecciosas: Modelado del metabolismo del grupo farmacológico para una prescripción personalizada.
- Salud mental: observaciones ansiedad/ depresión.
- Atención Primaria: Codificación automática en tiempo real de Atención Primaria para motivos asociados a derivaciones a especialistas.



- Oncología: (a) PLN de observaciones clínicas en cáncer y (b) reconocimiento de entidades y anotación semántica para cáncer colorrectal.
- Imagen: Hallazgos en informes de imagen
- Microbiología: Ayuda a la decisión – antibióticos
- Pacientes: (a) identificación de conceptos en reclamaciones, (b) mejora gestión de prescripción de pacientes crónicos
- Servicios web procesamiento de literatura médica: soporte de indización y clasificación semántica, TA, detección de idioma, detección de tópicos.



Corpus anotados: entidades clínicas y mapeo a SNOMED-CT

Paciente de 70 años de edad, minero jubilado, sin **alergias medicamentosas** conocidas, que presenta como antecedentes personales: **accidente laboral** antiguo con **fracturas vertebrales** y **costales**; intervenido de **enfermedad de Dupuytren en mano** derecha y **by-pass iliofemoral** izquierdo; **Diabetes Mellitus tipo II**, **hipercolesterolemia** e **hiperuricemia**; **enolismo** activo, **fumador** de 20 cigarrillos / día.

Es derivado desde Atención Primaria por presentar **hematuria macroscópica postmiccional** en una ocasión y **microhematuria** persistente posteriormente, con **micciones normales**.

En la **exploración física** presenta un **buen estado general**, con abdomen y genitales normales; **tacto rectal** compatible con **adenoma de próstata** grado I/IV.

En la **analítica de orina** destaca la existencia de 4 hematíes/ campo y 0-5 leucocitos/campo; resto de sedimento normal.

Hemograma normal; en la bioquímica destaca una **glucemia** de 169 mg/dl y triglicéridos de 456 mg/dl; función hepática y renal normal. **PSA** de 1.16 ng/ml.

Las **citologías de orina** son repetidamente sospechosas de malignidad.

En la **placa simple de abdomen** se valoran **cambios degenerativos en columna lumbar** y **calcificaciones vasculares** en ambos hipocondrios y en pelvis.

La ecografía urológica pone de manifiesto la existencia de **quistes corticales simples en riñón derecho**, vejiga sin alteraciones con buena capacidad y próstata con un peso de 30 g.

En la **UIV** de ambos ureteres se observa **esfuerzo**.

