



Hacia un indicador temprano de Deterioro Cognitivo válido como herramienta de screening

Rafael Martínez-Tomás, Mariano Rincón-Zamorano,
Alba Gomez-Valadés
Dpto. Inteligencia Artificial, Universidad Nacional de
Educación a Distancia
UNED
Madrid, Spain

Héctor Gómez, Susana Arias
Facultad de Ciencias Humanas y Educación
Universidad Técnica de Ambato
Ambato, Ecuador

Abstract—Las enfermedades neurodegenerativas son causantes de deficiencias cognitivas que disminuyen gravísimamente la calidad de vidas de pacientes y familiares, se han convertido en uno de los problemas más graves de la sociedad actual, cada vez más envejecida. En esta tipología de enfermedades, la identificación temprana, en los primeros estadios, puede aliviar y ralentizar el desarrollo de las mismas. Si bien, para alguna de estas enfermedades (ej. Alzheimer), existen ya indicadores tempranos claros, es importante contar con un mecanismo barato, rápido, poco intrusivo y eficiente para la identificación temprana. En este artículo se resumen los principios y resultados de nuestra investigación con este objetivo. Hemos trabajado con test neuropsicológicos, de producción semántica y gráficos, obteniendo resultados prometedores. También hemos experimentado con la disociación entre la orientación del contenido semántico/emocional oral y la expresividad que le acompaña, en este caso tonalidad, en su articulación. Por último, describimos nuestro trabajo en pro de una mejor integración de resultados de experimentos en estas áreas, que proceden de diferentes fuentes y en diferentes formatos, gracias al diseño de un marco ontológico que sin duda colaborará en una mayor reutilización de estos datos para análisis más completos y profundos.

Palabras clave—MCI, neuropsychological test, Alzheimer disease, machine learning, computer vision, bayesian networks.

I. INTRODUCCIÓN

Se estima que para el año 2020, 48,1 millones de personas en todo el mundo sufrirán problemas relacionados con el síndrome clínico que denominamos demencia [1] y se triplicará para 2050¹. Quizá por el alargamiento de la esperanza de vida, la enfermedad de Alzheimer (AD), que representa ya la primera causa de demencia neurodegenerativa [2], es el problema más preocupante y analizado de los últimos años relacionado con este síndrome. Es una patología degenerativa causante de deterioro cognitivo y de problemas de comportamiento y funcionales, que no solo produce un alto impacto en la calidad de vida del paciente sino también de sus familiares y cuidadores. AD se está atacando desde múltiples perspectivas, por un lado, buscando terapias para retardar la

evolución neurodegenerativa, y por otro, buscando su detección precoz. Desde el diagnóstico temprano se puede prestar un mayor apoyo al paciente, monitorizar la enfermedad y administrar los tratamientos que permitan mantener su calidad de vida durante más tiempo. Además, como en todos los procesos neurodegenerativos, en un ciclo virtuoso, cuanto antes se diagnostique, mejor se puede estudiar la evolución y consecuentemente conocer mejor la enfermedad.

El mejor y más temprano indicador de la enfermedad de Alzheimer es el índice de concentración de las proteínas tau y beta-amiloide en el fluido cerebroespinal, pero requiere de un método invasivo, la punción lumbar, muchas veces contraindicado. En general, el diagnóstico del deterioro cognitivo puede ser costoso temporal y económicamente, pues requiere la evaluación de información procedente de diferentes fuentes (neuropsicológica, test de laboratorio, neuroimagen, datos históricos, demográficos, personales, etc.) y cuya precisión y eficiencia vienen determinadas por el nivel de pericia del personal profesional. Por su lado, la evaluación neuropsicológica se ha revelado como una herramienta esencial para la detección temprana de deterioro cognitivo, permitiendo identificar déficits funcionales consecuencia de alteraciones en el entramado neuronal producido por enfermedades neurodegenerativas. El DSM-5, la 5ª edición del Diagnostic and Statistical Manual of Mental Disorders del 2013 de la American Psychiatric Association, define seis dominios de la función cognitiva: atención compleja, función ejecutiva, aprendizaje y memoria, lenguaje, función visoespacial y cognición social, cada uno con un determinado número de subdominios.

Nuestro grupo ha utilizado técnicas y metodologías propias de la Inteligencia Artificial (IA) para el análisis de estas evaluaciones neuropsicológicas con el objetivo de obtener herramientas para el screening de población, es decir, herramientas para la identificación temprana de la deficiencia cognitiva, poco invasivas, económicas y sin riesgos secundarios. En concreto, hemos utilizado modelado ontológico para estructurar el conocimiento disponible y poderlo utilizar en inferencia, redes bayesianas para combinar conocimiento de expertos con aprendizaje a partir de datos, y diferentes métodos de aprendizaje automático cuando el volumen de datos lo ha permitido.

