



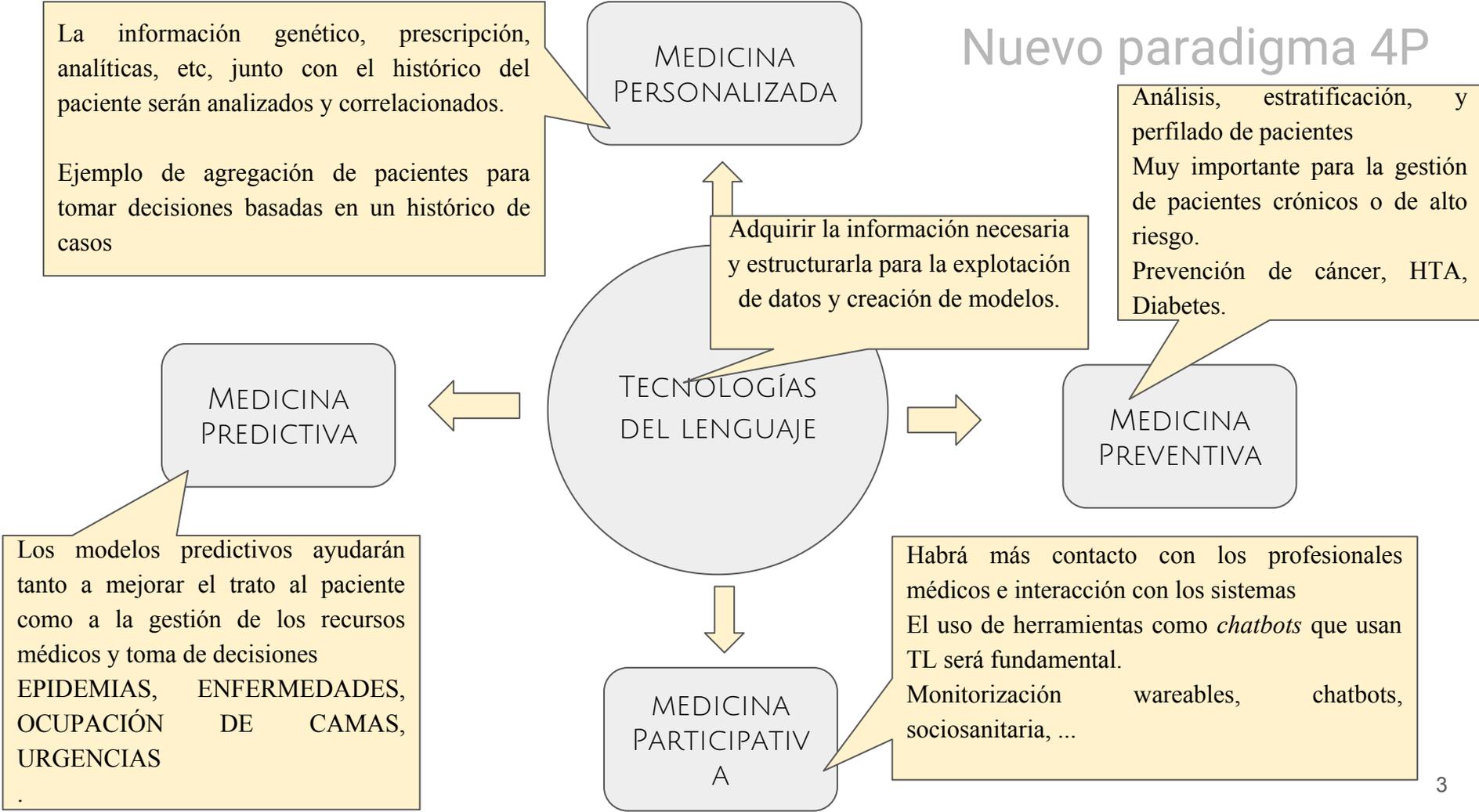
Tecnologías del lenguaje para la transformación digital

Enrique Mota Socio Director



“Sector de la Salud es la única gran industria que todavía no ha abrazado por completo la revolución digital”

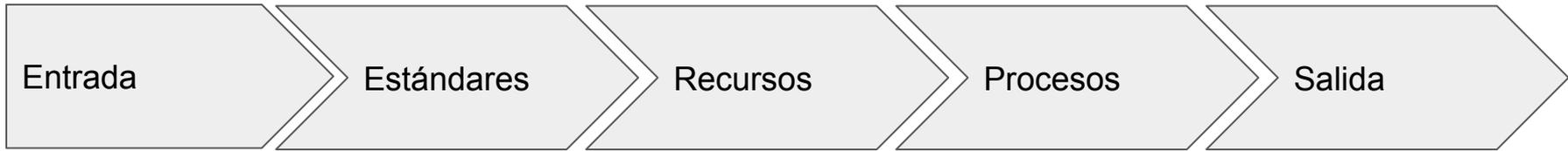
Nuevo paradigma 4P



Pipeline - Lógico

1. UMLS
2. SNOMED
3. CIE-10-ES
4. LOINC

- Extracción de entidades (Dx, Px, ...)
- Correlaciones
- Análisis
- Prección



- Limpieza y corrección
- Anonimización
- Identificación epígrafes, frases, entidades, ...
- Negaciones
- Duda
- Contextualización

Personalizados

- Diccionarios
- Tesoros
- Ontologías

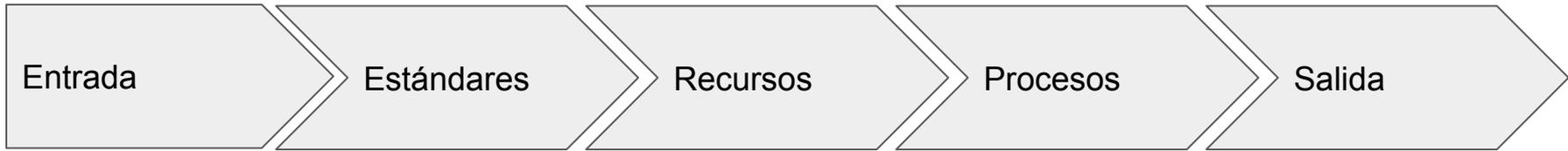
- Activadores de procesos
- Reporting
- Alertas y notificaciones

Pipeline - Herramientas

indizen 

- [CliniTerm](#)

- SpaCy
- NLTK
- Stanford NLP



- Data Lake
- NoSQL
- Data masking
- Data Cleansing
- ...

- Recursos personalizados mediante machine learning

- Mensajería HL7
- Conectores con Web Services
- ...

Casos empresariales

Indizen

- Codificación masiva y en origen
- Estratificación de pacientes (patologías, riesgos, ...)
- Predicción
- Data masking

Globales

- **Farr Institute:** Es una entidad inglesa que tiene como objetivo el empleo de la ciencia de datos en medicina, éste es un buen ejemplo para el ámbito público y además también hacen data mining de datos médicos.
- **Amazon:** acaba de apostar el sector sanitario, es interesante pensar lo que puede hacer una empresa en este sector (aplicar estrategias del retail a salud)