

2019 KNS 추계학술발표회

인공지능시스템 소프트웨어 확인 및 검증 방안



2019. 10. 23.

**한국원자력연구원
계측제어인간공학연구부**

구 서 룡



CONTENTS



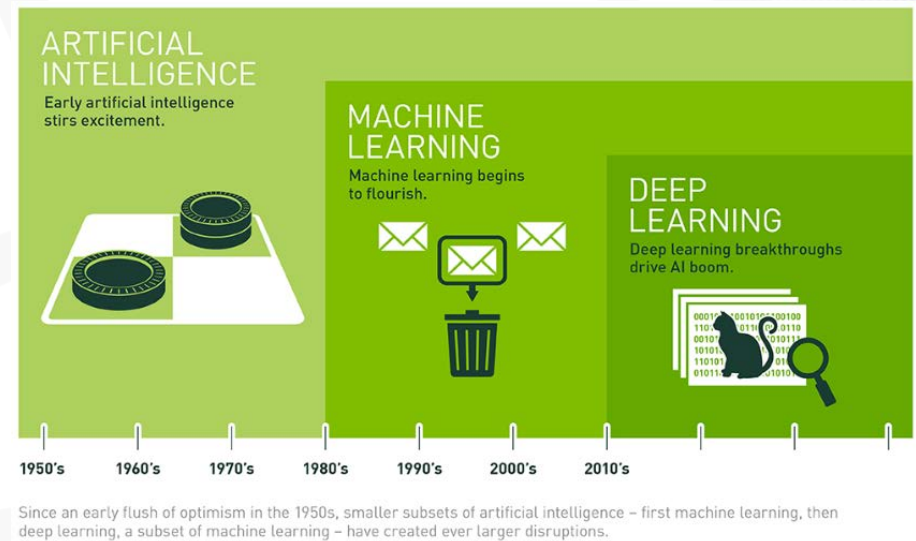
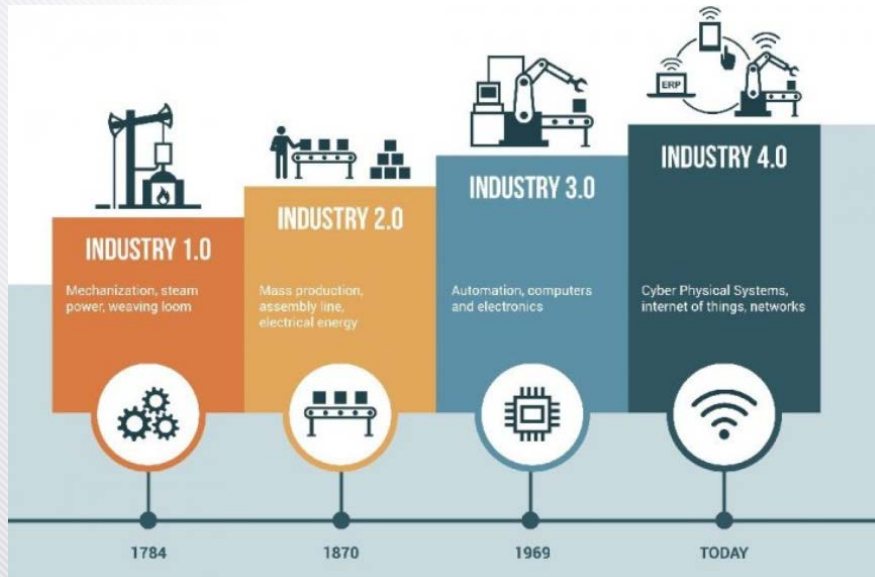
- 01 연구 배경 >
- 02 연구 개발 필요성 >
- 03 연구 개발 목표 및 내용 >
- 04 연구 개발 추진 체계 >
- 05 활용방안 및 기대효과 >

01 연구 배경

01 연구 배경

» 4차 산업혁명과 AI 기반의 새로운 시대

- 4차 산업혁명은 2016년 세계 경제 포럼에서 주창된 정보통신(ICT) 기술의 융합으로 이루어진 혁명 시대를 의미함
- 1950년대 정립된 AI는 1980년대까지의 두 차례의 암흑기에도 불구하고 컴퓨팅 파워 증가와 머신러닝/딥러닝 알고리즘의 진화로 AI 부흥기가 도래함