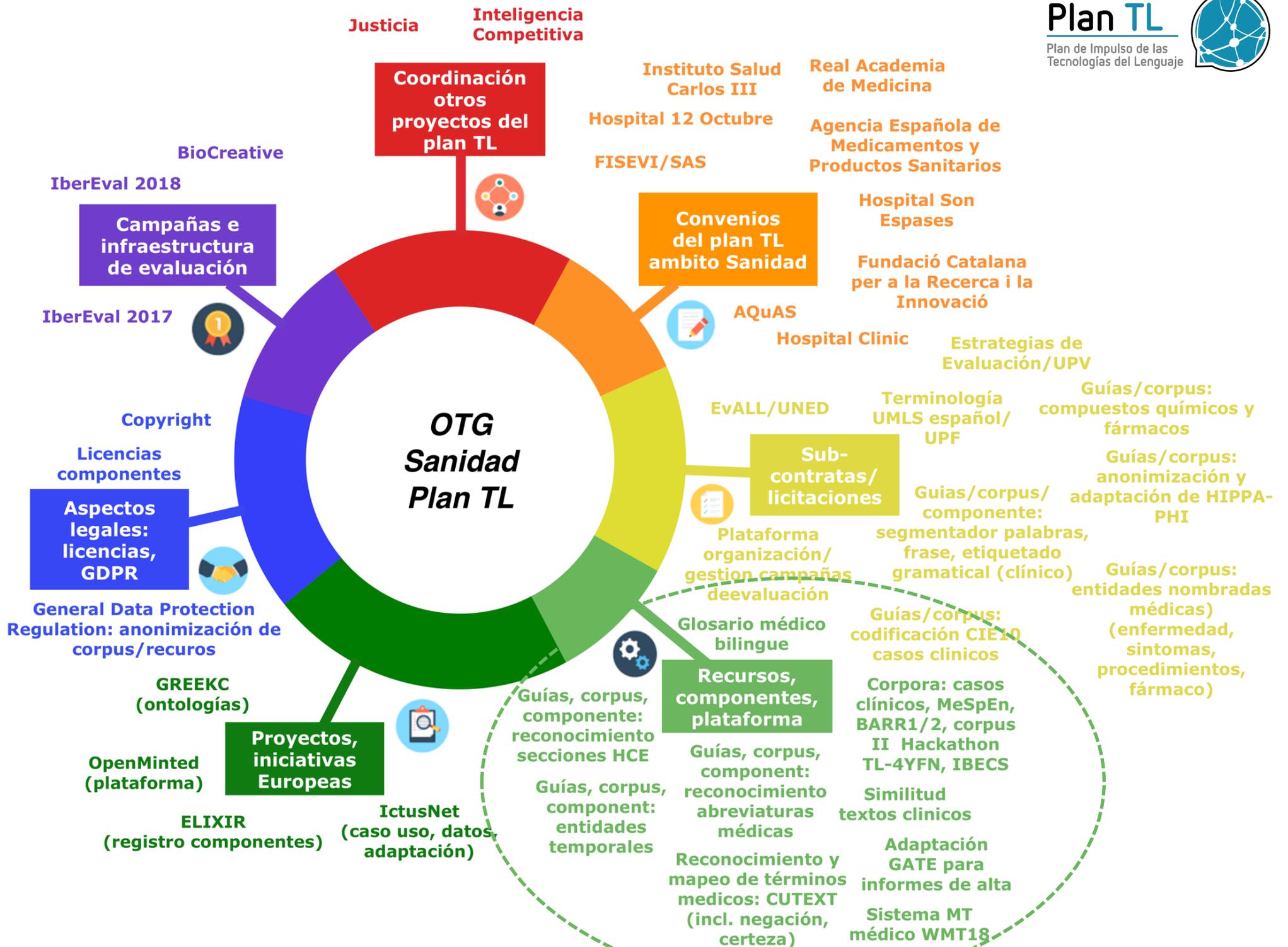
La **cistoscopia** superficial

de la **uretra** derecha y uréteres arrosariados con imágenes de adición en el tercio superior de la vejiga con buena capacidad, pero paredes trabeculadas en relación a vejiga de

se **resección transuretral** con el resultado anatomopatológico de **carcinoma urotelial**

- Guías de anotación
- Análisis de consistencia
- Selección de documentos
- Herramienta de anotación
- Corpus anotado
- Formato de corpus
- Licencia y metadatos
- Herramienta de evaluación
- Predicciones automáticas de etiquetas

