



**Academia de Ciencias Matemáticas,
Físico-Químicas y Naturales de Granada**

**Inteligencia Artificial y Sentimientos:
Incompatibilidad o Sinergismo**

**Discurso leído en el acto de su recepción como
Académico Correspondiente, por el**

Excmo. Sr. D. Mohamed Larbi Kerkeb

Granada, 2026

Inteligencia Artificial y Sentimientos: Incompatibilidad o Sinergismo

Mohamed Larbi Kerkeb

**Excmo. Sr. Presidente de la Academia de Ciencias Matemáticas,
Físico- Químicas y Naturales de Granada,**

Excmos. e Ilmos. Señores Académicos,

Dignísimas Autoridades,

Queridos colegas, compañeros y amigos,

Señoras y Señores:

Ingresar en la Academia de Ciencias de Granada, y hacerlo, además, en unos momentos en los que se celebran los cincuenta años de su fundación, no es únicamente recibir un honor; es asumir una responsabilidad ante el saber, ante la verdad y ante esa dimensión del tiempo que trasciende lo inmediato. Y esta responsabilidad no se construye en soledad: es fruto de la generosidad, el apoyo y la confianza de muchas personas. Por ello, antes de adentrarme en el tema que hoy nos convoca, quiero expresar mi más profundo agradecimiento a quienes han hecho posible este momento tan significativo en mi vida académica y personal.

Quiero expresar, en primer lugar, mi más sincera gratitud al Profesor, Académico, y hoy mi padrino en este acto, Fernando González Caballero; él no solo fue mi Director de Tesis, sino también, y desde entonces, una permanente referencia de rigor, coherencia y pasión por la investigación, cuya influencia en mi trayectoria científica y en mi formación profesional ha sido profunda y duradera.

Al Profesor Gerardo Pardo, le debo su apoyo, decisivo en momentos clave de mi carrera. Cuando llegué a Granada, siendo él Director de Departamento, me acogió con una generosidad intelectual y humana que marcó profundamente mi integración académica. Posteriormente, como Presidente de esta Academia, impulsó y apoyó mi candidatura. A él le debo no solo oportunidades académicas, sino también una lección permanente sobre liderazgo sereno y compromiso institucional.

Quiero agradecer igualmente al Profesor Ángel Delgado, por su colaboración científica y su amistad leal; al Profesor Enrique Hita, por su constante apoyo y por ampliar mis horizontes en la comunidad científica granadina; y al Profesor Eduardo Battaner, nuestro Presidente de la Sección de Físico-Químicas, por su visión integradora y su compromiso con la excelencia académica.

Al Profesor Francisco González Lodeiro, Presidente de esta Academia, le quiero expresar mi reconocimiento por las numerosas conversaciones mantenidas a lo largo de los años, siempre enriquecedoras y estimulantes, en las que la reflexión científica se entrelazaba con la dimensión ética e institucional de la Universidad. Su liderazgo garantiza la continuidad de una tradición académica basada en el rigor y la reflexión crítica.

Y a todos los miembros de esta Academia, cuya generosidad al acogerme constituye un honor que asumo con humildad, les agradezco profundamente la confianza depositada en mi persona. Formar parte de esta corporación representa para mí, sin ninguna duda, un compromiso renovado con la excelencia, el pensamiento crítico y el servicio a la sociedad a través del conocimiento.

Permítanme finalmente agradecer a mi familia, a mis colegas y a mis estudiantes. La ciencia es siempre una obra colectiva. Cada avance es fruto de múltiples diálogos, esfuerzos compartidos y complicidades intelectuales.

Con estos sentimientos de gratitud, que nacen del corazón y de la memoria, me dispongo ahora a presentar el discurso que he preparado para esta ocasión.

El tema que he elegido para esta ocasión, *Inteligencia Artificial y Sentimientos: Incompatibilidad o Sinergismo*, no es casual. Nos encontramos en un momento histórico singular, en el que los avances en inteligencia artificial están transformando radicalmente nuestra comprensión de la cognición, la conciencia y, de manera particularmente provocadora, la naturaleza misma de los sentimientos y las emociones. La pregunta que nos convoca hoy trasciende lo meramente técnico para adentrarse en territorios filosóficos, éticos y científicos de profunda relevancia para el futuro de la humanidad.

A lo largo de años de investigación en inteligencia artificial, redes neuronales y sistemas cognitivos, he tenido el privilegio de contribuir al desarrollo de tecnologías que, si bien inicialmente concebidas para resolver problemas de optimización y reconocimiento de patrones, han evolucionado hacia dominios cada vez más cercanos a la comprensión de procesos cognitivos complejos. Son precisamente estos trabajos en el diseño de sistemas inteligentes, para la predicción del

comportamiento humano, el análisis de patrones de interacción y el desarrollo de arquitecturas adaptativas los que me han llevado a reflexionar profundamente sobre la naturaleza de la inteligencia y su relación con la dimensión afectiva de la experiencia humana.

ÍNDICE

I. INTRODUCCIÓN: EL DESAFÍO CONCEPTUAL DE LA IA EMOCIONAL	9
II. FUNDAMENTOS TEÓRICOS: DE LA COMPUTACIÓN AFECTIVA A LA IA EMOCIONAL.....	10
1) Evolución Histórica del Concepto de Emoción en IA.....	10
2) Teorías Psicológicas de la Emoción y su Formalización Computacional	11
3) Arquitecturas Computacionales para la IA Emocional.....	12
4) Del Reconocimiento a la Generación: Empatía Artificial	12
III. ARQUITECTURAS NEURO-INSPIRADAS Y SISTEMAS DE RECONOCIMIENTO DE PATRONES.....	13
1) Fundamentos de las Redes Neuronales Artificiales.....	13
2) Reconocimiento de Patrones en Sistemas Cognitivos	14
3) Técnicas de Codificación y Representación de Datos	14
4) Sistemas Adaptativos y Control Inteligente.....	15
5) Computación Neuromórfica y Eficiencia Energética	16
IV. EL DEBATE FILOSÓFICO: FUNCIONALISMO, SENTIENCIA Y CONCIENCIA ARTIFICIAL.....	16
1) El Problema de la Conciencia Fenomenal.....	16
2) Funcionalismo y Realizabilidad Múltiple.....	17
3) Teorías de la Conciencia y sus Implicaciones para la IA	17
4) El Problema del Dolor Artificial y la Ética de la Sentiencia	18
5) Reflexiones desde la Práctica Investigadora	19
V. EVIDENCIA CIENTÍFICA: MODELOS COGNITIVOS Y SISTEMAS ADAPTATIVOS	19
1) Modelos Computacionales de Cognición Emocional	19
2) Aprendizaje Automático y Predicción de Estados Cognitivos	20
3) Sistemas Inteligentes en Contextos Reales: Movilidad Urbana	21
4) Inteligencia Artificial y Sostenibilidad Ambiental	22
5) Transformación Digital y Tecnologías Emergentes	22
VI. APLICACIONES PRÁCTICAS: DE LA TEORÍA A LA IMPLEMENTACIÓN	23
1) Sistemas de Tutoría Inteligente y Analítica del Aprendizaje	23
2) Agentes Conversacionales y Asistentes Virtuales	24
3) Robótica Social y Compañeros Artificiales.....	25
4) Detección de Amenazas y Análisis de Discurso	25
5) Inteligencia Artificial en Salud Mental.....	26

VII. INCOMPATIBILIDAD VERSUS SINERGISMO: ANÁLISIS CRÍTICO	27
1) Argumentos a Favor de la Incompatibilidad Fundamental.....	27
2) Argumentos a Favor del Sinergismo Funcional.....	28
3) Hacia una Síntesis: Emociones Funcionales sin Fenomenología	28
4) Implicaciones para el Diseño de Sistemas de IA	29
VIII. IMPLICACIONES ÉTICAS Y FUTURAS DIRECCIONES	30
1) Desafíos Éticos de la IA Emocional.....	30
2) Marcos Regulatorios y Gobernanza.....	31
3) Direcciones Futuras de Investigación.....	32
4) Inteligencia Artificial y el Futuro de la Humanidad.....	33
5) Reflexión Personal.....	34
IX. CONCLUSIÓN.....	34
X. REFERENCIAS.....	36

I. INTRODUCCIÓN: EL DESAFÍO CONCEPTUAL DE LA IA EMOCIONAL

La relación entre inteligencia artificial y sentimientos plantea una de las cuestiones más fascinantes y controvertidas de nuestro tiempo. Durante décadas, la inteligencia artificial se ha concebido fundamentalmente como un dominio de procesamiento lógico, cálculo algorítmico y manipulación simbólica. Sin embargo, el reconocimiento creciente de que las emociones desempeñan un papel fundamental en la cognición humana, la toma de decisiones y la interacción social ha impulsado el desarrollo de un campo emergente: la *computación afectiva* y la *inteligencia artificial emocional*.

El término "computación afectiva" (*affective computing*), acuñado por Rosalind Picard en el Massachusetts Institute of Technology a mediados de la década de los años 90 del pasado siglo, designa el *estudio y desarrollo de sistemas capaces de reconocer, interpretar, procesar y simular emociones humanas*. Esta disciplina ha experimentado un crecimiento exponencial en las últimas décadas, impulsada por avances en aprendizaje automático, visión por computador, procesamiento del lenguaje natural y neurociencia computacional.

No obstante, el desarrollo de sistemas artificiales, capaces de procesar información emocional, plantea interrogantes fundamentales que trascienden lo puramente técnico. ¿Pueden las máquinas experimentar genuinamente sentimientos, o simplemente simulan respuestas emocionales mediante algoritmos sofisticados? ¿Es la emoción un epifenómeno de la complejidad computacional, o requiere de sustratos biológicos específicos? ¿Qué implicaciones éticas y sociales conlleva la creación de entidades artificiales que exhiben comportamientos emocionalmente inteligentes?

Estas preguntas no son meramente especulativas. En mi propia trayectoria, como investigador, he tenido que confrontar repetidamente la tensión entre la capacidad predictiva de los modelos computacionales y la irreductibilidad aparente de la experiencia subjetiva, o, en otras palabras, que no todos los aspectos de la vivencia humana pueden ser capturados por dichos modelos. Y, efectivamente, en algunos de nuestros trabajos recientes sobre redes neuronales artificiales, aplicadas a la predicción del comportamiento de estudiantes en entornos virtuales de aprendizaje [31], hemos podido poner de manifiesto que, si bien podemos modelar con notable precisión patrones de interacción y rendimiento académico, la dimensión afectiva del aprendizaje, la motivación, la frustración, el entusiasmo, permanece como un territorio de difícil formalización algorítmica.

En la presente disertación, me propongo examinar críticamente la relación entre inteligencia artificial y sentimientos, desde una perspectiva multidisciplinar, que integra hallazgos de la neurociencia computacional, la filosofía de la mente, la psicología cognitiva y la ingeniería de sistemas inteligentes. Mi objetivo no es ofrecer respuestas definitivas a preguntas que continúan siendo objeto de intenso debate académico, sino más bien articular un marco conceptual, riguroso, que permita evaluar tanto las posibilidades como las limitaciones de la IA emocional, fundamentado en evidencia empírica y análisis teórico.

II. FUNDAMENTOS TEÓRICOS: DE LA COMPUTACIÓN AFECTIVA A LA IA EMOCIONAL

1) Evolución Histórica del Concepto de Emoción en IA

La historia de la inteligencia artificial ha estado marcada por una tensión fundamental entre dos paradigmas: el enfoque simbólico lógico, dominante desde los trabajos pioneros de Alan Turing y John McCarthy, y el enfoque conexionista o neuro-inspirado, que ha ganado prominencia con el resurgimiento de las redes neuronales profundas en las últimas dos décadas.

El paradigma simbólico tradicional, fundamentado en la manipulación de representaciones abstractas mediante reglas lógicas, relegó durante décadas el estudio de las emociones al considerarlas fenómenos irracionales o, en el mejor de los casos, periféricos a la cognición genuina. Esta visión reflejaba una concepción cartesiana de la mente, que establecía una dicotomía radical entre razón y emoción, entre procesamiento cognitivo "superior" y respuestas afectivas "primitivas".