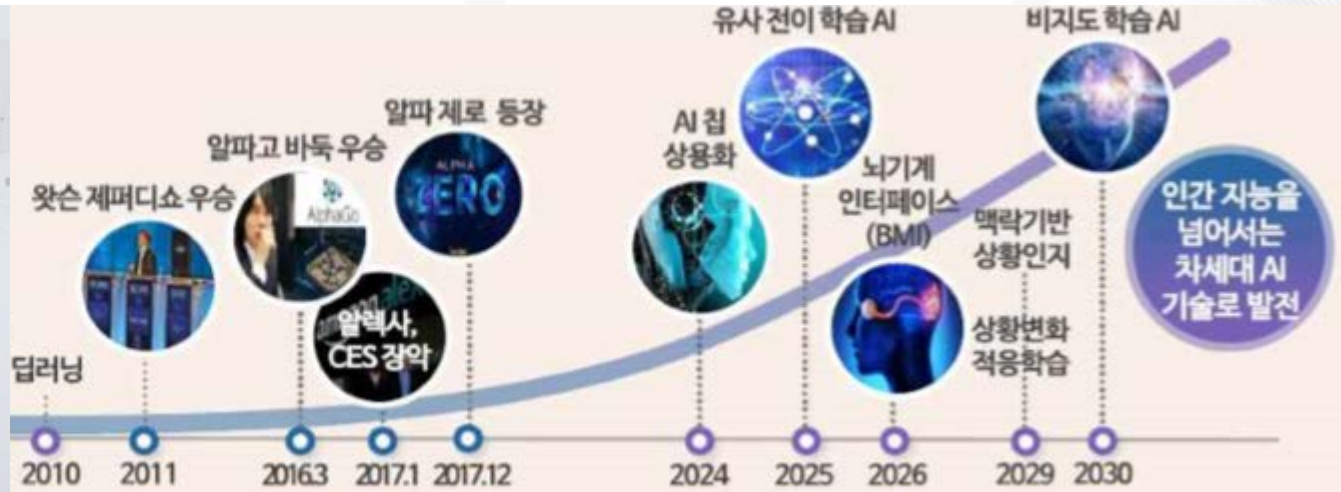


01 연구 배경

» 우리나라의 I-KOREA 4.0

- 세계적인 추세에 맞춰 과학기술과 ICT로 열어가는 사람 중심의 4차 산업혁명 I-KOREA 4.0 비전을 수립함
- 인공지능 기술의 발전은 2024년 AI칩 상용화를 시작으로 2030년부터는 인간 지능을 넘어서는 차세대 AI 기술로 발전할 것이라고 전망함

출처: I-KOREA 4.0 실현을 위한 인공지능(AI) R&D 전략, 과기정통부 (2018.05.)



01 연구 배경

» 미래원자력기술 활용 및 수출 경쟁력 확보

- 정부의 탈원전 기조 속에서도 가동원전 안전 확보 및 해체분야 연구와 방사선 등 타 분야 융합연구에 AI, 빅데이터 등 첨단 ICT기반 기술의 접목이 요구됨
- 중소형 원전 SMART 또는 미래형 원전에서 ICT기반 원자력 안전 혁신 분야에 적극 도입하여 원천적인 원전사고 방지 및 경쟁력 확보 가능함



미래원자력기술 발전전략 하반기 신규과제 <small>*자료: 한국연구재단</small>	
	신규 연구과제
원전 안전 기술	<ul style="list-style-type: none">원자력시스템 리스크 평가·관리원자력사고 신속대응 로봇지능형 원전 자율운전시스템빅데이터·IoT기반 운전·고장 감시예방시스템
방사선 기술	<ul style="list-style-type: none">치매 등 중증질환 조기진단X선 등 방사선의료기기 국산화AI 방사선 반응 예측시스템방사선 중립 3차원(D) 프린팅 소재
타 분야 융합연구	<ul style="list-style-type: none">우주선 동력원용 원자력 배터리원자력추진 우주선우주기지용 소형원자로동위원소 기반 장수명 전력시스템