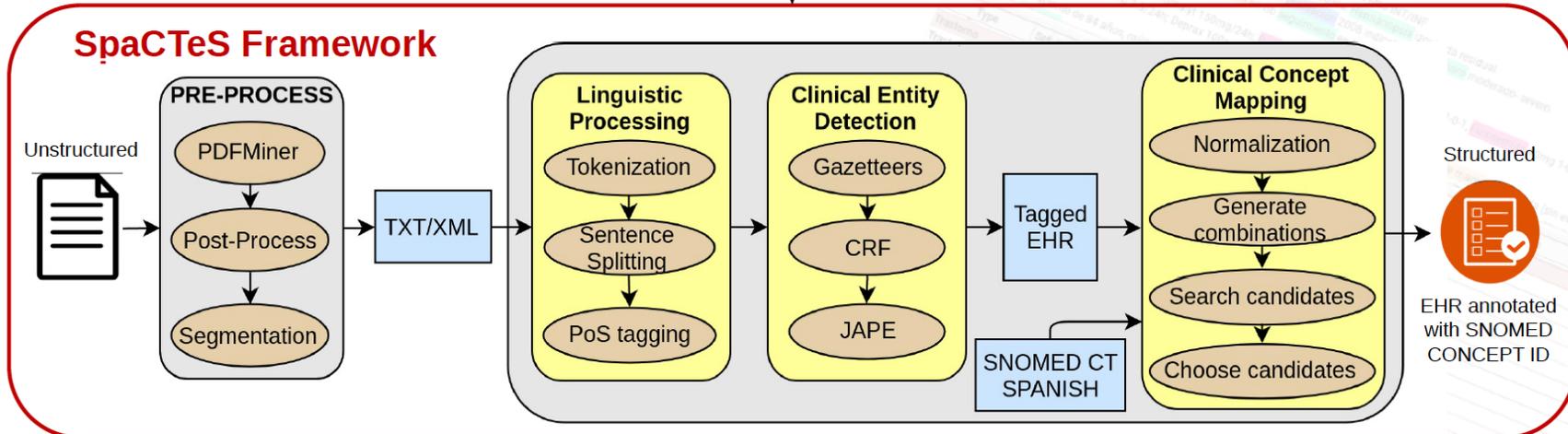
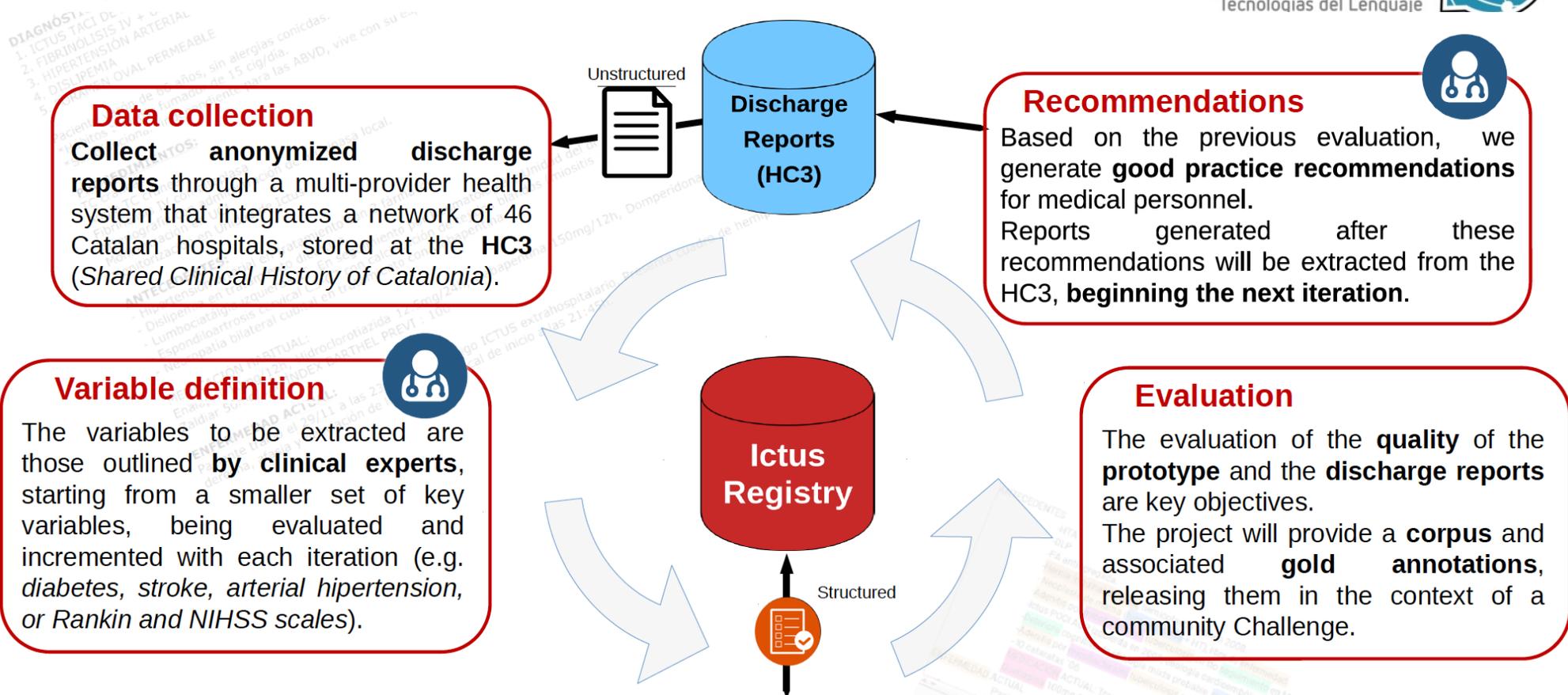
- SINTOMA
- ENFERMEDAD
- PROCEDIMIENTO_DIAGNOSTICO
- PROCEDIMIENTO_TERAPEUTICO
- FARMACO