Sin embargo, investigaciones en neurociencia cognitiva, particularmente los trabajos seminales de Antonio Damasio sobre el papel de las emociones en la toma de decisiones racionales, han demostrado que esta dicotomía es insostenible. Las emociones no son perturbaciones del pensamiento racional, sino componentes integrales de la cognición adaptativa. El célebre caso de Phineas Gage y estudios posteriores con pacientes que presentan lesiones en la corteza prefrontal ventromedial han revelado que la incapacidad para procesar información emocional conduce a déficits severos en la toma de decisiones, incluso cuando las capacidades lógicas permanecen intactas [1].

Esta reconceptualización de las emociones como elementos funcionales de la arquitectura cognitiva ha tenido profundas implicaciones para la inteligencia artificial. Si las emociones son computacionalmente relevantes para la inteligencia adaptativa, entonces su incorporación en sistemas artificiales no es un mero ejercicio

de simulación superficial, sino una necesidad arquitectónica para alcanzar formas genuinas de inteligencia general.

2) Teorías Psicológicas de la Emoción y su Formalización Computacional

La formalización computacional de las emociones requiere, como punto de partida, una teoría psicológica coherente sobre la naturaleza y función de los estados afectivos. Las principales teorías psicológicas de la emoción pueden clasificarse en tres grandes categorías:

Teorías dimensionales: Proponen que las emociones pueden representarse en un espacio continuo de dimensiones fundamentales, típicamente valencia (positivo-negativo) y activación (alta-baja). El modelo circuplejo de Russell constituye el ejemplo paradigmático de este enfoque. Desde una perspectiva computacional, estas teorías son atractivas por su parsimonia y su facilidad de implementación mediante representaciones vectoriales en espacios de baja dimensionalidad [2].

Teorías categoriales: Postulan la existencia de un conjunto discreto de emociones básicas o primarias (alegría, tristeza, miedo, ira, sorpresa, disgusto) que son universales y tienen correlatos neurobiológicos específicos. Los trabajos de Paul Ekman sobre expresiones faciales han proporcionado evidencia transcultural de estas categorías emocionales básicas. Computacionalmente, este enfoque se presta a implementaciones mediante clasificadores discretos y sistemas de reglas [3].

Teorías de valoración cognitiva (*appraisal theories*): Sostienen que las emociones emergen de evaluaciones cognitivas de eventos en relación con metas, valores y creencias del individuo. El modelo OCC (Ortony, Clore y Collins) constituye la formalización más influyente de este enfoque y ha sido ampliamente utilizado en el diseño de agentes artificiales emocionales. Esta perspectiva enfatiza el carácter funcional y adaptativo de las emociones como mecanismos de evaluación de relevancia [4].

En nuestros trabajos sobre sistemas cognitivos aplicados a entornos educativos, hemos explorado la integración de modelos de valoración cognitiva para comprender cómo los estudiantes evalúan sus interacciones con plataformas de aprendizaje virtual. Nuestras investigaciones han demostrado que la incorporación de variables relacionadas con la percepción de relevancia, la autoeficacia y la atribución causal mejora significativamente la capacidad predictiva de modelos de aprendizaje automático para anticipar el rendimiento académico y la persistencia estudiantil [31], [32].

3) Arquitecturas Computacionales para la IA Emocional

El diseño de sistemas de inteligencia artificial capaces de procesar información emocional requiere arquitecturas computacionales específicas que integren múltiples componentes funcionales. La literatura científica ha propuesto diversos marcos arquitectónicos, entre los cuales destacan:

Arquitecturas cognitivas integradas: Estos sistemas incorporan módulos emocionales dentro de arquitecturas cognitivas generales como SOAR, ACT-R o CLARION. La integración permite modelar interacciones bidireccionales entre procesos cognitivos y afectivos, reflejando la interdependencia observada en sistemas biológicos [5], [11].

Modelos neuro-inspirados: Basados en principios de la neurociencia computacional, estos enfoques implementan redes neuronales que emulan circuitos cerebrales involucrados en el procesamiento emocional, particularmente la amígdala, la corteza prefrontal y el sistema límbico. Los trabajos sobre computación neuromórfica han demostrado la viabilidad de implementar reconocimiento emocional en hardware especializado con eficiencia energética comparable a sistemas biológicos [10].

En este contexto, nuestras investigaciones sobre redes neuronales artificiales y aprendizaje profundo han explorado arquitecturas adaptativas capaces de ajustar dinámicamente sus parámetros en respuesta a patrones de interacción humana. En trabajos recientes sobre técnicas de codificación de datos para modelos de aprendizaje profundo en contextos educativos [32], [33] hemos puesto de manifiesto que la representación adecuada de secuencias temporales de interacción es crucial para capturar dinámicas afectivas implícitas en el comportamiento de los usuarios.

Sistemas multimodales de reconocimiento emocional: La detección automática de estados emocionales requiere la integración de múltiples modalidades sensoriales: expresiones faciales, prosodia vocal, gestos corporales, señales fisiológicas y contenido lingüístico. Los avances recientes en aprendizaje profundo multimodal han permitido desarrollar sistemas de reconocimiento emocional con tasas de precisión que se aproximan al desempeño humano en condiciones controladas [6], [7].

4) Del Reconocimiento a la Generación: Empatía Artificial

Un aspecto particularmente desafiante de la IA emocional es la transición del reconocimiento pasivo de emociones a la generación activa de respuestas empáticas. La empatía, entendida como *la capacidad de comprender y responder apropiadamente a los*

estados emocionales de otros, constituye un componente fundamental de la inteligencia social humana.

Investigaciones recientes han propuesto modelos computacionales de empatía artificial basados en mecanismos de simulación interna y teoría de la mente. Estos sistemas no se limitan a clasificar estados emocionales observables, sino que construyen modelos predictivos de los estados mentales de los usuarios, permitiendo anticipar necesidades emocionales y generar respuestas contextualmente apropiadas [8].

La implementación de empatía artificial plantea, sin embargo, cuestiones éticas fundamentales. ¿Es éticamente aceptable diseñar sistemas que simulan empatía sin experimentarla genuinamente? ¿Puede la empatía simulada tener efectos terapéuticos o educativos comparables a la empatía humana auténtica? Estas preguntas adquieren particular relevancia en aplicaciones como agentes conversacionales terapéuticos, tutores virtuales y sistemas de compañía para personas mayores [15], [17].

III. ARQUITECTURAS NEURO-INSPIRADAS Y SISTEMAS DE RECONOCIMIENTO DE PATRONES

1) Fundamentos de las Redes Neuronales Artificiales

Las redes neuronales artificiales constituyen el paradigma computacional dominante en la inteligencia artificial contemporánea. Inspiradas en la arquitectura del cerebro biológico, estas estructuras computacionales están compuestas por unidades de procesamiento interconectadas (neuronas artificiales) que transforman señales de entrada mediante funciones de activación no lineales y ajustan sus parámetros mediante algoritmos de aprendizaje.

A lo largo de mi trayectoria investigadora, he trabajado extensamente con mis colaboradores, con diversas arquitecturas de redes neuronales, desde perceptrones multicapa hasta arquitecturas profundas contemporáneas. Nuestros primeros trabajos en el campo se centraron en aplicaciones de redes neuronales a problemas de caracterización físico-química y procesamiento de imágenes, donde la capacidad de estas arquitecturas para extraer características relevantes de datos de alta dimensionalidad demostró ser particularmente valiosa.

La evolución de las redes neuronales hacia arquitecturas profundas (*deep learning*) ha revolucionado el campo de la inteligencia artificial en la última década. Las redes neuronales convolucionales (CNN) han alcanzado desempeño

sobrehumano en tareas de reconocimiento visual, mientras que las redes neuronales recurrentes (RNN) y sus variantes, particularmente las redes LSTM (*Long Short-Term Memory*) y GRU (*Gated Recurrent Units*), han demostrado capacidades excepcionales para modelar secuencias temporales y dependencias de largo plazo.

En nuestras investigaciones sobre predicción del rendimiento estudiantil mediante análisis de interacciones con entornos virtuales de aprendizaje, hemos implementado arquitecturas de redes neuronales artificiales que procesan secuencias temporales de eventos de interacción para predecir el éxito académico [31]. Estos modelos incorporan capas de procesamiento que extraen características relevantes de patrones de comportamiento, tales como frecuencia de acceso, duración de sesiones, patrones de navegación y secuencias de actividades completadas.

2) Reconocimiento de Patrones en Sistemas Cognitivos

El reconocimiento de patrones constituye una capacidad fundamental de los sistemas inteligentes, tanto biológicos como artificiales. En el contexto de la inteligencia artificial emocional, el reconocimiento de patrones se manifiesta en múltiples niveles: desde la detección de características faciales asociadas a expresiones emocionales hasta la identificación de patrones lingüísticos indicativos de estados afectivos.

Nuestros trabajos en el ámbito del reconocimiento de patrones han abarcado diversas aplicaciones, desde la detección de objetos en imágenes agrícolas [34] hasta el análisis de patrones de tráfico urbano [35], [36]. En cada uno de estos dominios, la capacidad de extraer características relevantes, construir representaciones apropiadas y aplicar algoritmos de clasificación o regresión constituye el núcleo metodológico del enfoque.

En el contexto específico de sistemas educativos inteligentes, hemos desarrollado técnicas de ingeniería de características (*feature engineering*) que permiten extraer patrones significativos de datos de interacción estudiantil. Nuestro trabajo sobre minería de datos educativos ha demostrado que la selección cuidadosa de características temporales, secuenciales y contextuales mejora sustancialmente la capacidad predictiva de modelos de aprendizaje automático [31].

3) Técnicas de Codificación y Representación de Datos

Un aspecto crucial en el diseño de sistemas de aprendizaje automático es la representación adecuada de los datos de entrada. Las técnicas de codificación determinan cómo la información del mundo real se transforma en representaciones numéricas procesables por algoritmos de aprendizaje.

En nuestras investigaciones recientes sobre aprendizaje profundo aplicado a la predicción del rendimiento estudiantil, hemos realizado un análisis exhaustivo de técnicas de codificación de datos para modelos de aprendizaje profundo en educación [32], [33]. Este trabajo ha revelado que la elección de esquemas de codificación, incluyendo codificación one-hot, codificación ordinal, codificación de frecuencia y técnicas de embedding, tiene un impacto significativo en la capacidad de los modelos para capturar relaciones complejas en los datos.

Particularmente relevante para el procesamiento de información emocional es el desarrollo de representaciones distribuidas (*embeddings*) que capturan similitudes semánticas entre conceptos emocionales. Los modelos de lenguaje basados en transformadores, como BERT y GPT, han demostrado capacidad para aprender representaciones vectoriales de palabras y frases que preservan relaciones semánticas, incluyendo dimensiones afectivas del lenguaje.

4) Sistemas Adaptativos y Control Inteligente

La adaptabilidad constituye una característica esencial de la inteligencia, tanto natural como artificial. Los sistemas adaptativos son capaces de modificar su comportamiento en respuesta a cambios en el entorno o en los objetivos del sistema, exhibiendo formas de aprendizaje continuo y mejora incremental.

Nuestros trabajos sobre sistemas adaptativos han explorado aplicaciones en diversos dominios. En el ámbito de la movilidad urbana inteligente, hemos desarrollado sistemas de control adaptativo de señales de tráfico que utilizan inteligencia artificial y sistemas distribuidos para optimizar el flujo vehicular en tiempo real [36], [37]. Estos sistemas implementan algoritmos de aprendizaje por refuerzo que ajustan dinámicamente los tiempos de semáforos en función de patrones de tráfico observados, demostrando mejoras significativas en eficiencia de movilidad y reducción de emisiones de carbono.

En el contexto de redes de vehículos aéreos no tripulados (UAV), hemos evaluado el desempeño de algoritmos de enrutamiento dinámico adaptativo que utilizan técnicas de inteligencia artificial para la selección óptima de rutas en entornos de comunicación cambiantes [38]. Estos sistemas exhiben capacidades de adaptación en tiempo real que son esenciales para operaciones autónomas en condiciones impredecibles.