02 연구 개발 필요성

02 연구 개발 필요성

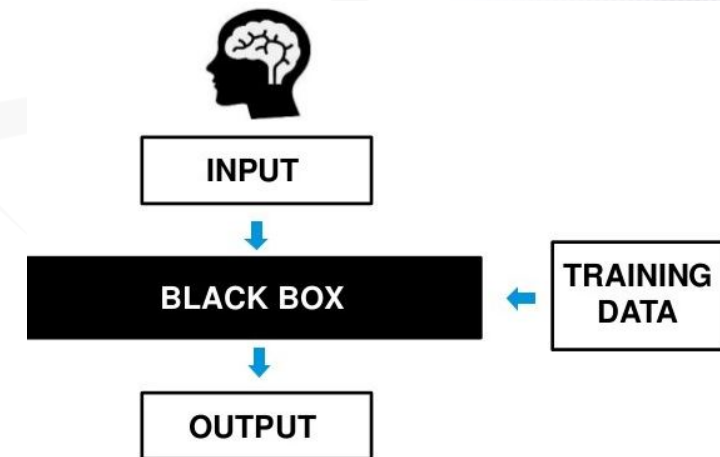
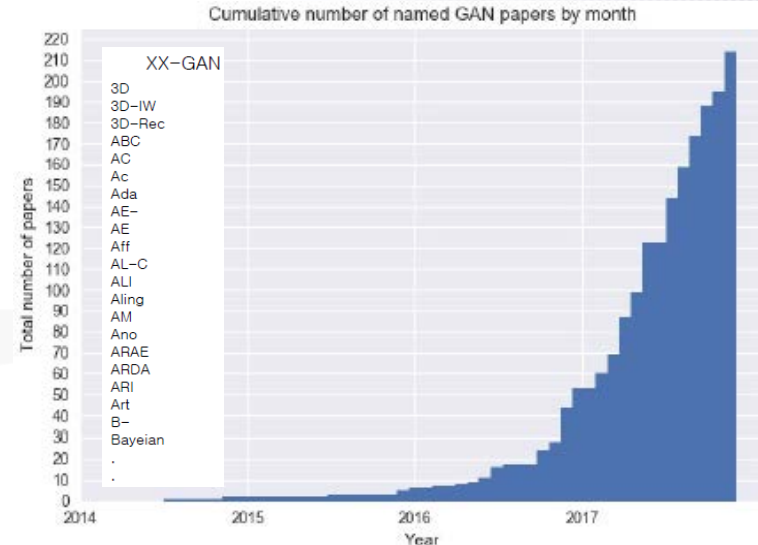
» AI 알고리즘의 특성

- Archive, GitHub 등 클라우드 공간에서의 알고리즘 오픈소스 공유 문화 형성됨
- 딥러닝 알고리즘(ex. GAN) 논문 수 폭발적 증가하고 있음
- 이미지 프로세싱 정확성 확인에 집중됨

» AI 알고리즘의 단점

- 생성된 모델(블랙박스)을 인간이 해석하거나 개선 어려움 – **비결정론적**
- 학습 데이터 범위 외 패턴에 대해서는 제대로 해석하지 못함 – **과적합 문제**

<https://deepphant.in/the-gan-zoo-79597dc8c347>



02 연구 개발 필요성

» 인공지능 자율 주행 자동차 기술발전과 안전성

- 윤리적인 문제에도 자율 주행 자동차에 인공지능 기술 발전은 지속됨
- 자율 주행 자동차의 안전성/신뢰성/품질확보를 위한 **새로운 기술표준 수립이 시급하다고 판단함**



인공지능 제어 차량의 기능 안전성: ISO 26262가 아니라면 대안은?

JOSEPH DAILEY, 글로벌 기능 안전성 책임자, MENTOR, A SIEMENS BUSINESS

초록

ISO 26262 국제 기능 안전 표준은 2011년에 처음 정립된 이후 자율주행차량의 자동차용 전기 및/또는 전자(E/E) 시스템 안전성을 최적화할 방법을 모색하는 엔지니어들에게 확실한 기준으로 삼을 수 있는 표준으로 급부상해왔습니다. 하지만 ISO 26262는 아키텍처와 아키텍처가 안전에 미치는 영향을 분석하는 데 엄격한 기준을 준수하도록 요구합니다. 따라서 자율주행 차량에서 주행 관련 중요한 의사 결정을 기계 학습에 의존하는 형태로 기술 통합이 이전함에 따라 새로운 개념을 충족하기 위해 필요한 요구사항과 기능 안전성 사이의 직접적인 연결 고리로써 이 표준의 역할이 약화되고 있습니다. 이 글에서는 ISO 26262가 정립되기까지 차량 기능 안전 표준이 어떻게 발전되어 왔는지 간략하게 역사적 배경을 소개하고, 이 표준이 인공지능(AI) 제어 차량과 관련하여 어떤 면에서 결함이 있는지에 대한 요소를 살펴보고자 합니다. 특히 완전 자동화된 미래를 앞두고 기존 표준 기반에서 직면하고 있는 여러 가지 문제에 대해 집중적으로 다룰 예정입니다. ISO 26262에서는 E/E 시스템 오류가 존재하는 경우 안전상의 위험 요소를 제거하기 위한 요구 사항을 규정하고 있지만, 이 글에서는 AI 시대에 걸맞은 새로운 기준에서 안전기능(so-called safety of the intended functionality, SOTIF)을 다루어야만 하는 이유를 설명합니다. 이는 첨단 자동차 기능을 차량에 적용할 때 결함이 없더라도 안전상의 위험 요소를 피하도록 검증하는 데 유익한 일입니다. 이 글에서는 현재 개발 중인 세계 최초의 SOTIF 표준인 ISO/WD PAS 21448 작성 프로젝트에 참여한 경험을 바탕으로 실제 사례를 들어보겠습니다.