¹ <https://www.ceafa.es/es/que-comunicamos/noticias/la-oms-advierte-que-para-el-2050-se-triplicara-la-cifra-de-personas-que-padecen-demencia>

Para conseguir este fin, hemos formado un equipo multidisciplinar en el que participan investigadores del Dpto. de Inteligencia Artificial y del Dpto. de Psicología Básica de la UNED, el Centro de Referencia de la Enfermedad de Alzheimer (CREA) de Salamanca y el Dpto. de Geriátrica del Hospital de Zamora. Asimismo, colaboramos con el grupo dirigido por Dr. Héctor Gómez de la Universidad de Ambato, la residencia del perpetuo socorro de Quito y otros centros asistenciales de Ecuador, y asimismo varios hospitales universitarios noruegos.

En este artículo hacemos un repaso por nuestros trabajos más relevantes en esta línea, y describimos las técnicas empleadas y los resultados obtenidos. En el apartado 2 se describe nuestro trabajo en dos tipos de test neuropsicológicos. Primeramente, mostramos nuestro trabajo en el análisis de la producción oral de categorías semánticas mediante el uso de redes bayesianas (BNs), lo que permite combinar conocimiento experto con el corpus de datos disponible. A continuación, nos centramos en otro tipo de test, también muy habitual en la evaluación cognitiva, que se fundamenta en la copia a mano alzada de figuras geométricas esquemáticas. Aquí, para el análisis, se hace uso de técnicas de visión artificial y aprendizaje automático. En el apartado 3, describimos nuestro trabajo para la identificación de déficits expresivos que puedan asociarse con deterioros cognitivos, para lo cual se analiza la orientación semántica/emocional de la conversación y la correspondiente entonación, buscando inconsistencias en las asociaciones mediante diferentes técnicas de aprendizaje automático. Ya en el apartado 4, se presenta otra línea de trabajo, en la cual utilizamos modelado ontológico para crear herramientas de apoyo a la investigación en consonancia con la tendencia actual de los proyectos de investigación en neurociencia -proyectos multidisciplinarios, heterogéneos y multicentro-, dónde la compartición de datos y la accesibilidad son cruciales. Por último, en el apartado de conclusiones, resumimos los logros obtenidos, los problemas encontrados y las limitaciones de nuestras implementaciones, extrayendo conclusiones generales y futuras líneas de investigación.

II. APOYO A LA IDENTIFICACIÓN TEMPRANA DEL DETERIORO COGNITIVO DESDE TEST NEUROPSICOLÓGICOS

Los test neuropsicológicos (Neuro-Psychological tests, NPT) se han mostrado muy útiles en la evaluación de los diferentes dominios y subdominios de la función cognitiva mediante la realización de tareas estandarizadas [4], pero es un tema de debate si unas tareas determinadas son las mejores [5], y, por tanto, qué test o combinación de ellos, y qué medidas para su evaluación, permiten clasificar tipologías y determinar etapas del desarrollo del deterioro con más precisión. El debate, parece claro que incluye una dosis de subjetividad [6] [7] y también de propiedad intelectual (test con propiedad intelectual). También existe una inercia a seguir utilizando los mismos test para facilitar la comparativa con otros estudios y para mantener la metodología en estudios longitudinales. En todo caso, hay un consenso generalizado en que los NPTs son una herramienta que, con unos requerimientos muy simples (un papel y un lápiz en algunos casos), es fundamental para determinar claves de la organización de la actividad cerebral y su implicación en los trastornos cognitivos. La lista de test

podría optimizarse o acortarse, con el objetivo de reducir costes económicos y de tiempo, de cara a su aplicación en campañas de screening en amplios sectores de la población.

A. Análisis de la producción oral con Redes Bayesianas

En [8,9], asumimos que un método de diagnóstico temprano del AD, basado en el análisis de la producción oral para identificar el deterioro de la memoria semántica, era un método que cumplía con nuestros objetivos: es un método simple, barato, y accesible a la población general y sirve como un primer filtro para test posteriores más caros y precisos. Además, nuestra hipótesis era que las BNs eran la herramienta apropiada para alcanzar este objetivo, ya que combinan el conocimiento a priori de los expertos (modelo cualitativo de la BN, su estructura) con el conocimiento aprendido de los conjuntos de datos de casos resueltos (modelo cuantitativo de la BN, probabilidades a priori).