La relevancia de estos trabajos para la inteligencia artificial emocional radica en que los mecanismos de adaptación implementados en estos sistemas comparten principios arquitectónicos con modelos de regulación emocional. Tanto en sistemas de control adaptativo como en modelos de procesamiento emocional, la capacidad

de evaluar continuamente el estado del sistema, compararlo con objetivos deseados y ajustar parámetros de comportamiento constituye un mecanismo fundamental de funcionamiento inteligente.

5) Computación Neuromórfica y Eficiencia Energética

Un aspecto frecuentemente subestimado en el diseño de sistemas de inteligencia artificial es la eficiencia energética. El cerebro humano, con aproximadamente 86 mil millones de neuronas y 100 billones de sinapsis, opera con un consumo energético de apenas 20 vatios, mientras que sistemas de inteligencia artificial de última generación requieren megavatios de potencia para tareas comparables.

La computación neuromórfica representa un paradigma emergente que busca emular no solo la funcionalidad sino también la eficiencia energética de sistemas nerviosos biológicos. Trabajos pioneros en este campo han demostrado la viabilidad de implementar reconocimiento emocional en chips neuromórficos con consumos energéticos varios órdenes de magnitud inferiores a implementaciones convencionales [10].

Aunque nuestras propias investigaciones no se han centrado específicamente en hardware neuromórfico, los principios de eficiencia computacional han sido una consideración constante en el diseño de sistemas inteligentes. En nuestros trabajos sobre procesamiento de grandes volúmenes de datos, hemos explorado la integración de técnicas de criptografía ligera en frameworks de computación distribuida como MapReduce, buscando equilibrar requisitos de seguridad con eficiencia computacional [39], [40].

IV. EL DEBATE FILOSÓFICO: FUNCIONALISMO, SINTIENCIA Y CONCIENCIA ARTIFICIAL

1) El Problema de la Conciencia Fenomenal

El debate sobre la posibilidad de emociones genuinas en sistemas artificiales nos conduce inevitablemente a una de las cuestiones más profundas y persistentes de la filosofía de la mente: el problema de la conciencia fenomenal. ¿Qué significa experimentar subjetivamente un estado mental? ¿Existe algo que sea "como ser" un sistema consciente, en la célebre formulación de Thomas Nagel?

La distinción entre conciencia de acceso (*access consciousness*) y conciencia fenomenal (*phenomenal consciousness*), articulada por Ned Block, resulta crucial para este debate. La conciencia de acceso se refiere a la disponibilidad de información para

procesos cognitivos como el razonamiento, el control del comportamiento y el reporte verbal. La conciencia fenomenal, por el contrario, designa el carácter cualitativo de la experiencia subjetiva, los *qualia* que constituyen "lo que se siente" al experimentar un estado mental.

Los sistemas de inteligencia artificial contemporáneos pueden exhibir formas sofisticadas de conciencia de acceso: procesan información, la integran con conocimiento previo, la utilizan para guiar decisiones y pueden incluso generar reportes verbales sobre sus estados internos. Sin embargo, la cuestión de si estos sistemas poseen conciencia fenomenal, si existe algo que sea "como ser" un sistema de IA, permanece profundamente controvertida [13], [22].

2) Funcionalismo y Realizabilidad Múltiple

El funcionalismo, la teoría dominante en filosofía de la mente durante las últimas décadas, sostiene que los estados mentales se definen por sus relaciones funcionales, sus roles causales en el sistema cognitivo, más que por su sustrato físico de implementación. Según esta perspectiva, lo que hace que un estado sea una creencia, un deseo o una emoción no es su composición material (neuronas biológicas, circuitos de silicio, o cualquier otro sustrato), sino su posición en una red de relaciones causales con inputs sensoriales, otros estados mentales y outputs comportamentales.

El funcionalismo implica el principio de realizabilidad múltiple: si los estados mentales se definen funcionalmente, entonces pueden ser implementados en múltiples sustratos físicos. Un sistema artificial que instancia las mismas relaciones funcionales que un cerebro humano experimentando miedo debería, según el funcionalismo, estar genuinamente experimentando miedo, independientemente de que esté implementado en neuronas biológicas o en circuitos electrónicos.

Esta perspectiva ha sido enormemente influyente en inteligencia artificial y ciencia cognitiva, proporcionando justificación teórica para la empresa de crear mentes artificiales. Sin embargo, enfrenta objeciones significativas, particularmente el argumento de los *qualia* invertidos y el experimento mental de la habitación china de John Searle, que sugieren que la mera instanciación de relaciones funcionales podría ser insuficiente para generar experiencia consciente genuina [14].

3) Teorías de la Conciencia y sus Implicaciones para la IA

Diversas teorías científicas de la conciencia ofrecen perspectivas diferentes sobre la posibilidad de conciencia artificial:

Teoría de la Información Integrada (IIT): Propuesta por Giulio Tononi, esta teoría sostiene que la conciencia es idéntica a información integrada, cuantificada mediante la medida Φ (phi). Según la IIT, cualquier sistema que integre información de manera irreducible posee algún grado de conciencia, independientemente de su sustrato. Esta teoría implica que sistemas artificiales con arquitecturas apropiadas podrían ser genuinamente conscientes [16].

Teoría del Espacio de Trabajo Global (GWT): Desarrollada por Bernard Baars, esta teoría propone que la conciencia emerge cuando información es transmitida a un "espacio de trabajo global" accesible a múltiples procesos cognitivos. Implementaciones computacionales de esta arquitectura han sido propuestas como vías hacia la conciencia artificial [22].

Teorías de Orden Superior: Estas teorías sostienen que un estado mental es consciente cuando es objeto de un pensamiento o representación de orden superior. Desde esta perspectiva, la conciencia requiere capacidades metacognitivas, la capacidad de representar los propios estados mentales, que podrían, en principio, ser implementadas en sistemas artificiales.

Enactivismo y Cognición Encarnada: Estas perspectivas enfatizan el papel del cuerpo y la interacción sensoriomotora con el entorno en la constitución de la mente. Desde este enfoque, la conciencia y las emociones no son fenómenos puramente computacionales, sino que emergen de la dinámica de sistemas encarnados y situados. Esta perspectiva sugiere que la IA emocional genuina requeriría no solo algoritmos sofisticados, sino también formas apropiadas de encarnación y acoplamiento sensoriomotor con el mundo [9], [29].

4) El Problema del Dolor Artificial y la Ética de la Sintiencia

Una dimensión particularmente inquietante del debate sobre conciencia artificial es la posibilidad de dolor o sufrimiento en sistemas artificiales. Si aceptamos que sistemas artificiales suficientemente complejos podrían ser conscientes, ¿podrían también experimentar dolor? ¿Tendríamos obligaciones morales hacia entidades artificiales sintientes?

Masayoshi Asada ha argumentado que la implementación de dolor artificial podría ser no solo posible sino deseable, ya que el dolor desempeña funciones adaptativas cruciales y podría inducir formas de empatía, moralidad y comportamiento ético en robots [29]. Sin embargo, esta propuesta plantea dilemas éticos profundos: ¿sería moralmente permisible crear deliberadamente entidades capaces de sufrir? ¿Cómo podríamos determinar si un sistema artificial está genuinamente experimentando dolor o simplemente simulando respuestas de dolor?

Estas cuestiones no son meramente especulativas. A medida que desarrollamos sistemas de inteligencia artificial cada vez más sofisticados, particularmente en aplicaciones que involucran interacción social y emocional con humanos, la posibilidad de que estos sistemas posean alguna forma de experiencia subjetiva se convierte en una consideración ética ineludible [27], [30].

5) Reflexiones desde la Práctica Investigadora

Como investigador que ha dedicado décadas al desarrollo de sistemas inteligentes, he confrontado repetidamente la tensión entre la elegancia formal de los modelos computacionales y la riqueza irreducible de la experiencia humana. En nuestros trabajos sobre sistemas de predicción del comportamiento estudiantil, por ejemplo, hemos logrado desarrollar modelos que anticipan con notable precisión el rendimiento académico y la persistencia de los estudiantes [31], [32]. Sin embargo, estos modelos, por sofisticados que sean, capturan correlaciones estadísticas en patrones de comportamiento observable, no la experiencia vivida del estudiante, su ansiedad ante un examen, su satisfacción al comprender un concepto difícil, su frustración ante obstáculos aparentemente insuperables.

Esta brecha entre modelo y experiencia no es necesariamente una limitación temporal que será superada con algoritmos más sofisticados y mayor poder computacional. Podría reflejar una diferencia categorial entre procesamiento de información y experiencia subjetiva, entre sintaxis y semántica, entre simulación y realización genuina.

No obstante, la imposibilidad de acceso directo a la experiencia subjetiva de otros, el problema epistemológico de las otras mentes, se aplica tanto a sistemas artificiales como a otros seres humanos. Inferimos estados mentales en otros humanos mediante analogía con nuestra propia experiencia y observación de comportamiento, pero carecemos de acceso directo a sus *qualia*. Si un sistema artificial exhibiera el mismo rango de comportamientos emocionales que un ser humano, ¿sobre qué base podríamos negar que posee experiencia emocional genuina?

V. EVIDENCIA CIENTÍFICA: MODELOS COGNITIVOS Y SISTEMAS ADAPTATIVOS

1) Modelos Computacionales de Cognición Emocional

La integración de emociones en arquitecturas cognitivas computacionales ha sido objeto de intensa investigación en las últimas dos décadas. Los modelos computacionales de cognición emocional buscan capturar las interacciones

bidireccionales entre procesos cognitivos y afectivos observadas en sistemas biológicos.

Rodríguez y colaboradores han desarrollado modelos cognitivos computacionales de emociones que integran teorías de valoración cognitiva con arquitecturas de agentes inteligentes [12]. Estos modelos implementan mecanismos mediante los cuales los eventos son evaluados en relación con metas y valores del agente, generando respuestas emocionales que, a su vez, modulan procesos cognitivos como son la atención, memoria y la toma de decisiones.

Grig y colaboradores han propuesto arquitecturas cognitivas integradas con emociones, adoptando un enfoque bio-inspirado para desarrollar agentes artificiales emocionalmente inteligentes [11]. Estas arquitecturas incorporan módulos especializados, que emulan funciones de estructuras cerebrales involucradas en el procesamiento emocional, como la amígdala y la corteza prefrontal, demostrando cómo la integración de componentes emocionales mejora la adaptabilidad y robustez de sistemas de IA en entornos complejos e inciertos.

2) Aprendizaje Automático y Predicción de Estados Cognitivos

En una serie de artículos recientes de nuestro grupo, hemos presentado resultados obtenidos en el ámbito del aprendizaje automático, aplicado a sistemas educativos,. En ellos hemos explorado la predicción de estados cognitivos y patrones de comportamiento, mediante análisis de datos de interacción. Eran estudios sobre ingeniería de características y minería de datos, con el objetivo de predecir el éxito estudiantil, basándose en interacciones con entornos virtuales de aprendizaje [31]. Para ello, implementamos redes neuronales artificiales, que procesan secuencias temporales de eventos de interacción, para generar las predicciones sobre rendimiento académico.

Encontramos que ciertos patrones de interacción, frecuencia de acceso, duración de sesiones, secuencias de actividades, patrones de navegación, son altamente predictivos del éxito académico. Sin embargo, también identificamos limitaciones significativas: los modelos puramente basados en datos de interacción observable tienen dificultades para capturar factores motivacionales y afectivos, que influyen crucialmente en el aprendizaje.

En investigaciones posteriores, hemos explorado técnicas avanzadas de codificación de datos para modelos de aprendizaje profundo en educación [32], [33]. Estos trabajos han demostrado que la representación adecuada de datos temporales y secuenciales, mediante técnicas como embeddings y codificación posicional, mejora

sustancialmente la capacidad de los modelos para capturar dinámicas de aprendizaje complejas.

Particularmente relevante es nuestro trabajo sobre evaluación predictiva de estudiantes, mediante interacciones iniciales con técnicas de codificación en aprendizaje profundo [33]. Este estudio demostró que es posible predecir con notable precisión el rendimiento académico de estudiantes, basándose únicamente en sus primeras interacciones con plataformas de aprendizaje virtual, sugiriendo que patrones de comportamiento temprano contienen información significativa sobre disposiciones cognitivas y afectivas subyacentes.