Corpus anotados: abreviaturas BARR2

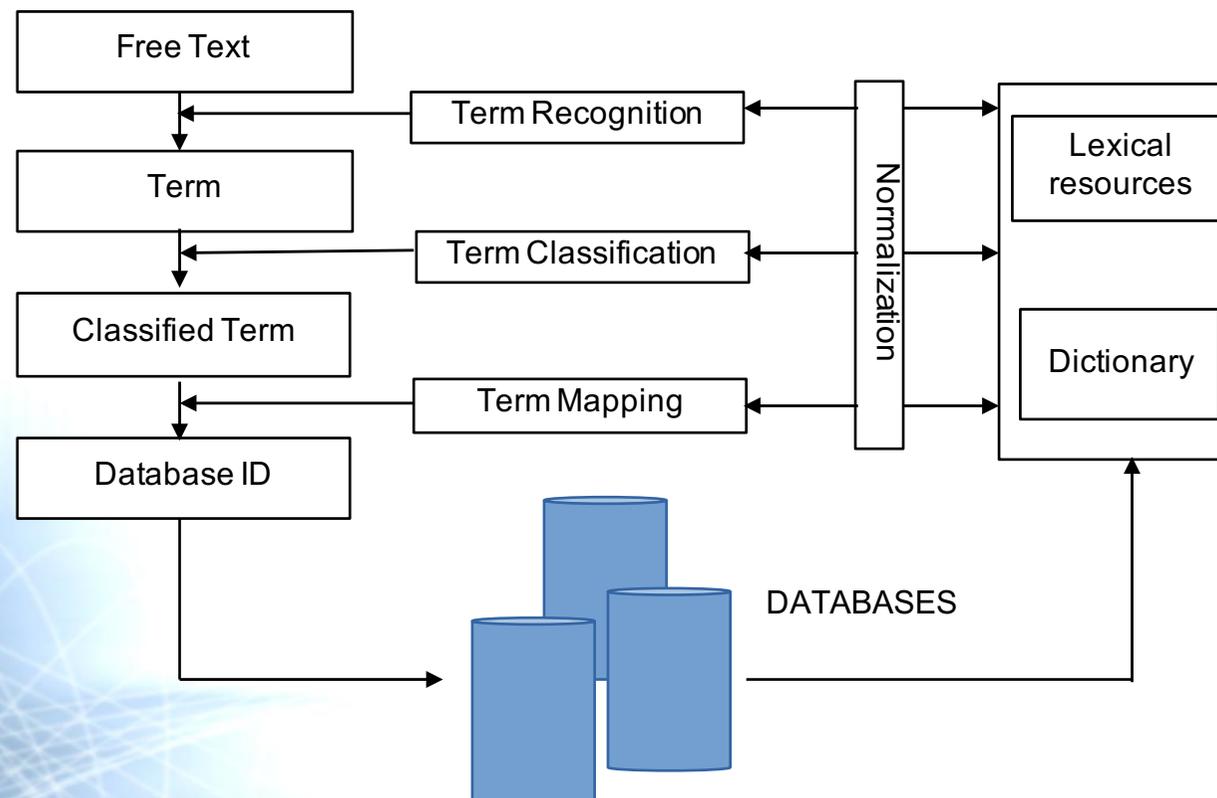
- 1 Varón de 59 años que ingresó en septiembre de 2008 con diagnóstico de quilotórax posquirúrgico. Antecedentes de cáncer de sigma, practicándose en julio de 2008 una metastasectomía hepática, siendo necesario realizar desde entonces varias toracocentesis evacuadoras de quilotórax (volumen extraído 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 quilotórax3: 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