En esta investigación, hicimos uso de dos fuentes de información para el diseño cuantitativo de la BN. Por un lado, el corpus de definiciones orales de Peraita y Grasso [10], que proporciona información sobre las diferencias en la producción semántica entre las personas sanas y los pacientes con CI (diagnosticados como AD leve y moderado). Peraita y Grasso analizaron las características semánticas de las definiciones del corpus, clasificando las definiciones de los pacientes dentro de 11 bloques conceptuales básicos: taxonómicos, tipos, partes funcionales, evaluativos, lugar/habitat, comportamiento, causas/generaciones procedimientos, ciclo de vida y otros. El test consistía en describir seis conceptos, pertenecientes a dos categorías semánticas diferenciales, criaturas vivas y artefactos, contar el número de elementos producidos por cada individuo para cada bloque conceptual y usar este número para obtener probabilidades a priori. Los autores del corpus detallaron los resultados de los casos recopilados (81 personas procedentes de diferentes hospitales de Madrid, 42 sanas y 39 diagnosticadas con AD leve o moderado) y analizaron cómo ciertas categorías semánticas eran representadas mentalmente, empleando modelos teóricos de características semánticas obtenidas de las tareas lingüísticas explícitas.

Por otro lado, el estudio epidemiológico de Fernández et al. [11] relacionó edad, sexo y nivel educativo con AD, lo que nos permitió relacionar la producción lingüística con el deterioro cognitivo (CI) y los estudios epidemiológicos con el AD y, debido a que el CI es un indicador temprano del AD, relacionar el CI con AD. Así, combinando esas dos fuentes pudimos cuantificar la influencia de las variables del modelo.

Se desarrolló una aplicación software en Java para gestionar todo el proceso de adquisición de datos para la BN. La inferencia bayesiana fue implementada a través de los módulos de inferencia desarrollados en Elvira². Entrenamos nuestra red BN con nuestro modelo de aprendizaje propio, empleando como *gold estándar* el diagnóstico llevado a cabo por los neurólogos, quienes seguían los Criterios NINCDS-ADRDA del Alzheimer. Dado el pequeño número de casos disponibles en el corpus, se empleó validación cruzada leave-one-out para evaluar el modelo. La curva ROC obtenida se

² Actualmente denominado Open Markov: <http://www.openmarkov.org>



podía usar para estimar el punto de corte óptimo de acuerdo al compromiso entre verdaderos positivos y falsos positivos. Dado que el objetivo era el diagnóstico temprano del AD, lo ideal sería que el sistema detectara todos los casos de AD, aunque el número de personas sanas clasificadas como AD fuera alto. De esta forma, el umbral óptimo en la distribución de probabilidad posterior continua sería aquella que maximizara el número de verdaderos positivos y mantuviera el número de falsos positivos lo suficientemente bajo; sin embargo, para obtener el umbral óptimo, hubiera sido necesario un análisis de coste que quedó fuera del trabajo. Por ello, consideramos nuestra BN como otro método para apoyar el diagnóstico del AD junto a otros test complementarios, y seleccionamos el punto de corte óptimo como el punto de la curva ROC con la máxima precisión. La Tabla 1 muestra las métricas de rendimiento para este umbral.

TABLE I. MÉTRICAS DE RENDIMIENTO OBTENIDAS CON LA BN PROPUESTA

Performance Metrics	Values
TP rate, TP/(TP + FN)	0.8718
FP rate, FP/(FP + TN)	0.0476
Precision, TP/(TP + FP)	0.9444
Accuracy, (TP + TN)/(P + N)	0.9136
AUC, area under ROC curve	0.9621

El estudio concluyó con la confirmación de las ventajas del modelo BN y, por lo tanto, consideramos que son los modelos más adecuados para los sistemas de diagnóstico y para el estudio de la enfermedad a partir del análisis de la producción oral de las características semánticas. Concluimos que:

- El modelo de BN, su estructura y sus parámetros, es reconocible por el experto.
- El modelo cuantitativo puede ser ajustado a partir de una base de datos de casos y estudios, permitiendo, por tanto, la recopilación tanto el conocimiento del experto como del conocimiento implícito en los datos disponibles.
- El modelo de BN, junto al algoritmo de aprendizaje automático, puede ser empleado con otro conjunto de categorías semánticas, tan amplio como sea posible o deseado (lógicamente, el corpus debe poseer los casos suficientes).