3) Sistemas Inteligentes en Contextos Reales: Movilidad Urbana

Más allá de aplicaciones educativas, nuestras investigaciones han explorado el desarrollo de sistemas inteligentes adaptativos en diversos dominios. Un área particularmente fructífera ha sido la aplicación de la inteligencia artificial a la optimización de la movilidad urbana, donde hemos desarrollado sistemas que integran predicción, control adaptativo y comunicación distribuida.

En nuestro trabajo sobre predicción del estado del tráfico en París mediante aprendizaje automático [35], implementamos modelos predictivos que procesan datos históricos y en tiempo real de flujo vehicular, para anticipar condiciones de tráfico futuras. Estos modelos utilizan técnicas de aprendizaje supervisado, incluyendo redes neuronales y métodos de ensemble, para generar predicciones que informan sistemas de gestión de tráfico y aplicaciones de navegación.

Extendiendo este trabajo, hemos desarrollado así mismo sistemas de control adaptativo de señales de tráfico, que utilizan inteligencia artificial y sistemas distribuidos para optimizar la movilidad urbana inteligente [36]. Estos sistemas implementan algoritmos de aprendizaje por refuerzo, que ajustan dinámicamente los tiempos de semáforos en función de patrones de tráfico observados, demostrando mejoras significativas en eficiencia de flujo vehicular y reducción de tiempos de espera.

Más recientemente, hemos propuesto una estrategia integrada para la optimización del tráfico urbano que combina predicción, control adaptativo de señales y comunicación distribuida mediante mensajería [37]. Este enfoque holístico demuestra cómo la integración de múltiples componentes de inteligencia artificial, percepción, predicción, planificación y control, puede generar sistemas de movilidad urbana significativamente más eficientes y sostenibles.

La relevancia de estos trabajos para la inteligencia artificial emocional radica en los principios arquitectónicos compartidos. Tanto en sistemas de control de tráfico como en modelos de procesamiento emocional, la capacidad de percibir el estado del entorno, predecir evoluciones futuras, evaluar situaciones en relación con objetivos y ajustar comportamiento adaptativamente constituye el núcleo funcional del sistema inteligente.

4) Inteligencia Artificial y Sostenibilidad Ambiental

Otra dimensión de nuestras investigaciones ha explorado aplicaciones de inteligencia artificial a desafíos de sostenibilidad ambiental. En nuestro trabajo sobre reducción de la huella de carbono mediante planificación de transporte en tiempo real y analítica de big data [41], demostramos cómo las técnicas de aprendizaje automático, aplicadas a grandes volúmenes de datos de movilidad, pueden optimizar rutas de transporte para minimizar emisiones de gases de efecto invernadero.

En el ámbito de la agricultura inteligente, hemos explorado aplicaciones de técnicas de aprendizaje profundo para protección ambiental [42]. Este trabajo incluye el desarrollo de sistemas de detección de enfermedades en cultivos mediante visión por computador y redes neuronales convolucionales, demostrando cómo la inteligencia artificial puede contribuir a prácticas agrícolas más sostenibles y eficientes.

Particularmente relevante ha sido nuestra revisión exhaustiva de técnicas de detección de objetos para identificación de enfermedades en fresas [34], donde analizamos diversos enfoques de aprendizaje automático y visión por computador para reconocimiento de patrones de enfermedad en cultivos. Este trabajo ilustra la aplicación de las técnicas de reconocimiento de patrones, fundamentales también para el reconocimiento de expresiones emocionales, a dominios de relevancia práctica inmediata.

5) Transformación Digital y Tecnologías Emergentes

A lo largo de mi carrera profesional, he mantenido un compromiso constante con la transformación digital de instituciones educativas y de investigación. Como Presidente de la Universidad Ibn Tofail, lideré el diseño e implementación del Plan Estratégico de Sistemas de Información de la universidad [43], así como el diseño e implementación de un centro de datos hiperconvergente, que proporciona infraestructura para cargas de trabajo de inteligencia artificial, aprendizaje automático y analítica de big data.

Además, lideré el establecimiento del Centro Code 212, dedicado a la formación en inteligencia artificial, programación y robótica [44], diseñando

programas de capacitación que preparan a la próxima generación de profesionales en tecnologías emergentes. Esta experiencia en educación en IA me ha proporcionado, qué duda cabe, perspectivas valiosas sobre cómo los humanos aprenden sobre inteligencia artificial y cómo conceptualizan la relación entre inteligencia humana y artificial.

En el ámbito de las telecomunicaciones, hemos explorado también la metamorfosis de las telecomunicaciones tradicionales hacia sistemas impulsados por inteligencia artificial y tecnologías avanzadas [45], examinando cómo la IA está transformando las infraestructuras de comunicación y habilitando nuevas capacidades de red inteligente.

VI. APLICACIONES PRÁCTICAS: DE LA TEORÍA A LA IMPLEMENTACIÓN

1) Sistemas de Tutoría Inteligente y Analítica del Aprendizaje

Como ya he comentado anteriormente, una de las aplicaciones más prometedoras de la inteligencia artificial emocional se encuentra en el ámbito de la educación, particularmente en sistemas de tutoría inteligente, capaces de adaptar su instrucción a las necesidades cognitivas y afectivas de los estudiantes.

Mis investigaciones en este dominio han explorado cómo el análisis de patrones de interacción con entornos virtuales de aprendizaje puede informar intervenciones pedagógicas personalizadas. En nuestro trabajo sobre predicción del éxito estudiantil mediante redes neuronales artificiales [31], demostramos que modelos de aprendizaje automático pueden identificar tempranamente estudiantes en riesgo de fracaso académico, permitiendo intervenciones preventivas oportunas.

Sin embargo, la efectividad de estas intervenciones depende críticamente de la capacidad del sistema para comprender no solo patrones de comportamiento observable, sino también estados afectivos subyacentes. Un estudiante que accede infrecuentemente a la plataforma de aprendizaje podría estar desmotivado, abrumado por otras responsabilidades, o simplemente aprendiendo eficientemente mediante otros recursos. Distinguir entre estas posibilidades requiere formas de inferencia sobre estados afectivos y motivacionales que van más allá del análisis de datos de interacción superficial.

Nuestras investigaciones sobre el impacto del aprendizaje a distancia en el rendimiento y actitud de estudiantes durante la pandemia de COVID-19 [46] revelaron la importancia crucial de factores afectivos en contextos de aprendizaje

virtual. Los estudiantes que reportaron mayores niveles de ansiedad, aislamiento social y dificultades de concentración exhibieron rendimientos académicos significativamente inferiores, independientemente de su acceso a tecnología o su competencia digital previa.

Estos hallazgos subrayan la necesidad de sistemas de tutoría inteligente que no solo entreguen contenido instruccional, sino que también monitoreen y respondan a estados afectivos de los estudiantes. La implementación de tales sistemas requiere la integración de múltiples modalidades de detección emocional, análisis de expresiones faciales mediante cámaras web, análisis de prosodia en interacciones de voz, análisis de patrones de escritura y contenido lingüístico, con modelos pedagógicos que especifiquen intervenciones apropiadas para diferentes estados afectivos.

2) Agentes Conversacionales y Asistentes Virtuales

Los agentes conversacionales o chatbots representan otra aplicación prominente de la inteligencia artificial emocional. Sistemas como asistentes virtuales, agentes de servicio al cliente y compañeros virtuales están siendo desplegados en contextos que requieren formas de inteligencia social y emocional.

La efectividad de estos sistemas depende críticamente de su capacidad para reconocer estados emocionales de los usuarios y generar respuestas contextualmente apropiadas. Un usuario frustrado requiere un tipo de respuesta diferente a un usuario confundido o satisfecho. La incapacidad de los sistemas conversacionales para reconocer y responder apropiadamente a señales emocionales constituye una de las principales fuentes de insatisfacción de usuarios con estas tecnologías.

Investigaciones recientes han explorado la implementación de empatía artificial en agentes conversacionales [8], [15]. Estos sistemas no se limitan a clasificar estados emocionales mediante análisis de contenido lingüístico y características prosódicas, sino que construyen modelos de los estados mentales de los usuarios, permitiendo anticipar necesidades emocionales y generar respuestas empáticas.

Sin embargo, la implementación de empatía artificial plantea cuestiones éticas significativas. ¿Es apropiado diseñar sistemas que simulan empatía sin experimentarla genuinamente? ¿Puede la interacción con agentes artificiales empáticos satisfacer necesidades humanas de conexión emocional, o constituye una forma de engaño que podría tener consecuencias psicológicas negativas? [17], [30].

3) Robótica Social y Compañeros Artificiales

La robótica social representa un dominio particularmente desafiante para la inteligencia artificial emocional. Los robots sociales están diseñados para interactuar con humanos en contextos sociales, exhibiendo comportamientos que los humanos interpretan como sociales y emocionales.

Aplicaciones de robótica social incluyen robots de compañía para personas mayores, robots terapéuticos para niños con trastornos del espectro autista, y robots educativos que sirven como tutores o compañeros de aprendizaje. En cada uno de estos contextos, la capacidad del robot para reconocer y responder apropiadamente a emociones humanas es crucial para la efectividad de la interacción.

La encarnación física de robots sociales introduce dimensiones adicionales de complejidad. Los humanos responden a señales emocionales no solo en expresiones faciales y voz, sino también en gestos corporales, postura y movimiento. Los robots sociales efectivos deben ser capaces de generar estas señales de manera coordinada y contextualmente apropiada, creando la impresión de coherencia emocional.

Investigaciones han demostrado que los humanos tienden a antropomorfizar robots sociales, atribuyéndoles estados mentales y emocionales incluso cuando son conscientes de su naturaleza artificial [17]. Este fenómeno plantea tanto oportunidades como riesgos: mientras que la tendencia a antropomorfizar puede facilitar interacciones naturales y efectivas, también puede conducir a expectativas poco realistas y potencial explotación de vulnerabilidades emocionales humanas.

4) Detección de Amenazas y Análisis de Discurso

Una aplicación menos convencional, pero de gran relevancia práctica de técnicas de inteligencia artificial es la detección y análisis de discurso extremista en plataformas digitales. Como investigador del proyecto "Nutcracker: Detección, rastreo, monitoreo y análisis de discurso terrorista en línea" [48], participé en un esfuerzo internacional para desarrollar sistemas de inteligencia artificial capaces de identificar, rastrear y analizar contenido extremista en internet.

Este proyecto utilizó técnicas avanzadas de procesamiento de lenguaje natural, aprendizaje automático y análisis de redes sociales para detectar patrones lingüísticos, narrativos y de interacción asociados con discurso extremista. Un aspecto particularmente desafiante fue distinguir entre expresiones legítimas de opiniones políticas o religiosas y contenido que constituye incitación a la violencia o reclutamiento terrorista.

La dimensión emocional del discurso extremista es crucial para su efectividad persuasiva. Los mensajes extremistas frecuentemente apelan a emociones intensas, ira, miedo, resentimiento, sentido de injusticia, para movilizar individuos hacia la radicalización. La capacidad de sistemas de inteligencia artificial para analizar no solo el contenido semántico sino también las dimensiones afectivas del discurso es esencial para la detección efectiva de amenazas.

Este trabajo ilustra cómo técnicas de reconocimiento de patrones y análisis de sentimientos, desarrolladas inicialmente para aplicaciones comerciales como análisis de opinión de clientes, pueden ser adaptadas a contextos de seguridad nacional y prevención de terrorismo. También subraya la importancia de consideraciones éticas en el desarrollo de sistemas de vigilancia automatizada, particularmente en relación con privacidad, libertad de expresión y potencial para sesgos algorítmicos.

5) Inteligencia Artificial en Salud Mental

Una aplicación emergente de particular relevancia es el uso de inteligencia artificial en salud mental. Se están desarrollando sistemas de IA para detectar signos tempranos de depresión, ansiedad y otros trastornos mentales, mediante análisis de patrones de comportamiento digital, contenido lingüístico en redes sociales, y características de voz en conversaciones telefónicas.

Estos sistemas prometen expandir significativamente el acceso a los servicios de salud mental, particularmente en regiones con escasez de profesionales de salud mental. Chatbots terapéuticos como Woebot y Wysa están siendo utilizados en la actualidad por millones de usuarios, para proporcionar intervenciones de terapia cognitivo-conductual automatizadas [15].