Adaptación y uso de GATE: informes de alta Ictus



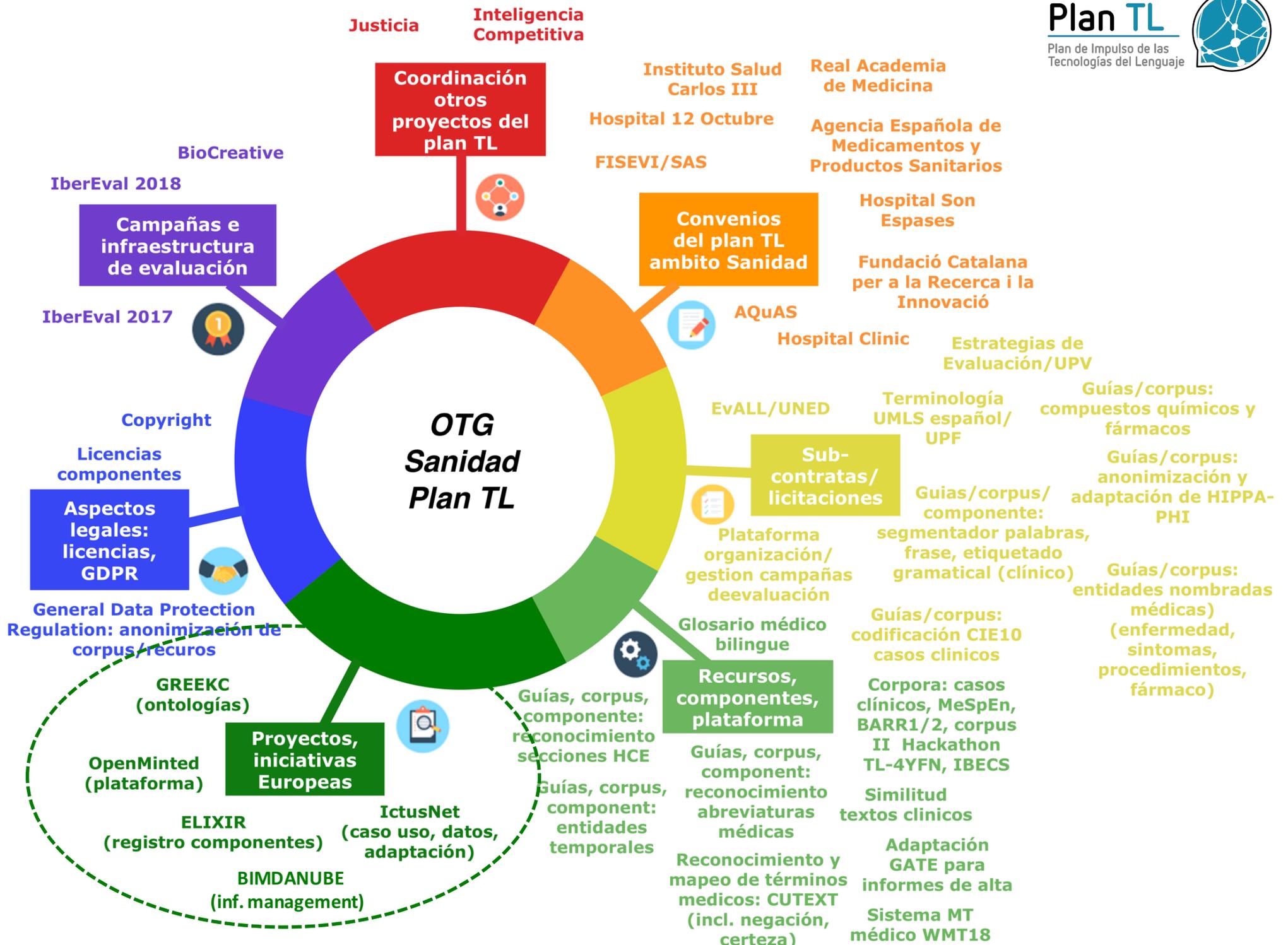
CUTEXT extracción automática y mapeo de términos médicos