Además, este método podría ser extendido con nuevas herramientas para proporcionar una gran cantidad de conocimiento sobre el deterioro de la memoria semántica irregular y contribuir al diagnóstico temprano de una forma extremadamente económica y accesible:

- Se podrían incorporar nuevas variables a nuestro modelo de BN, por ejemplo, variables sociodemográficas y clínicas asociadas, donde la cooperación entre equipos multidisciplinarios podría ser esencial.
- Nuestro sistema podría ser ampliado con un mecanismo de explicación, para lo cual sería especialmente útil medir la sensibilidad del modelo a los cambios en las evidencias.
- Nuestro método puede ser ampliado con análisis de decisiones (diagramas de influencia) para maximizar la utilidad esperada de varias opciones abiertas en la decisión.

Por ejemplo, podría maximizarse la utilidad esperada de las recomendaciones para hacer un test neuropsicológico concreto o exploraciones complementarias como los análisis bioquímicos, PET, etc.

- Disponer de más casos permitiría un análisis más profundo y seleccionar las variables más discriminantes para reducir el tiempo de los test.

B. Estudio de los test neuropsicológicos gráficos empleando visión artificial y aprendizaje automático.

Una parte de las pruebas neuropsicológicas estandarizadas diseñadas para evaluar el deterioro cognitivo leve (MCI) incluyen la reproducción o copia de figuras geométricas (test neuropsicológicos gráficos, G-NPT). La idea principal de este método es que ciertas distorsiones respecto a los patrones pueden indicar diferentes grados y perfiles de MCI. Además, no solo se puede evaluar la distorsión respecto al patrón, sino también la evolución en el proceso de ejecución a lo largo de los años. Existen figuras de este tipo en distintos test, como el Mini Examen Cognoscitivo (MEC) [12], el Test Barcelona [13] o la figura compleja de Rey [14].

Los procesos y funciones cognitivas —dominios cognitivos— que supuestamente se evalúan durante la ejecución de esas figuras son la función ejecutiva, la percepción visual y/o visoespacial, las habilidades motrices y la memoria especial. En cada uno de esos procesos o funciones, los rasgos o componentes (control, inhibición, planificación, etc.) involucrados en los patrones a analizar pueden ser definidos, por ejemplo, siguiendo las indicaciones del Dr. Peña Casanova con respecto a los dibujos de gráficos alternantes y bucles.

Mientras que en otros tipos de pruebas estandarizadas que evalúan la memoria episódica, la fluencia verbal, etc. es mucho más fácil obtener datos normativos para puntuar las pruebas sin subjetividad, en estos test que involucran la reproducción y copia de figuras, es mucho más difícil, imponiéndose la subjetividad y cierta discrecionalidad por parte del evaluador. Aunque existe un criterio de evaluación, existe también un gran componente de subjetividad que puede socavar la fiabilidad si las discrepancias entre evaluadores no son corregidas [15]. Como resultado, puede haber un importante margen de error en la detección de ciertos problemas de habilidades motrices, percepción visual o visoespacial, etc. dentro de un entorno para la detección temprana de MCI.

En [16] utilizamos técnicas de visión artificial y aprendizaje automático para automatizar este análisis. Aunque el objetivo a largo plazo era trabajar con todas las figuras de las pruebas mencionadas anteriormente, aquí nos centramos en la figura “pico-meseta”, que es una de las subpruebas del Test Barcelona. El valor diagnóstico de esta prueba, seleccionada para este análisis exploratorio automático, se basa en el hecho de que en su ejecución están involucrados algunos componentes de la función ejecutiva, como la serialización, la planificación, la flexibilidad, la inhibición, así como la capacidad práxica. La tarea consiste en copiar una figura en la que se alternan picos y mesetas. Las distorsiones que pueden emerger en la ejecución de esta figura pueden ser de diferentes tipos: variaciones en el tamaño de la figura realizada, alteración

de las características por la adición o eliminación, garabateo, perseveraciones, rotaciones, acercamiento al patrón, etc. Esos errores o alteraciones en la reproducción del dibujo pueden ser marcadores de disfunciones más severas, que pueden ser de tipo apráxico, en los cuales las funciones ejecutivas también están involucradas, como en la demencia tipo Alzheimer [17].

Dependiendo de los resultados obtenidos en varios NPTs, los participantes fueron clasificados de acuerdo a dos perfiles cognitivos, individuos sanos (n=16), aquellos con puntuaciones normales dentro de las escalas de referencia, e individuos con MCI (n=24). Se desarrolló un algoritmo de visión artificial para 1) segmentar el dibujo, 2) extraer las líneas del dibujo, 3) reconocer los componentes del patrón (picos y mesetas), 4) caracterizar el dibujo y las discrepancias dibujo-patrón (15 características relacionadas con la posición respecto al patrón, diferencia de tamaño, inclinación del dibujo, fallos en el patrón pico-meseta, etc.) y 5) finalmente, establecer un diagnóstico mediante un clasificador implementado con un árbol de decisión J48. El clasificador obtuvo una precisión de 0.77 y una exhaustividad de 0.77.