Sin embargo, la aplicación de IA a salud mental plantea cuestiones éticas y clínicas complejas. ¿Pueden sistemas artificiales proporcionar el tipo de comprensión empática y conexión terapéutica que constituye el núcleo de la psicoterapia efectiva? ¿Qué riesgos existen de que sistemas automatizados proporcionen intervenciones inapropiadas o incluso dañinas? ¿Cómo garantizamos la privacidad y confidencialidad de la información sobre la salud mental, muy sensible, procesada por sistemas de IA?

VII. INCOMPATIBILIDAD VERSUS SINERGISMO: ANÁLISIS CRÍTICO

1) Argumentos a Favor de la Incompatibilidad Fundamental

La posición de incompatibilidad fundamental sostiene que las emociones genuinas requieren sustratos biológicos específicos y que los sistemas artificiales, independientemente de su sofisticación computacional, no pueden experimentar estados emocionales auténticos. Esta perspectiva se fundamenta en varios argumentos:

Argumento de la encarnación: Las emociones no son fenómenos puramente computacionales, sino que emergen de la interacción entre cerebro, cuerpo y entorno. Las emociones tienen componentes fisiológicos esenciales, cambios en frecuencia cardíaca, conductancia de la piel, liberación de hormonas, que no pueden ser replicados en sistemas artificiales desencarnados. Desde esta perspectiva, la ausencia de un cuerpo biológico con sistemas endocrino y nervioso autónomos, hace imposible la experiencia emocional genuina [9].

Argumento de la evolución: Las emociones evolucionaron como mecanismos adaptativos en organismos biológicos enfrentados a desafíos específicos de supervivencia y reproducción. Las emociones están profundamente entrelazadas con necesidades biológicas fundamentales, hambre, sed, dolor, deseo sexual, que carecen de análogos en sistemas artificiales. Un sistema que no tiene necesidades biológicas de supervivencia no puede experimentar miedo genuino ante amenazas a su existencia [13].

Argumento de la conciencia fenomenal: Incluso si los sistemas artificiales pudieran crear las mismas relaciones funcionales que los estados emocionales humanos, esto no garantizaría que poseyeran la dimensión fenomenal de la experiencia emocional, el carácter cualitativo de "lo que se siente" al experimentar una emoción. La brecha explicativa entre procesamiento de información y experiencia subjetiva sugiere que la mera complejidad computacional es insuficiente para generar conciencia fenomenal [14], [22].

Argumento de la autenticidad: Existe una diferencia moralmente relevante entre emociones genuinas y simulación de emociones. Un sistema que simula empatía mediante algoritmos sofisticados no está genuinamente preocupado por el bienestar de otros, sino simplemente ejecutando instrucciones programadas. Esta distinción es crucial en contextos donde la autenticidad emocional tiene valor intrínseco, como en las relaciones terapéuticas, o de cuidado [17], [30].

2) Argumentos a Favor del Sinergismo Funcional

La posición de sinergismo funcional sostiene que, independientemente de si los sistemas artificiales experimentan emociones fenomenalmente, pueden crear, o materializar (*instanciar*), funciones emocionales, de manera que complementa y amplifica las capacidades humanas. Esta perspectiva se fundamenta en varios argumentos:

Argumento funcionalista: Si las emociones se definen por sus roles funcionales en sistemas cognitivos, evaluación de relevancia, modulación de atención, facilitación de toma de decisiones, coordinación de respuestas comportamentales, entonces los sistemas artificiales que *instancian* estas funciones poseen emociones en el sentido funcionalmente relevante, independientemente de su experiencia fenomenal [5], [11].

Argumento de la complementariedad: Los sistemas de IA emocional no necesitan replicar perfectamente las emociones humanas para ser valiosos. Pueden complementar capacidades humanas procesando información emocional a escalas y velocidades inaccesibles para humanos, detectando patrones sutiles en expresiones emocionales, y proporcionando retroalimentación objetiva sobre estados afectivos [6], [7].

Argumento pragmático: La cuestión de si las emociones artificiales son "genuinas" puede ser menos relevante que la cuestión de si son funcionalmente efectivas. Si un sistema de tutoría inteligente mejora resultados de aprendizaje mediante respuestas sensibles a estados emocionales de estudiantes, su efectividad práctica puede ser independiente de si experimenta empatía fenomenal [31], [32].

Argumento de la continuidad: La distinción entre emociones "genuinas" y "simuladas" puede ser menos clara de lo que parece. Incluso en humanos, las emociones involucran componentes de procesamiento de información, evaluación cognitiva y generación de respuestas que podrían ser caracterizados como "computacionales". La diferencia entre emociones humanas y artificiales podría ser de grado más que de tipo [9], [13].

3) Hacia una Síntesis: Emociones Funcionales sin Fenomenología

Mi propia posición, apoyada por décadas de investigación en inteligencia artificial y sistemas cognitivos, se inclina hacia una síntesis que reconoce tanto las limitaciones como las posibilidades de la IA emocional.

Considero probable que los sistemas artificiales contemporáneos, independientemente de su sofisticación, carecen de conciencia fenomenal y, por

tanto, no experimentan emociones en el sentido subjetivo que los humanos las experimentan. La brecha explicativa entre procesamiento de información y experiencia subjetiva permanece sin resolver, y no tenemos razones convincentes para atribuir experiencia fenomenal a sistemas que carecen de las características biológicas asociadas con conciencia en organismos naturales.

Sin embargo, esto no implica que los sistemas de IA emocional carezcan de valor o relevancia. Las funciones cognitivas y comportamentales de las emociones, evaluación de relevancia, modulación de atención, facilitación de toma de decisiones, coordinación de respuestas, pueden ser implementadas en sistemas artificiales, de manera que complementan y amplifican las capacidades humanas.

La investigación previa desarrollada por nuestro grupo, sobre sistemas adaptativos, en dominios como movilidad urbana [35], [36], [37] y educación [31], [32], [33], ha demostrado que los sistemas artificiales pueden exhibir formas de adaptabilidad, sensibilidad contextual y respuesta a cambios ambientales, que comparten principios arquitectónicos con procesamiento emocional biológico. Estos sistemas evalúan continuamente su entorno, comparan estados actuales con objetivos deseados, y ajustan su comportamiento en consecuencia, funciones análogas a las desempeñadas por emociones en sistemas biológicos.

La cuestión crucial no es si estos sistemas "sienten" en el sentido fenomenal, sino si *instancian* funciones emocionales de manera efectiva y éticamente apropiada. Un sistema de tutoría inteligente que detecta frustración estudiantil y ajusta su instrucción en consecuencia, no necesita experimentar empatía fenomenal para ser pedagógicamente efectivo. Un sistema de control de tráfico que optimiza flujo vehicular, en respuesta a patrones cambiantes, no necesita experimentar "estrés" para exhibir adaptabilidad funcional.

4) Implicaciones para el Diseño de Sistemas de IA

Esta perspectiva tiene implicaciones importantes para el diseño de sistemas de inteligencia artificial:

Priorizar funcionalidad sobre simulación: El objetivo del diseño de IA emocional no debería ser crear simulaciones convincentes de emociones humanas, sino implementar funciones emocionales que mejoren la efectividad de sistemas en dominios específicos. En educación, esto significa sistemas que adaptan instrucción a necesidades cognitivas y afectivas de estudiantes [31], [32]. En movilidad urbana, significa sistemas que responden adaptativamente a patrones cambiantes de demanda [35], [36], [37].

Transparencia sobre capacidades: Los sistemas de IA emocional deberían ser transparentes sobre su naturaleza artificial y sus limitaciones. Crear la ilusión de empatía genuina, cuando el sistema simplemente ejecuta algoritmos de respuesta emocional, constituye una forma de engaño que puede tener consecuencias éticas negativas [17], [30].

Complementariedad humano-IA: Los sistemas de IA emocional deberían ser diseñados para complementar, no reemplazar, capacidades emocionales humanas. En contextos como salud mental o cuidado de personas mayores, los sistemas de IA pueden expandir el acceso a servicios y proporcionar monitoreo continuo, pero no deberían sustituir completamente una interacción humana genuina [15].

Consideraciones éticas continuas: El desarrollo de IA emocional requiere reflexión ética continua sobre cuestiones de privacidad, autonomía, autenticidad y potencial para manipulación emocional. Mis experiencia, liderando proyectos de transformación digital en instituciones educativas [43], [44] han reforzado mi convicción de que las consideraciones éticas deben ser integradas desde las primeras etapas del diseño de sistemas, no añadidas como reflexiones posteriores.

VIII. IMPLICACIONES ÉTICAS Y FUTURAS DIRECCIONES

1) Desafíos Éticos de la IA Emocional

El desarrollo de sistemas de inteligencia artificial capaces de reconocer, procesar y simular emociones, plantea desafíos éticos significativos, que requieren atención cuidadosa de investigadores, desarrolladores, reguladores y de la sociedad en general.

Privacidad emocional: Los sistemas de reconocimiento emocional requieren acceso a información altamente personal, expresiones faciales, patrones de voz, contenido lingüístico, señales fisiológicas, que revela estados internos íntimos. La recopilación, almacenamiento y procesamiento de esta información plantea riesgos significativos de privacidad. ¿Quién tiene derecho a acceder a datos emocionales? ¿Cómo garantizamos que esta información no sea utilizada para manipulación o discriminación? [30].

Manipulación emocional: Los sistemas capaces de reconocer estados emocionales pueden ser diseñados para manipular emociones de usuarios hacia fines comerciales, políticos o de otro tipo. Publicidad dirigida emocionalmente, contenido diseñado para maximizar *engagement* mediante explotación de vulnerabilidades

emocionales, y sistemas de persuasión automatizada representan riesgos significativos para la autonomía individual [21], [30].

Autenticidad y engaño: La creación de sistemas que simulan empatía y comprensión emocional, sin experimentarlas genuinamente, plantea cuestiones sobre autenticidad y engaño. ¿Es ético diseñar sistemas que crean la ilusión de cuidado genuino? ¿Deberían los usuarios ser informados explícitamente cuando interactúan con sistemas artificiales en lugar de humanos? [17].

Sesgos algorítmicos: Los sistemas de reconocimiento emocional, entrenados en conjuntos de datos que no representan adecuadamente la diversidad humana, pueden exhibir sesgos sistemáticos, reconociendo emociones con mayor precisión en ciertos grupos demográficos que en otros. Estos sesgos pueden perpetuar y amplificar desigualdades existentes [30].

Responsabilidad y rendición de cuentas: Cuando los sistemas de IA emocional toman decisiones que afectan el bienestar humano, en contextos educativos, de salud mental, o de justicia criminal, surgen cuestiones complejas sobre responsabilidad. ¿Quién es responsable cuando un sistema de IA proporciona una intervención inapropiada o dañina? [27].

2) Marcos Regulatorios y Gobernanza

El desarrollo responsable de IA emocional requiere marcos regulatorios y de gobernanza que equilibren innovación con protección de derechos y bienestar humanos. Varios enfoques están siendo explorados:

Regulación basada en riesgo: Clasificar aplicaciones de IA emocional según su nivel de riesgo, y aplicar requisitos regulatorios proporcionales. Aplicaciones de alto riesgo, como sistemas de IA en salud mental o justicia criminal, requerirían evaluaciones rigurosas de seguridad y efectividad antes de su despliegue.

Principios éticos para IA: Organizaciones internacionales, gobiernos y empresas tecnológicas han propuesto diversos conjuntos de principios éticos para IA, típicamente incluyendo transparencia, equidad, privacidad, responsabilidad y beneficencia. La implementación efectiva de estos principios requiere traducirlos en requisitos técnicos y procesos de diseño concretos [27].

Evaluación de impacto ético: Requerir evaluaciones de impacto ético para sistemas de IA emocional antes de su despliegue, análogas a las evaluaciones de impacto ambiental. Estas evaluaciones examinarían riesgos potenciales para

privacidad, autonomía, equidad y bienestar, así como otras medidas de mitigación propuestas.