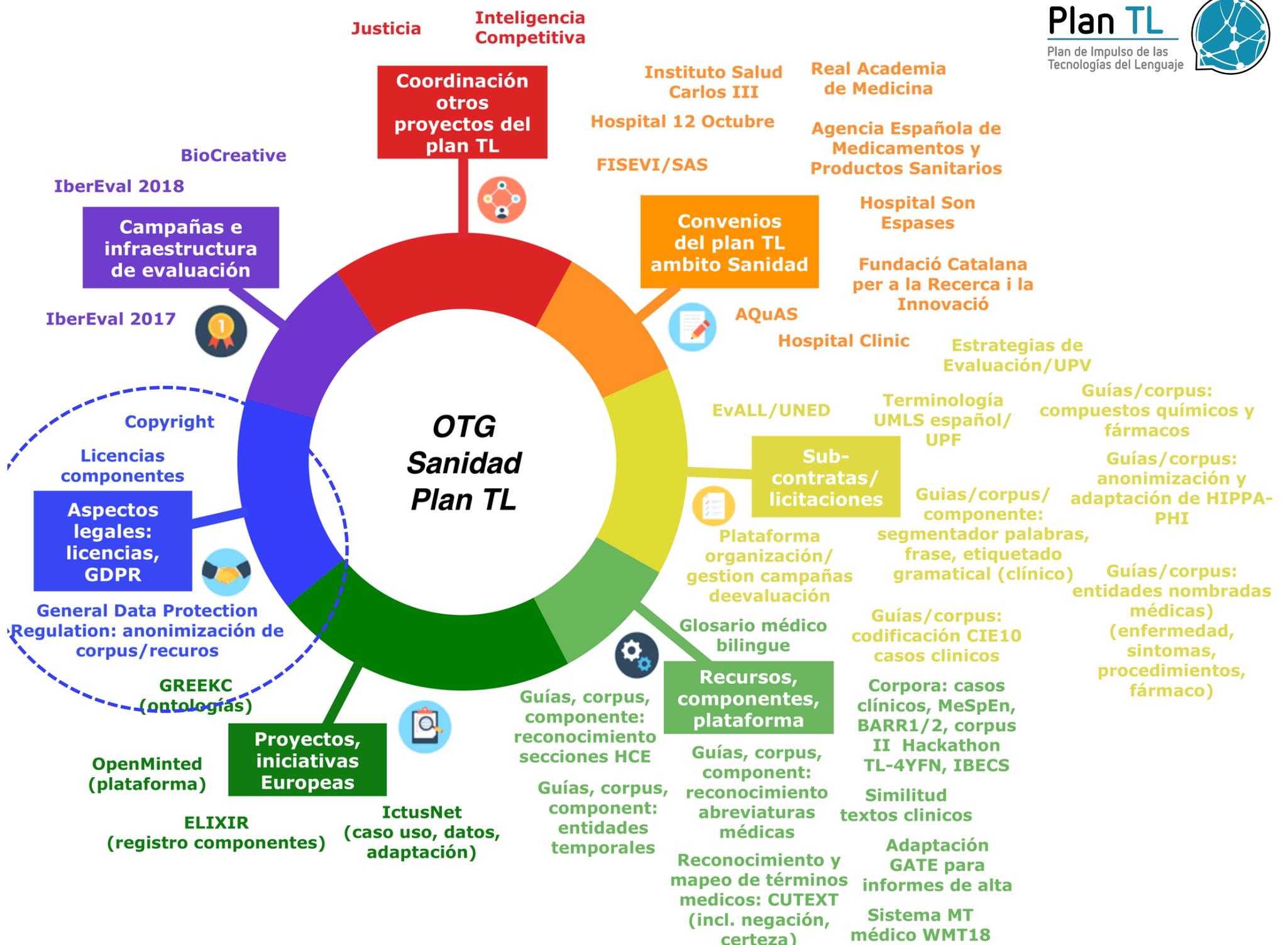
Desde el texto hasta
identificadores de conceptos
médicos



Módulos integrados

- Conversión a minúsculas.
- Conversión al lema.
- Separación de signos y números.
- Eliminación de tildes.
- Expansión de acrónimos.
- Sustitución de caracteres extraños.
- Sustitución de palabras por su sinónimo más común.
- Modificador: Negación
- Modificador: conjunciones
- Modificador certeza clínica
- Calculo de similitud textual





Corpus sintético de Casos Clínicos en Español : anonimización

Nombre: **Luis Miguel**

Apellidos: **Suárez Castro**

NASS: **15 50201565 22**

Domicilio: **Avda. de los Empresarios 15, 1F**

Localidad/ Provincia: **Cádiz**

CP: **11319**

NHC: **76194955**

Datos asistenciales

Fecha de nacimiento: **28/05/1992**

País: **España**

Edad: **25 años** Sexo: **Varón**

Fecha de Ingreso: **03/06/2018**

Especialidad: Urología

Médico: **Manuel Soto Delgado** N°Col: **11 11 26484**

Antecedentes: sin antecedentes personales de interés

Historia Actual: **Varón** de **25 años** de edad que consulta por tumoración en ambos testes de dos semanas de evolución.