El problema fundamental del diagnóstico del AD a partir del aprendizaje automático, al igual que en la mayoría de los estudios neurológicos, es la ausencia de un conjunto de datos lo suficientemente grande para construir un sistema fiable de diagnóstico. Tenemos que reconocer que, con esta pequeña muestra, solo pudimos concluir que estas características propuestas para el diagnóstico del MCI eran buenas candidatas, pero que por sí solas no permitían distinguir diferentes tipos de MCI. Para trabajos futuros, estamos ampliando la muestra y queremos combinar características de diferentes figuras para mejorar el rendimiento de la clasificación. De esta forma, este trabajo intenta ser un estudio piloto de un trabajo mucho más amplio, en el cual, a partir de los datos longitudinales ya disponibles, podremos realizar diferentes tipos de análisis:

- Análisis longitudinales de controles sanos: comparar, a lo largo de una serie de años, la ejecución de los dibujos de los sujetos de control sanos, para verificar su estabilidad en la realización de las figuras.

- Análisis longitudinales de los MCIs estables: el mismo tipo de aproximación, pero en otro grupo de sujetos, para intentar responder a las siguientes preguntas: ¿Es la ejecución de las figuras estable? ¿De todas ellas? ¿Por cuánto tiempo?

- Análisis longitudinales de los MCIs que evolucionan bien a AD bien otra demencia.

- Análisis transversales de diferentes tipos de MCI: estudios de grupo con MCI amnésico, multidominio y no amnésico.

- Relación entre la ejecución de la figura y la apraxia ideomotora: para estudiar la relación entre la ejecución de las figuras y la apraxia ideomotora obtenida por los mismos sujetos en otros test.

- Análisis sociodemográficos: para analizar las relaciones entre las variables sociodemográficas, tales como la edad, el género o el nivel educativo, y la ejecución de los patrones de las figuras mencionadas anteriormente.

Pero, por otro lado, los G-NPTs sufren de la ambigüedad inherente a la valoración por diferentes especialistas, tanto en las características significativas de los dibujos como de la clasificación del paciente según el resultado de la misma. Para tratar este problema parece adecuado la aplicación de técnicas de IA, para tratar con las ambigüedades de forma más natural, coherente con la interpretación humana, evitando tomar decisiones "crisp" (según la teoría de conjuntos clásica), por ejemplo, con técnicas difusas o probabilistas.

Sobre determinados G-NPTs, el uso de técnicas de ML ha demostrado altos niveles de precisión, pero se ha trabajado sobre cada test de forma independiente por la incertidumbre en la asociación dibujo-patrón. Hay muchos tipos y es necesaria una asociación de los G-NPTs con las capacidades cognitivas y motoras, y de estas con regiones y dinámica cerebral. Se echa de menos la definición de ontologías que los modelen, organicen y enlacen.

III. ANÁLISIS DEL DÉFICIT DE EXPRESIVIDAD USANDO ANÁLISIS DE SENTIMIENTO

En [18] planteamos la hipótesis de que el contenido semántico/emocional del discurso y las manifestaciones físicas asociadas (gestos, entonación...) se alinean en personas cognitivas normales, pero no en aquellas personas afectadas por la demencia. Por ello, planteamos este elemento para identificar pacientes con problemas cognitivos en la etapa inicial.

El estudio de cambios de comportamiento, analizado con resultados favorables a partir de texto escrito en personas mayores [19], podría ayudar a prevenir enfermedades ligadas a dichos cambios. El análisis de la voz se ha usado para indicar signos de la enfermedad en ancianos [20]. Nosotros tomamos en cuenta la semántica de una conversación y la emoción en la tonalidad de la voz conjuntamente, en vez de por separado, para verificar si el contenido semántico-emocional del texto de un entrevistado correlaciona con el tono de voz. Si no están relacionados, podríamos suponer que se debería avisar a los familiares o cuidadores acerca de los cambios en el comportamiento del paciente.

Para el análisis, seleccionamos dos características asociadas con la semántica de un texto: la polaridad y la orientación semántica. La polaridad (positiva, negativa o neutra) está basada en diccionarios especializados y usamos la herramienta *SentiStrength*³ para medirla. La orientación semántica es una medida propuesta por [26] que ayuda a clasificar frases como "Excelente" o "Pobre". La orientación semántica funciona con las opiniones que las personas tienen acerca de las palabras o frases en un contexto predeterminado. Se busca esta orientación en frases similares mediante Google⁴ Api.