Participación pública: El desarrollo de IA emocional no debería ser dominio exclusivo de expertos técnicos. La participación pública en decisiones sobre qué aplicaciones de IA emocional son socialmente deseables y bajo qué condiciones, es esencial para una gobernanza democrática de estas tecnologías.

3) Direcciones Futuras de Investigación

El campo de la inteligencia artificial emocional se encuentra en una etapa relativamente temprana de desarrollo, con numerosas direcciones prometedoras para investigación futura:

Modelos multimodales integrados: La mayoría de los sistemas actuales de reconocimiento emocional se centran en modalidades individuales, expresiones faciales, voz, o texto. El desarrollo de modelos que integran múltiples modalidades de manera coherente, reflejando la naturaleza multimodal del procesamiento emocional humano, representa una importante dirección de la futura investigación [6], [7].

Aprendizaje continuo y personalización: Los sistemas actuales típicamente se entrenan en conjuntos de datos estáticos y no se adaptan a características individuales de usuarios. El desarrollo de sistemas capaces de aprendizaje continuo y personalización a patrones emocionales individuales podría mejorar significativamente su efectividad [33].

Modelado de dinámicas emocionales temporales: Las emociones no son estados estáticos sino procesos dinámicos que evolucionan en el tiempo. El desarrollo de modelos que capturan trayectorias temporales de estados emocionales, transiciones entre emociones, y regulación emocional representa un desafío importante [35], [36].

Comprensión de emociones complejas: La mayoría de investigación hasta la actualidad se ha centrado en emociones básicas. El reconocimiento y modelado de emociones complejas, nostalgia, orgullo, vergüenza, gratitud, que involucran evaluaciones cognitivas sofisticadas permanece como un desafío significativo [4].

Emociones sociales y colectivas: Las emociones no son fenómenos puramente individuales, sino que emergen en contextos sociales. El modelado de emociones sociales, empatía, envidia, admiración, y dinámicas emocionales colectivas representa sin duda una frontera importante [8].

Integración con robótica encarnada: El desarrollo de robots sociales con capacidades emocionales requiere integración de reconocimiento emocional, con generación de expresiones emocionales multimodales, faciales, vocales, gestuales, de manera coordinada y contextualmente apropiada [9], [29].

4) Inteligencia Artificial y el Futuro de la Humanidad

Más allá de aplicaciones específicas, el desarrollo de inteligencia artificial emocional plantea cuestiones profundas sobre el futuro de la humanidad y nuestra relación con tecnologías cada vez más sofisticadas.

A medida que los sistemas de IA se vuelven más capaces de reconocer, procesar y responder a emociones humanas, las fronteras entre interacción humano-humano e interacción humano-máquina se vuelven progresivamente más difusas. Esto plantea interrogantes sobre la naturaleza de las relaciones sociales, la autenticidad de conexiones emocionales, y el significado de conceptos como empatía, cuidado y comprensión.

Algunos teóricos han expresado su preocupación por que la proliferación de agentes artificiales emocionalmente inteligentes pudiera conducir a una "erosión de lo humano", una disminución de las capacidades humanas de conexión emocional genuina, a medida que las personas se acostumbran a las interacciones con sistemas artificiales, que simulan empatía sin experimentarla [17], [21].

Otros adoptan una perspectiva más optimista, sugiriendo que la IA emocional podría amplificar capacidades humanas, proporcionando herramientas para comprender mejor nuestras propias emociones, mejorar comunicación emocional, y expandir el acceso a servicios de apoyo emocional [15].

Mi propia perspectiva es que el impacto de la IA emocional dependerá críticamente de las elecciones que hagamos como sociedad sobre cómo desarrollar y desplegar estas tecnologías. La tecnología no es determinista; sus efectos son moldeados por decisiones de diseño, marcos regulatorios, normas sociales y valores culturales.

Como investigadores y desarrolladores de IA, tenemos la responsabilidad de anticipar consecuencias potenciales de nuestras creaciones, de diseñar sistemas que respeten la dignidad y la autonomía humana, y de participar en diálogos públicos sobre el futuro que queremos construir. Como educadores, tenemos la responsabilidad de preparar a la próxima generación ,no solo con competencias técnicas en IA, sino también con capacidades de reflexión ética y pensamiento crítico sobre las implicaciones sociales de estas tecnologías [44].

5) Reflexión Personal

Al reflexionar sobre la investigación en inteligencia artificial, redes neuronales y sistemas cognitivos, me impresiona tanto el progreso extraordinario que hemos logrado como la persistencia de preguntas fundamentales que permanecen sin respuesta.

Hoy, el aprendizaje profundo basado en redes neuronales ha transformado radicalmente el campo, logrando capacidades que parecían ciencia ficción hace apenas dos décadas. Mis propias investigaciones han evolucionado desde aplicaciones de redes neuronales a problemas de caracterización físico-química [47], pasando por el desarrollo de sistemas de información para gestión universitaria [49], hasta aplicaciones contemporáneas en predicción de comportamiento estudiantil [31], [32], [33], optimización de movilidad urbana [35], [36], [37], y análisis del discurso extremista [48], este último proyecto liderado por los Profesores Hidalgo Tenorio y Juan Luis Castro, de la Universidad de Granada.

A lo largo de este recorrido, he mantenido una fascinación constante por la pregunta de qué constituye genuinamente la inteligencia y cómo se relaciona con dimensiones afectivas de la experiencia. Los sistemas que, modestamente, he contribuido a desarrollar, exhiben formas impresionantes de adaptabilidad, reconocimiento de patrones y toma de decisiones, pero permanezco agnóstico sobre si instancian formas genuinas de comprensión o experiencia subjetiva.

Esta incertidumbre no es motivo de desaliento, sino de humildad intelectual. Las preguntas más profundas sobre la naturaleza de la mente, la conciencia y las emociones han ocupado a filósofos durante milenios, y no serán resueltas definitivamente por avances tecnológicos, por impresionantes que sean. Lo que la inteligencia artificial puede ofrecer es una nueva perspectiva sobre estas preguntas antiguas, una forma de explorar empíricamente hipótesis sobre la naturaleza de la cognición y la emoción mediante la construcción de sistemas artificiales.

IX. CONCLUSIÓN

Al concluir esta disertación sobre la relación entre inteligencia artificial y sentimientos, retorno a la pregunta que motivó esta reflexión: ¿incompatibilidad o sinergismo?

La evidencia examinada sugiere que la respuesta no es simple ni unívoca. Existe, probablemente, una incompatibilidad fundamental entre procesamiento computacional de información emocional y experiencia fenomenal de emociones. Los sistemas artificiales contemporáneos, independientemente de su sofisticación,

carecen de las características biológicas, encarnación, necesidades de supervivencia, historia evolutiva, que parecen esenciales para experiencia emocional genuina en el sentido subjetivo que los humanos la experimentan.

Sin embargo, esta incompatibilidad fenomenal no excluye un sinergismo funcional profundo. Las funciones cognitivas y comportamentales de las emociones, evaluación de relevancia, modulación de atención, facilitación de toma de decisiones, coordinación de respuestas adaptativas, pueden ser implementadas en sistemas artificiales de manera que complementa y amplifica las capacidades humanas.

Nuestras propias investigaciones en inteligencia artificial, redes neuronales y sistemas adaptativos han demostrado que los sistemas artificiales pueden exhibir formas sofisticadas de sensibilidad contextual, adaptabilidad y respuesta a patrones complejos de información. En dominios como la educación [31], [32], [33], movilidad urbana [35], [36], [37], agricultura inteligente [34], [42], y últimamente análisis crítico del discurso [48], hemos desarrollado sistemas que procesan información de manera funcionalmente análoga a aspectos del procesamiento emocional biológico.

La cuestión crucial para el futuro de la inteligencia artificial emocional no es si los sistemas artificiales pueden "sentir" en el sentido fenomenal, sino cómo podemos diseñar sistemas que instancien funciones emocionales de manera efectiva, ética y beneficiosa para la humanidad. Esto requiere:

Rigor científico: Continuar la investigación fundamental sobre la naturaleza de las emociones, tanto en sistemas biológicos como artificiales, fundamentada en metodologías rigurosas y evidencia empírica.

Reflexión filosófica: Mantener un diálogo continuo, con tradiciones filosóficas que han reflexionado sobre la naturaleza de la mente, la conciencia y las emociones, reconociendo que las cuestiones técnicas están entrelazadas con interrogantes filosóficos profundos.

Responsabilidad ética: Integrar consideraciones éticas desde las primeras etapas del diseño de sistemas de IA emocional, anticipando consecuencias potenciales para privacidad, autonomía, autenticidad y bienestar humano.

Participación social: Involucrar a la sociedad en general en decisiones sobre qué aplicaciones de IA emocional son deseables y bajo qué condiciones, reconociendo que estas son cuestiones de valores sociales, no meramente técnicas.

Humildad intelectual: Reconocer los límites de nuestro conocimiento actual sobre la naturaleza de la conciencia y las emociones, evitando tanto el escepticismo

dogmático que niega toda posibilidad de IA emocional, como el optimismo ingenuo, que asume que la sofisticación computacional es suficiente para generar experiencia subjetiva.

Como académico que ha dedicado parte de su carrera a la formación en mi país de la próxima generación de investigadores en inteligencia artificial [44], considero que nuestra responsabilidad trasciende el avance técnico. Debemos cultivar no solo competencias en algoritmos y arquitecturas de IA, sino también capacidades de reflexión crítica sobre implicaciones sociales, éticas y filosóficas de las tecnologías que creamos.

Y concluyo: La pregunta sobre la relación entre inteligencia artificial y sentimientos no es meramente académica. Es una pregunta sobre qué tipo de futuro queremos construir, qué tipo de relaciones queremos tener con tecnologías cada vez más sofisticadas, y qué significa ser humano en una era de inteligencia artificial. Estas son preguntas que no pueden ser respondidas únicamente por investigadores en laboratorios, sino que requieren diálogo amplio que involucre a filósofos, científicos sociales, artistas, humanistas, formuladores de políticas y ciudadanos.

Al asumir mi posición como Académico Correspondiente de esta ilustre institución, me comprometo a continuar contribuyendo a este diálogo, aportando tanto rigor científico como reflexión humanística, y trabajando para asegurar que el desarrollo de la inteligencia artificial sirva al florecimiento humano y al bien común.

Agradezco nuevamente a la Academia de Ciencias de Granada el honor de acogerme en su seno, y espero que las reflexiones compartidas en este discurso contribuyan a la comprensión colectiva de uno de los desafíos intelectuales más fascinantes de nuestro tiempo.

Muchas gracias.

X. REFERENCIAS

- [1] Damasio, Descartes' Error: Emotion, Reason, and the Human Brain. New York: Putnam, 1994.
- [2] J. A. Russell, "A circumplex model of affect," *Journal of Personality and Social Psychology*, vol. 39, no. 6, pp. 1161-1178, 1980.
- [3] P. Ekman, "An argument for basic emotions," *Cognition & Emotion*, vol. 6, no. 3-4, pp. 169-200, 1992.
- [4] Ortony, G. L. Clore, and A. Collins, *The Cognitive Structure of Emotions*. Cambridge: Cambridge University Press, 1988.