Exploración física: Masa dura que ocupa más de la mitad del testículo izquierdo. Teste derecho aumentado de consistencia sobre todo en su parte posterior.

Resumen de pruebas complementarias: Eco escrotal : Tumoración sólida de 37x21 mm, hipoecoica en teste derecho. Dos lesiones sólidas en teste izquierdo de 32x17 mm y de 7 mm de características ecográficas similares a las encontradas en teste derecho. Todo ello compatible compatible con neoplasia testicular bilateral multicéntrica. Varicocele izquierdo.

Marcadores tumorales: AFP: normal; β -HCG: 6,15 mUI/ml (valor normal: < de 2,5 mUI/ml); LDH: normal.

Estudio de extensión (TAC toracoabdominopélvico): crecimientos adenopáticos a nivel retroperitoneal (N2) y lesiones metastásicas hepáticas (M1).

Tratamiento quirúrgico: Orquiectomía radical bilateral por vía inguinal. El paciente desechó la posibilidad de colocar prótesis.

Diagnóstico Principal: seminoma con elevado índice mitótico e invasión vascular (pT2) en ambos testes (seminoma anaplásico).

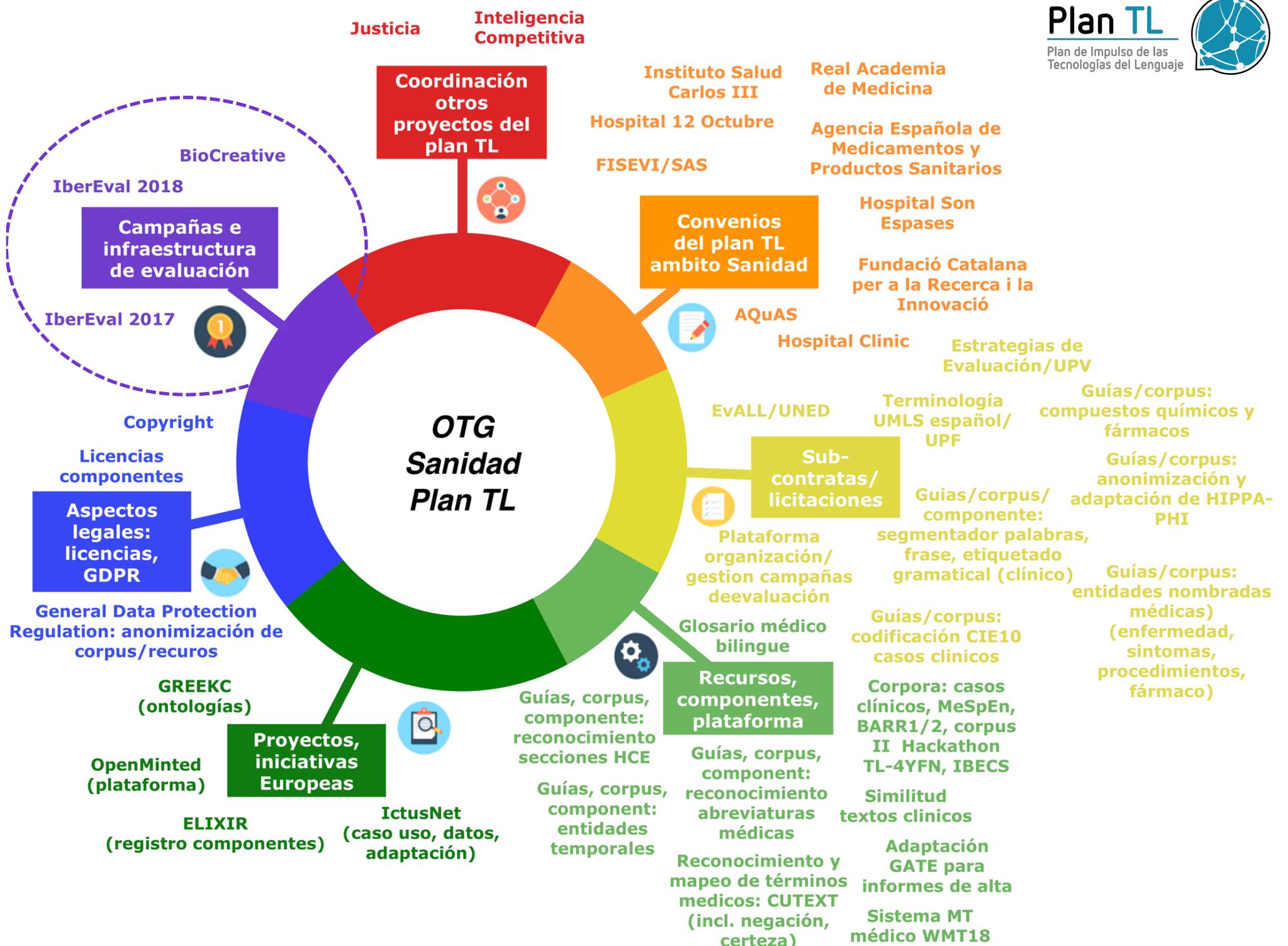
Tratamiento: quimioterapia sistémica con el esquema etopósido y cisplatino (4 ciclos).