Las variaciones de la entonación que una persona le da a una frase mientras habla se asocia con variaciones emocionales. Para ello se usó la herramienta *Speech Recognition System*⁵. De aquí, la emoción de la voz se puede

³ <http://sentistrength.wlv.ac.uk/>

⁴ <http://ajax.googleapis.com/ajax/services/search/web?v=1.0&q=%s>

⁵ <https://www.linkedin.com/pulse/speech-emotion-recognition-system-matlab-source-code-luigi-rosa>



comparar con la polaridad y con la orientación ya que se clasifican dentro de bandas positivas -excelentes y negativas-pobres. Nos ha interesado especialmente el análisis e interpretación de los datos, más que la investigación en las técnicas de extracción de características.

Las entrevistas (texto+audio) del grupo de control, formado por pacientes ancianos sanos, se obtuvo de Charlotte⁶. Los resultados mostraron que no había una gran variabilidad entre los valores etiquetados por los expertos y las herramientas empleadas. No se obtuvo disociación entre polaridad, orientación semántica y entonación. Después se probó con pacientes con AD en una etapa temprana y sí se obtuvo disociación relevante (p -valor $<0,05$). Los valores obtenidos de polaridad, orientación semántica y entonación utilizaron para construir varios clasificadores que distinguieran entre personas sanas y AD. Los clasificadores se probaron con variables individuales y combinadas, proporcionando mejores resultados para el F1Score cuando se trabajó con las 3 variables simultáneamente. Los resultados demuestran que es posible utilizar el método para el diagnóstico, y que las variables trabajan mejor juntas que cuando son usadas por separado en los clasificadores.

TABLE II. COMBINING SENTI-STRENGTH-PMI-TONALITY

Algorithm	Precision	Recall	F1Score
J48	0,88	0,78	0,83
Multilayer perceptron	0,75	0,86	0,80
Bayes Net	0,79	0,91	0,85

IV. UN MARCO SEMANTICO INCREMENTAL PARA MEJORAR LA ACCESIBILIDAD EN SISTEMAS BIOBANKING

Como ha quedado patente a lo largo del artículo, uno de los problemas con el que nos encontramos es la falta de una muestra significativa para los estudios. Por ello, es fundamental poder integrar *datasets* de distintos estudios, con muy diferentes características y contextos. El intercambio y accesibilidad de los datos permite unir los esfuerzos entre diferentes grupos de investigación, así como en los estudios de replicación, cruciales para el progreso en el campo. Las soluciones de los archivos de datos de investigación están evolucionando rápidamente para abordar esas necesidades, sin embargo, la integración de los datos distribuidos es todavía complicada debido a la necesidad de acuerdos entre los diferentes modelos de datos. Para abordar estos problemas, las ontologías son ampliamente empleadas en investigación biomédica para obtener vocabularios comunes y descripciones lógicas, pero su aplicación podría sufrir de problemas de escalabilidad, sesgos de dominio, y pérdida de acceso de los datos de bajo nivel. Con el objetivo de mejorar los modelos de aplicación semántica en los sistemas de *biobanking*, hemos diseñado e implementado un marco semántico incremental que aprovecha los últimos avances en ontologías biomédicas y en la plataforma XNAT [21]. Se trata de una aproximación metodológica para abordar el problema de habilitar modelos basados en la semántica en sistemas de archivos de investigación ya implementados. Esto mejora la gestión de los

⁶ <http://newsouthvoices.uncc.edu/nsv/southernpiemont>

datos, desde datos de bajo nivel a conceptos semánticos y lógicos. Construido con tecnologías de la Web Semántica y empleando ontologías biomédicas, el marco proporciona un modelo homogéneo de acceso a datos y razonamiento sobre datos neurológicos multimodales.

El diseño del marco sigue una aproximación de abajo a arriba por capas, permitiendo trabajar con los datos a diferentes niveles de descripción. El marco añade capacidades de razonamiento a partir de las relaciones implícitas y las definiciones lógicas para derivar nuevos datos, y realiza comprobaciones de coherencia de los datos para el Control de Calidad. El empleo de los principios de enlazado de datos permite el enlace entre datos, abriendo las puertas a conjuntos de datos externos de referencia. Además, el tener un conjunto de datos altamente enlazados facilita la inspección desde diferentes conceptualizaciones (proyecto, tema, enfermedad, etc.), una característica altamente deseable para el descubrimiento de patrones y el estudio de las relaciones entre enfermedades a medida que el conjunto de datos crece.