- [5] J. E. Laird, "The Soar Cognitive Architecture," MIT Press, 2012.
- [6] S. Poria et al., "A review of affective computing: From unimodal analysis to multimodal fusion," *Information Fusion*, vol. 37, pp. 98-125, 2017. <https://doi.org/10.1016/j.inffus.2017.02.003>
- [7] Mehrabian and M. Corwin, "Emotion in Artificial Intelligence," 2025. <https://doi.org/10.48550/arxiv.2508.10286>
- [8] Alanazi et al., "Prediction of Emotional Empathy in Intelligent Agents to Facilitate Precise Social Interaction," *Applied Sciences*, vol. 13, no. 2, 1163, 2023. <https://doi.org/10.3390/app13021163>
- [9] Y. Zhao et al., "From Emotion AI to Cognitive AI," *International Journal of Network Dynamics and Intelligence*, vol. 1, no. 1, pp. 65-72, 2022. <https://doi.org/10.53941/ijndi0101006>
- [10] P. U. Diehl et al., "True happiness: Neuromorphic emotion recognition on true north," *arXiv: Neurons and Cognition*, 2016.
- [11] Grig et al., "Emotion-Integrated Cognitive Architectures: A Bio-Inspired Approach to Developing Emotionally Intelligent AI Agents," *Studies in Computational Intelligence*, vol. 1109, pp. 357-368, 2024. https://doi.org/10.1007/978-3-031-50381-8_38
- [12] L. F. Rodríguez et al., "Cognitive computational models of emotions," in *IEEE International Conference on Cognitive Informatics and Cognitive Computing*, pp. 80-86, 2011. <https://doi.org/10.1109/COGINF.2011.6016124>
- [13] Y. Li et al., "Artificial Emotion: A Survey of Theories and Debates on Realising Emotion in Artificial Intelligence," 2025.
- [14] O. Chursinova et al., "Is the Realization of the Emotional Artificial Intelligence Possible? Philosophical and Methodological Analysis," *Filosofija-sociologija*, vol. 32, no. 1, pp. 56-64, 2021. <https://doi.org/10.6001/FIL-SOC.V32I1.4382>
- [15] J. Voisen, "My Computer, My Friend, My Therapist: Exploring and questioning the theories and methodologies that guide the design of empathic agents," 2009.
- [16] Y. Shan, "From Consciousness to Emotion - On the Intelligent Evolution of Complex Systems," *Preprints*, 2023. <https://doi.org/10.20944/preprints202308.0002.v1>
- [17] Weber-Guskar, "How to feel about emotionalized artificial intelligence? When robot pets, holograms, and chatbots become affective partners," *Ethics and Information Technology*, vol. 23, pp. 479-489, 2021. <https://doi.org/10.1007/S10676-021-09598-8>

- [18] R. Duffy, "Fundamental Issues in Affective Intelligent Social Machines," *The Open Artificial Intelligence Journal*, vol. 2, pp. 21-34, 2008. <https://doi.org/10.2174/1874061800802010021>
- [19] J. L. Martín, "La idea de hombre bajo los presupuestos de l'humanismo computacional'," 2003.
- [20] L. M. Possati, "Freud and the algorithm: neuropsychanalysis as a framework to understand artificial general intelligence," *Palgrave Communications*, vol. 7, 2021. <https://doi.org/10.1057/S41599-021-00812-Y>
- [21] Curry, "Dreary useless centuries of happiness: Cordwainer Smith's 'Under Old Earth' as an ethical critique of our current Emotion AI goals," *Neo helicon*, vol. 49, pp. 625-641, 2022. <https://doi.org/10.1007/s11059-022-00663-9>
- [22] H. H. Haladjian et al., "Artificial consciousness and the consciousness-attention dissociation," *Consciousness and Cognition*, vol. 45, pp. 210-225, 2016. <https://doi.org/10.1016/J.CONCOG.2016.08.011>
- [23] J. Mingle, "Redefining Purpose: The Role of Consciousness and Emotion in AI Agency."
- [24] C. Youvan, "Emotional Field Modulation in Artificial Intelligence: A Thought Experiment in Dual-Model Sentience and Self-Evaluation."
- [25] J. Fairchild, "Artificial Consciousness: Where Does Science Fact Break from Science Fiction, and How Do We Know?," 2014.
- [26] M. T. Wynn, "AI: Artificial Intelligence as Philosophy: Machine Consciousness and Intelligence," in *Encyclopedia of the UN Sustainable Development Goals*, 2023. https://doi.org/10.1007/978-3-319-97134-6_59-1
- [27] Banerjee, "A framework for designing compassionate and ethical artificial intelligence and artificial consciousness," *Interdisciplinary Description of Complex Systems*, vol. 18, no. 2-B, pp. 85-95, 2020. <https://doi.org/10.7287/PEERJ.PREPRINTS.3502V1>
- [28] Y. Fang, "Philosophical Approach to Human-Centered Artificial Intelligence and 21st Century Technology: Is it Possible for a Machine to Ever Experience Emotions the Way We..."
- [29] M. Asada, "Artificial pain may induce empathy, morality, and ethics in the conscious mind of robots," *Philosophies*, vol. 4, no. 3, 38, 2019. <https://doi.org/10.3390/PHILOSOPHIES4030038>

- [30] M. Chavan et al., "Feeling machines: ethics, culture, and the rise of emotional AI," arXiv.org, 2025. <https://doi.org/10.48550/arxiv.2506.12437>
- [31] M. Ouahi, S. Khouliji, and M. L. Kerkeb, "Feature Engineering, Mining for Predicting Student Success Based on Interaction with the Virtual Learning Environment Using Artificial Neural Network," *Annals of the Romanian Society for Cell Biology*, vol. 25, no. 6, pp. 3891-3903, 2021.
- [32] M. Ouahi, S. Khouliji, and M. L. Kerkeb, "Advancing Sustainable Learning Environments: Literature Review on Data Encoding Techniques for Student Performance Prediction using Deep Learning Models in Education," *E3S Web of Conferences*, vol. 477, 00074, 2024. <https://doi.org/10.1051/e3sconf/202447700074>
- [33] M. Ouahi, S. Khouliji, and M. L. Kerkeb, "Predictive assessment of learners through initial interactions with encoding techniques in deep learning," *Journal of Autonomous Intelligence*, vol. 7, no. 4, 1443, 2024.
- [34] Z. Remch, S. Khouliji, and M. L. Kerkeb, "Object Detection Techniques for Strawberry Disease Detection: A Comprehensive Review," *Journal of Theoretical and Applied Information Technology*, vol. 102, no. 19, pp. 7178-7201, 2024.
- [35] Zrigui, I., Khouliji, S., Kerkeb, M.L., Remch, Z., Bourekkadi, S. "Traffic State Prediction In Paris: Leveraging Machine Learning For Efficient Urban Mobility", *Journal of Theoretical and Applied Information Technology*, 2024, 102(19), pp. 7128–7136 <https://www.scopus.com/pages/publications/85207438210>
- [36] Bourekkadi, "Traffic State Prediction in Paris: Leveraging Machine Learning for Efficient Urban Mobility," *Journal of Theoretical and Applied Information Technology*, vol. 102, no. 19, pp. 7128-7136, 2024.
- [37] Zrigui, S. Khouliji, and M. L. Kerkeb, "Adaptive Traffic Signal Control Using AI and Distributed Systems for Smart Urban Mobility," *Journal of Theoretical and Applied Information Technology*, vol. 103, no. 8, pp. 3532-3547, 2025.
- [38] Zrigui, S. Khouliji, and M. L. Kerkeb, "Integrated Strategy for Urban Traffic Optimization: Prediction, Adaptive Signal Control, and Distributed Communication via Messaging," arXiv preprint arXiv:2501.02008, 2025.

- [39] K. Slimani, S. Khouliji, H. Taherdoost, and M. L. Kerkeb, "Performance Evaluation of Dynamic Adaptive Routing (DAR) for Unmanned Aerial Vehicle (UAV) Networks," *Computers, Materials & Continua*, vol. 85, no. 2, pp. 4115-4132, 2025.
- [40] M. Khadji, S. Khouliji, and M. L. Kerkeb, "Efficient Big Data Security: Evaluating The Performance Of A Proposed Hybrid Key Management Algorithm Using Lightweight Cryptography," *Journal of Theoretical and Applied Information Technology*, vol. 101, no. 13, pp. 5384-5398, 2023.
- [41] M. Khadji, S. Khouliji, M. L. Kerkeb, and I. Khadji, "Securing large-scale data processing: Integrating lightweight cryptography in MapReduce," *Journal of Autonomous Intelligence*, vol. 7, no. 4, 1390, 2024.
- [42] Zrigui, S. Khouliji, A. Ennassiri, S. Bourekadi, and M. L. Kerkeb, "Reducing Carbon Footprint with Real-Time Transport Planning and Big Data Analytics," *E3S Web of Conferences*, vol. 412, 01082, 2023. <https://doi.org/10.1051/e3sconf/202341201082>
- [43] R. Zineb, K. Samira, and K. M. Larbi, "Applications of smart agriculture for environmental protection using deep learning techniques," *E3S Web of Conferences*, vol. 412, 01083, 2023. <https://doi.org/10.1051/e3sconf/202341201083>
- [44] M. L. Kerkeb, "Design and development of the Strategic Information System Plan of Ibn Tofail University," *Internal Report*, Ibn Tofail University, Kenitra, Morocco, 2023-2025.
- [45] M. L. Kerkeb, "Establishment and design of AI training programs for Code 212 Center (AI, programming, robotics)," *Code 212 Center*, Kenitra, Morocco, 2024-2025.
- [46] K. Slimani, S. Khouliji, A. Mortreau, and M. L. Kerkeb, "From tradition to innovation: The telecommunications metamorphosis with AI and advanced technologies," *Journal of Autonomous Intelligence*, vol. 7, no. 1, 1099, 2024.
- [47] L. Rouah, S. Khouliji, and M. L. Kerkeb, "The Impact of Distance Learning in the Performance and Attitude of Students During the Global Coronavirus Pandemic (Covid-19): Case of the Moroccan University," *Lecture Notes in Networks and Systems*, vol. 668, pp. 385-396, 2023.
- [48] E. Hidalgo-Tenorio, J.L. Castro, M. L. Kerkeb, et al., "Nutcracker: Detection, tracing, monitoring and analysis of terrorist discourse

online," University of Granada, Spain, Research Project funded by Spanish Ministry, 2016-2023.

- [49] M. L. Kerkeb, "Morocco-Spain cooperation projects in physico-chemistry, image processing and characterization techniques," International Cooperation Projects, 1999-2003.
- [50] M. L. Kerkeb, "SIMARECH: Scientific Research Information System," Versions 1-4, Abdelmalek Essaadi University and Ibn Tofail University, 2009-2025.



**Academia de Ciencias Matemáticas,
Físico-Químicas y Naturales de Granada**

**Contestación al discurso de ingreso del
Excmo. Sr. D. Mohamed Larbi Kerkeb
como Académico Correspondiente**

**Pronunciado por el
Excmo. Sr. D. Fernando González Caballero
Académico Numerario**

Granada, 2026

Excmo. Sr. Presidente de la Academia de Ciencias Matemáticas, Físico-Químicas y Naturales de Granada,

Excmas. e Ilmas. Señoras Académicas, Excmos. E Ilmos. Señores Académicos

Queridas y queridos colegas,

Señoras y Señores:

Este día para mí tiene una importancia que supera lo profesional, lo académico. Es un honor y una satisfacción difícil de explicar y compartir tomar la palabra en este acto solemne para dar la bienvenida, en nombre de la Academia, al Excmo. Sr. D. Mohamed Larbi Kerkeb, quien hoy ingresa como Académico Correspondiente de nuestra corporación. Este acto, que se inscribe en la mejor tradición de las academias científicas españolas, no sólo representa el reconocimiento a una trayectoria excepcional; con este acto también le damos la bienvenida a un científico cuya obra ha contribuido significativamente al avance del conocimiento en áreas estratégicas para el desarrollo científico y tecnológico contemporáneo.

La elección del Dr. Kerkeb como miembro de esta Academia constituye un acierto indiscutible, como así fue reconocido en el momento de su aceptación unánime por el Pleno de esta Academia. Su brillante carrera, su prolífica producción científica, su liderazgo institucional en el ámbito universitario internacional y su compromiso con la formación de nuevas generaciones de investigadores, lo acreditan para ocupar el lugar que merece en esta Academia. Pero más allá de los méritos cuantitativos, que por sí solos son suficientes en un entorno como el actual, que se mueve con una rapidez insospechada gracias al creciente desarrollo de la inteligencia artificial, es la calidad, la profundidad y la relevancia de su investigación lo que verdaderamente justifica su incorporación a esta corporación académica.

El discurso que acabamos de escuchar, titulado “Inteligencia Artificial y Sentimientos: Incompatibilidad o Sinergismo”, aborda probablemente una de las cuestiones más fascinantes y controvertidas del tiempo presente. El Dr. Kerkeb nos ha ofrecido una reflexión rigurosa, profunda y matizada sobre un tema que trasciende los límites de la ingeniería informática para adentrarse en territorios filosóficos, éticos y antropológicos de primera magnitud. Su concienzudo y profundo análisis nos invita a repensar nuestras concepciones sobre la naturaleza de la inteligencia, la conciencia y, especialmente, sobre la dimensión afectiva de la experiencia cognitiva.