Guías de anotación y corpus de Anonimización adaptando HIPPA-PHI al escenario español.

Estudio legal proceso y resultado de anonimización (GDPR)

Corpus de HCE de personas fallecidas

Estadísticas y n-gramas derivados de HCE





[Home](#) [Team participation](#) [Important dates](#) [Organization](#) [Datasets](#) [Evaluation](#) [Resources](#) [Proceedings](#) [FAQ](#) [BARR 2017](#)



Biomedical Abbreviation Recognition and Resolution 2nd Edition (BARR2)

IberEval 2018 | SEPLN 2018

18 September 2018. Seville, Spain

Description

Healthcare professionals do face serious time constraints when carrying out their daily workload. Moreover, a considerable fraction of medical and clinical terms, including names of diseases, symptoms or even clinical tests, correspond to long complex phrases. Consequently, abbreviations and acronyms, hereafter referred to as short forms, are widely encountered in clinical reports. This is true even for cases of highly ambiguous short forms that could potentially have many different senses.

Clinical short forms, as a result of their underlying semantic complexity, are problematic. Correct interpretation of short forms is not only difficult for patients, but often represent a challenge even for medical professionals and health documentation experts.

Contact

Ander Intxaurreondo
ander.intxaurreondo@bsc.es

News

- **June 21th, 2018:** BARR2 final test set clinical cases revealed.
- **May 28th, 2018:** BARR2 background and test sets released.
- **May 25th, 2018:** BARR2 development set released.
- **May 18th, 2018:** BARR2 evaluation script released.
- **May 17th, 2018:** BARR2 training set released.

- Infoday sobre Actuaciones del Plan de Impulso de las Tecnologías del Lenguaje en Sanidad
- Taller TASS-2018-Task 3. eHealth Knowledge Discovery
- Taller IberEval: DIANN: anotación de discapacidades
- Taller IberEval: Second Biomedical Abbreviation Recognition and Resolution track (BARR2)
- SEPLN: sesión específica sobre Análisis de textos médicos (3 trabajos)
- Charla Invitada 1: Wendy Chapman
“How Close are We to Using NLP Out of the Box”

Conclusiones y actuaciones futuras

- Facilitar acceso a los recursos, guías de anotación, corpus, terminologías y componentes generados para sanidad
- Promover la organización y participación en campañas de evaluación y de tareas de TL en sanidad
- Convenios/licitaciones: involucrar y colaborar con la industria de TL, grupos académicos de TL y fomentar interacción con informática de la salud/ informática médica y actores de sanidad
- Identificación y participación en propuestas europeas y nacionales relevantes para el plan TL en Sanidad
- Fomentar la formación de TL del dominio biomédico y de sanidad

Gracias

www.PlanTL.es

PlanTecnologiasLenguaje@minetad.es