Nuestra propuesta se diferencia de trabajos anteriores en que se centra en las consultas y el razonamiento avanzados sin perder los datos de bajo nivel, mientras aprovecha las ventajas de las plataformas de archivos ya disponibles y ampliamente utilizadas. En particular, elegimos XNAT como la columna vertebral para el manejo de los datos y las imágenes clínicas, debido a su amplio conjunto de características y su diseño flexible y personalizable.

Este marco está siendo utilizado en el JPND (EU Joint Program for Neurodegenerative Disease)⁷/APGeM project⁸, destinado a encontrar biomarcadores tempranos para la enfermedad de Alzheimer y otras demencias relacionadas [22] Comprende una cantidad significativa de datos de diferentes subdominios y modalidades, como neuroimágenes, bioquímica, rastreos clínicos/ neuropsicológicos y genéticos, estableciendo un escenario adecuado para impulsar y probar el marco en el curso de la investigación neurológica actual.

V. CONCLUSIONES

El problema de identificar etapas tempranas de enfermedades neurodegenerativas se puede atacar desde múltiples perspectivas. Aunque hemos experimentado con el déficit de la producción gestual en personas con demencia, este factor, medido en nuestro trabajo mediante la disociación entre gesto, tonalidad, y contenido semántico-emocional del discurso, no nos sirve como indicador neuropsicológico temprano de la enfermedad porque características subjetivas, como la motivación, la depresión, el estrés o simplemente el estado de ánimo, pueden hacer que personas sanas presenten signos de disociación semejantes.

Por eso nos estamos centrando más en identificar el deterioro cognitivo mediante test neuropsicológicos, que han demostrado ser una solución rápida y económica en la que vale la pena emplear recursos. Hemos investigado en la automatización y evaluación automática de diferentes

⁷ <http://www.neurodegenerationresearch.eu/>

⁸ <http://www.neurodegenerationresearch.eu/publication/apgem/>

mecanismos, el test de producción oral y test gráficos. Hay otros test, como el test de fluidez verbal, que aportan conclusiones diferentes y que tendrán que ser analizados en detalle y, seguramente, automatizados en próximos trabajos.

En cualquier caso, cada día se incrementa de manera no lineal la cantidad de datos de historias clínicas almacenadas, por lo que resulta natural que los nuevos trabajos de investigación traten de concentrar toda esta información. La manera de investigar está cambiando desde una investigación individualista, en la que es necesario proponer un modelo que explique o que sea consistente los pocos datos disponibles conocidos, hacia un modelo cooperativo, en el que los estudios multicentro, multidisciplinares, multidimensionales y heterogéneos, serán lo habitual, y el modelado del conocimiento y el aprendizaje automático a partir de datos cobrará cada vez más importancia. Para ello, como muestra el último trabajo presentado, es fundamental un marco ontológico que oriente la integración de diferentes fuentes, formatos de datos y resultados, para un mejor aprovechamiento en experimentos más completos y profundos, que ayuden a conocer mejor la enfermedad y su afectación sobre funciones cognitivas y áreas cerebrales.

RECONOCIMIENTOS

- Proyecto 018-ABEL-CM-2013, financiado por el programa NILS Science and Sustainability con el The Intervention Centre del Oslo University Hospital
- Ayuda predoctoral PEJD-2017-PRE/TIC-4406 cofinanciada por la UNED y el Fondo Social Europeo a través del Programa Operativo de Empleo Juvenil y la Iniciativa de Empleo Juvenil.