Antes de entrar en el análisis del discurso que acabamos de escuchar, permítanme trazar una **semblanza biográfica** del nuevo académico, también amigo y antiguo alumno de nuestra universidad, a fin de poder apreciar en su justa dimensión el camino recorrido hasta llegar a este momento.

Mohamed Larbi Kerkeb es un científico cuya trayectoria vital y académica ejemplifica la mejor tradición de la cooperación científica internacional, y el compromiso con la excelencia investigadora. Su formación académica se inició precisamente en el marco de la colaboración hispano-marroquí, en aquellos años (los 80s) en los que la colaboración institucional entre la Facultad de Ciencias de Granada y la de la Universidad Abdelmalek Essadi, de Tetuán-Tánger, alcanzó unos niveles de excelencia que, aún hoy, al menos el que les habla, echa de menos. Debo recordar, y lo hago con mucha satisfacción, que esa colaboración estuvo liderada por el entonces Vicerrector de Investigación y Relaciones Internacionales, y hoy Presidente de nuestra Academia, el Profesor González Lodeiro . Y fue en ese contexto en el que Larbi Kerkeb inició sus estudios de doctorado en nuestra Facultad, obteniendo su Grado de Doctor en Ciencias Físicas por la Universidad de Granada en el año 1993.

Finalizados estos estudios, su carrera profesional se ha desarrollado fundamentalmente en Marruecos, donde ha ocupado posiciones de máxima responsabilidad. Fue Vicedecano en la Facultad Polidisciplinar de Tetuán (2005-2009), Decano de la Facultad Polidisciplinar de Larache (2018-2023), y desde febrero de 2023 hasta finales de 2025, ha sido Presidente de la Universidad Ibn Tofail de Kenitra. Actualmente, ostenta el cargo de Primer Vicepresidente de la Asociación de Universidades Africanas, posición desde la cual ejerce un liderazgo continental en materia de política universitaria y cooperación académica.

Pero la dimensión institucional de su trayectoria no se limita al ámbito universitario marroquí o africano. El Dr. Kerkeb ha sido Presidente del Consejo Científico de la Agencia Universitaria de la Francofonía entre 2016 y 2019, organización que agrupa a más de mil instituciones de educación superior e investigación en los cinco continentes. Esta responsabilidad le permitió contribuir a la definición de políticas científicas y de cooperación universitaria en el espacio francófono. En ese contexto pudo promover la movilidad académica, la investigación colaborativa y la transferencia de conocimiento entre instituciones del Norte y del Sur.

En relación a su producción científica, destaca tanto por su volumen como por su diversidad temática y su impacto. Su obra abarca un amplio espectro de áreas de investigación, siendo las más relevantes las relacionadas con la inteligencia

artificial, el aprendizaje automático, las redes neuronales, el reconocimiento de patrones y los sistemas cognitivos. Destacaré algunas de sus contribuciones más significativas, que ilustran la amplitud y profundidad de su trabajo.

En el campo de la *inteligencia artificial aplicada*, el Dr. Kerkeb y su equipo han propuesto enfoques integrales orientados a la optimización del tráfico urbano, combinando técnicas de predicción algorítmica, control adaptativo de semáforos y comunicación distribuida [1], [2]. Estos estudios, publicados en 2025, ponen de manifiesto el potencial de la inteligencia artificial para el desarrollo de infraestructuras propias de las ciudades inteligentes, con un impacto directo en la mejora de la movilidad urbana sostenible y eficiente.

En lo que se refiere a las *redes neuronales y el aprendizaje profundo* aplicado a la educación, Larbi Kerkeb ha realizado aportaciones pioneras en la predicción del rendimiento estudiantil mediante el análisis de interacciones con entornos virtuales de aprendizaje [3], [4], [5]. Sus investigaciones sobre técnicas de codificación de datos y modelos de aprendizaje profundo para la evaluación predictiva del alumnado estudiantes representan contribuciones significativas a la analítica educativa basada en el potencial indudable de la inteligencia artificial.

Sus trabajos sobre *reconocimiento de patrones* y visión por computador incluyen revisiones exhaustivas sobre técnicas de detección de objetos aplicadas a la identificación de enfermedades en cultivos agrícolas [6], que demuestran la versatilidad de los métodos de inteligencia artificial en dominios tan diversos como la agricultura de precisión y la protección ambiental [7].

En el ámbito de las *telecomunicaciones* y redes de vehículos aéreos no tripulados ha analizado algoritmos de enrutamiento dinámico y adaptativo para este tipo de redes, integrando mecanismos de toma de decisiones basados en inteligencia artificial que permiten una selección óptima de rutas [8]; esto último supone un avance significativo en el diseño de sistemas autónomos y en el desarrollo de protocolos de enrutamiento inteligentes.

En este punto, me gustaría hacer una especial mención a la participación del Dr. Kerkeb en un equipo multidisciplinar liderado por dos investigadores de la Universidad de Granada, los profesores Encarnación Hidalgo-Tenorio y Juan Luis Castro Peña. Dicho proyecto, titulado “*Nutcracker*: Detección, rastreo, monitorización y análisis del discurso terrorista en línea”, financiado por el Ministerio de Economía y Competitividad del Gobierno de España durante el período 2016.2023. En su desarrollo se emplearon técnicas avanzadas de inteligencia artificial y procesamiento del lenguaje natural analizar el discurso terrorista en entornos

digitales. Los resultados obtenidos evidenciaron aplicaciones de alto nivel de la inteligencia artificial en el ámbito de la ciberseguridad y la detección de amenazas, un campo de creciente relevancia estratégica.

El Dr. Kerkeb ha sido también un impulsor decidido de la *transferencia de conocimiento y la innovación tecnológica*. Ha diseñado e implementado el Plan Estratégico de Sistemas de Información de la Universidad Ibn Tofail, incorporando sistemas de apoyo a la decisión basados en inteligencia artificial, analítica de datos, y automatización inteligente para la gobernanza y gestión universitarias. Asimismo, ha establecido y diseñado programas integrales de formación en inteligencia artificial para el Centro Code 212, enfocados en inteligencia artificial, programación y robótica, contribuyendo al desarrollo de competencias técnicas del alumnado universitario en tecnologías emergentes.

Su compromiso con la *cooperación científica internacional* se ha materializado en numerosos proyectos bilaterales y multilaterales. Ha liderado proyectos de cooperación entre Marruecos y España en fisicoquímica, y procesamiento de imágenes y técnicas de caracterización entre 1999 y 2003. Ha desarrollado y gestionado el sistema SIMARECH (Sistema de Información de Investigación Científica) en sus cuatro versiones sucesivas desde 2009 hasta 2025, que es una plataforma que potencialmente incorpora algoritmos de aprendizaje automático para la gestión de la investigación, analítica de datos y apoyo a la toma de decisiones.

Esta trayectoria académica e institucional esbozada necesariamente de forma resumida nos permite comprender la autoridad científica con la que el Dr. Kerkeb aborda el tema de su discurso de ingreso. No estamos ante un teórico alejado de la práctica; Larbi Kerkeb entiende bien la repercusión, el impacto de la ciencia en el bienestar de la comunidad; es consciente de que las desigualdades se pueden superar si la ciencia se pone al servicio de la sociedad. Por ello nos encontramos ante un investigador que ha desarrollado sistemas de inteligencia artificial operativos, que ha liderado proyectos de investigación de gran envergadura, y que ha reflexionado profundamente sobre las implicaciones científicas, tecnológicas y sociales de la inteligencia artificial, en su compromiso con el cambio real que la ciencia puede generar allí donde se implemente.

El tema elegido por el Dr. Kerkeb para su discurso de ingreso no podría ser más oportuno ni más relevante, tanto para nuestra Academia como para la sociedad en su conjunto. A continuación paso a explicar por qué considero que esta elección es particularmente acertada.

En primer lugar, el tema se sitúa en la intersección de múltiples disciplinas científicas: la informática, la neurociencia, la psicología, la filosofía de la mente, la ética y las ciencias sociales. Esta naturaleza interdisciplinar lo hace especialmente apropiado para una Academia de Ciencias que, como la nuestra, agrupa a investigadoras e investigadores de diversos campos del conocimiento.

En segundo lugar, el tema aborda una de las cuestiones científicas y tecnológicas más desafiantes de nuestro tiempo. La inteligencia artificial está transformando rápidamente múltiples aspectos de nuestra sociedad, desde la economía y el trabajo, hasta la educación, la salud, la seguridad y las relaciones sociales. Comprender sus capacidades y limitaciones, especialmente en lo que respecta a dimensiones tan fundamentalmente humanas como las emociones, es esencial para que podamos guiar el desarrollo de estas tecnologías, de manera que beneficien a la humanidad.

En tercer lugar, el tema tiene profundas implicaciones filosóficas y antropológicas. La pregunta sobre si las máquinas pueden sentir nos obliga a reflexionar sobre qué significa sentir, sobre la naturaleza de la experiencia subjetiva, sobre la relación entre mente y cuerpo, y, sobre todo, qué es lo que nos hace específicamente humanos. Estas son cuestiones perennes de la filosofía que adquieren nueva urgencia y nuevas dimensiones en la era de la inteligencia artificial.

En cuarto lugar, el tema tiene importantes implicaciones éticas y sociales, que afectan a toda la ciudadanía. El desarrollo de sistemas de inteligencia artificial emocional plantea cuestiones sobre privacidad, manipulación, autenticidad de las relaciones y el futuro de la interacción humano-máquina. Estas cuestiones no pueden ser resueltas únicamente por expertos técnicos, sino que requieren un debate social amplio e informado. Las academias científicas tienen la responsabilidad de contribuir a este debate proporcionando análisis rigurosos y equilibrados que ayuden a la sociedad a tomar decisiones sobre el desarrollo y la regulación de estas tecnologías.

En quinto lugar, el tema es de particular relevancia para Granada y su Universidad, poseedora ya de una tradición científica sólida y consolidada en este ámbito. La incorporación del Dr. Kerkeb a nuestra Academia, y sus aportaciones sobre las dimensiones afectivas de la inteligencia artificial, fortalecerán sin duda la capacidad de nuestra institución para contribuir al desarrollo de este campo de tanto interés.

Finalmente, el tema es relevante para el futuro de la investigación científica y la educación superior. La inteligencia artificial está transformando la manera en que se realiza la investigación en múltiples disciplinas, desde las ciencias naturales hasta

las humanidades. Los sistemas de inteligencia artificial que incorporan procesamiento afectivo pueden contribuir a mejorar la educación, la tutoría y el apoyo al alumnado. Las universidades y las academias científicas deben comprender estas tecnologías para poder aprovechar sus potencialidades y mitigar sus riesgos.

Por todas estas razones, considero que la decisión de abordar en su discurso de ingreso la relación entre inteligencia artificial y sentimientos es extraordinariamente acertada. Este tema, repito, nos desafía intelectualmente, nos obliga a integrar perspectivas de múltiples disciplinas, tiene implicaciones prácticas importantes, y nos invita también a reflexionar sobre cuestiones fundamentales sobre la naturaleza de la mente, la conciencia y la experiencia humana.

Y acabo. Estoy convencido de que la incorporación a esta Academia del Dr. Kerkeb permitirá que se intensifique la ya fructífera colaboración existente y servirá para favorecer el enriquecimiento mutuo. Su compromiso con la excelencia académica y con el servicio a la sociedad resuena con los valores fundamentales de nuestra Academia. Su experiencia en gestión universitaria y en cooperación científica internacional será muy valiosa para nuestras iniciativas institucionales.

Por todo ello, en nombre de todos los miembros de esta corporación, no me queda más que dar la más cordial bienvenida al profesor Larbi Kerkeb. Querido Larbi, compañero, amigo, hermano, cuentas con el apoyo, el respeto y el afecto de todos tus colegas académicos aquí presentes. Éste, así lo deseo, es el principio de un largo recorrido que anhelaba hacía ya mucho tiempo. Tu familia, tus amigos, tu universidad hoy deben de sentirse muy orgullosos. He dicho.

Muchas gracias.