REFERENCES

- [1] M. Prince, R. Bryce, E. Albanese, A. Wimo, W. Ribeiro and C.P. Ferry. "The global prevalence of dementia: a systematic review and metaanalysis". *Alzheimer's & Dementia*, vol. 9(1), pp.63-75, 2013.
- [2] H.W. Querfurth and F.M. La Ferla. "Alzheimer's disease". *The New England Journal Medicine*, vol. 362, no. 4, pp. 329-344, 2010.
- [3] S. G. Gauthier. "Alzheimer's disease: the benefits of early treatment". *European Journal of Neurology*, vol. 12(s3), pp.11-16, 2005.
- [4] M.D. Lezak. *Neuropsychological assessment*. Oxford University Press, USA, 2004.
- [5] G. Gainotti, D. Quaranta, M.G. Vita and C. Marra. "Neuropsychological predictors of conversion from mild cognitive impairment to Alzheimer's disease". *Journal of Alzheimer's Disease*, vol. 38(3), pp. 481-495, 2014.
- [6] H.C. Fichman, R.M. Oliveria and C.S. Fernandez. "Neuropsychological and neurobiological markers of the preclinical stage of Alzheimer's disease". *Psychology & Neuroscience*, vol. 4(2), pp. 245-253, 2011.
- [7] R.H. Logie, M.A. Parra and S. Della Salla. "From cognitive science to dementia assessment". *Policy Insights from the Behavioral and Brain Sciences*, vol. 2(1), pp. 1-91, 2015.
- [8] J.M. Guerrero, R. Martínez-Tomas, M. Rincon and H. Peraita. "Diagnosis of Cognitive Impairment Compatible with Early Diagnosis of Alzheimer's Disease". *Methods of information in medicine*, vol. 55(1), pp. 42-49, 2016.
- [9] J.M. Guerrero, R. Martínez-Tomas and H. Peraita. "Bayesian Network Based Model for the Diagnosis of Deterioration of Semantic Content Compatible with Alzheimer's Disease". *Foundations on Natural and Artificial Computation. Lecture Notes on Computer Science*, vol 6686, pp.461-470, Springer, 2011
- [10] H. Peraita and L. Grasso. *Corpus lingüístico de definiciones de categorías semánticas de personas mayores sanas y con la enfermedad de Alzheimer: una investigación transcultural hispano-argentina*. Fundación BBVA, 2010.
- [11] M. Fernández-Martínez, J. Castro-Flores, S. Pérez-delas-Heras, A. Mandaluniz-Lekumberri, M. Gordejuela and J. Zarranz. "Prevalencia de la demencia en mayores de 65 años en una comarca del País Vasco". *Revista de Neurología*, vol. 46 (2), pp. 89-96, 2008.
- [12] A. Lobo, J. Ezquerro, J., F. Gómez, J.M. Sala, A. Seva. "El mini-examen cognoscitivo. Un test sencillo, práctico, para detectar alteraciones intelectivas en pacientes médicos". *Actas Luso Españolas de Neurología, Psiquiatría y Ciencias Afines* vol. 3, pp. 189-202, 1979.
- [13] J. Peña-Casanova. *Programa integrado de exploración neuropsicológica "test Barcelona"*. In *Normalidad, semiología y patología neuropsicológica*. Masso, Barcelona, 1991.
- [14] A. Rey. *Rey. Test de copia y de reproducción de memoria de figuras geométricas complejas*. TEA, Madrid, 2003)
- [15] S. Urbina. *Claves para la evaluación con tests psicológicos (Kaufman, A.S., Kaufman, N.L (trad.))*. TEA Ediciones (publicado originalmente en 2004), Madrid, 2007.
- [16] M. Rincón, Sara García-Herraz, M.C. Díaz-Mardomingo, R. Martínez-Tomás & H. Peraita. *Automatic drawing analysis of figures included in neuropsychological tests for the assessment and diagnosis of mild cognitive impairment*. In *International Work-Conference on the Interplay Between Natural and Artificial Computation*, pp. 508- 515. Springer, Cham. 2015.
- [17] R.Q. Freeman. T. Giovannetti, M. Lamar, B.S. Cloud, R.A. Stern, E. Kaplan, D.J. Libon. "Visuoconstructional problems in dementia: contribution of executive systems functions". *Neuropsychology* vol. 14(3), pp. 414-426, 2000.
- [18] S. Arias, R. Martínez-Tomas, H. Gómez, V. Hernandez, J. Sanchez, J. Barbosa, J. Mocha. *The dissociation between polarity, semantic orientation, and emotional tone as an early indicator of cognitive impairment*. *Frontiers in Computational Neurosciences*, vol. 10 (95), 2016.
- [19] H. Cole-Lewis, T. Kershaw. "Text Messaging as a Tool for Behaviour Change in Disease Prevention and Management." *Epidemiology*, nº 10.1093, 2010.
- [20] L. Naranjo, C. Pérez, Y. Campos-Roca and J. Martín. "Addressing Voice recording replications for Parkinson's disease detection". *Expert Systems with Applications* vol. 46, pp. 286-292, 2015.
- [21] S. Timon, M. Rincón, R. Martínez-Tomás. "Extending XNAT Platform with an Incremental Semantic Framework". *Frontiers in Neuroinformatics*. vol 11, art. 5, 2017.
- [22] T. Fladby, L. Pålhaugen, P. Selnes, K. Waterloo, G. Bråthen, E. Hessen. "Detecting at-risk alzheimer's disease cases". *Journal of Alzheimer's Disease*. Vol. 60(1), pp. 97-105, 2017.