본 글자 주 이 글은 **임베디드 월드(embedded world) 2018**에서 발표되었으며 Mentor.com의 약해를 받아 재공하는 것입니다.

Mentor
A Siemens Business

A U T O M O T I V E W H I T E P A P E R

www.mentor.com

02 연구 개발 필요성

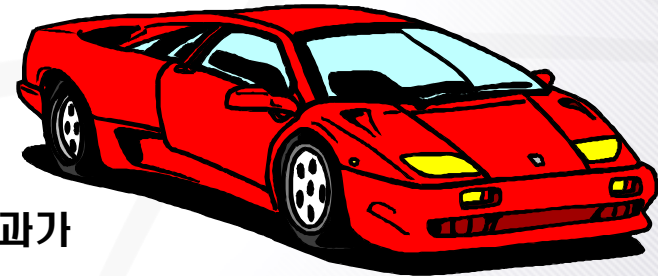
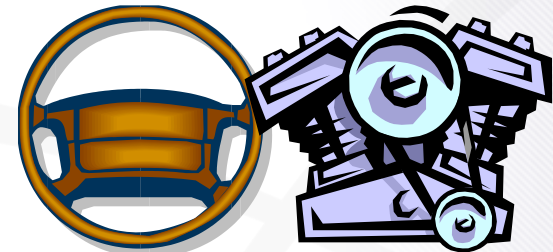
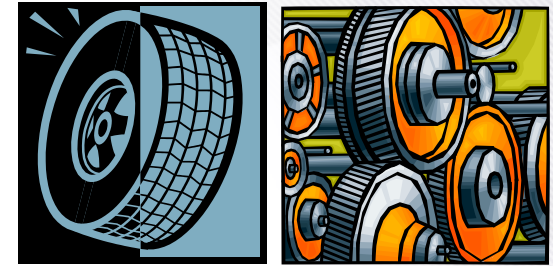
» 소프트웨어 확인 및 검증의 중요성 (1/2)

➤ 소프트웨어 확인 (Verification)

- 소프트웨어 요구사항 및 소프트웨어 설계의 적정성을 평가하는 과정
- 현 단계의 산출물이 바로 이전 단계의 산출물과 일치하는가 확인하는 과정
- Are we building the product right?

➤ 소프트웨어 검증 (Validation)

- 소프트웨어가 지정된 기능을 정확히 수행하는가 여부를 밝히는 활동
- 사용자가 원하는 제품을 만들었는지, 요구사항 분석의 결과가 사용자의 요구를 만족하는지를 검증하는 활동
- Did we build the right product?



02 연구 개발 필요성

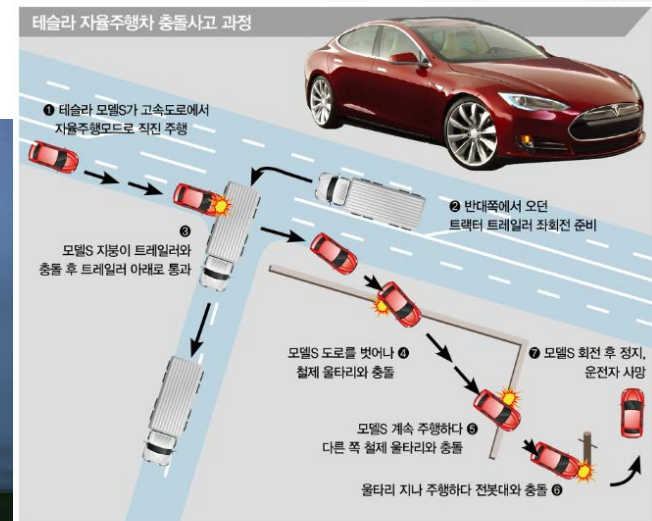
» 소프트웨어 확인 및 검증의 중요성 [2/2]

- 시스템의 작동을 결정하기 위한 논리 수행을 소프트웨어로 처리하는 의존성이 빠르게 증가하고 있음
- 시 기반 시스템은 대부분의 시스템 기능 수행을 소프트웨어로 처리하고 있어 소프트웨어의 오동작은 치명적인 손실을 야기함

Therac 25



Arian-5 Explosion



자료: 로이터·NHTSA

03 연구 개발 목표 및 내용

03 연구 개발 목표 및 내용

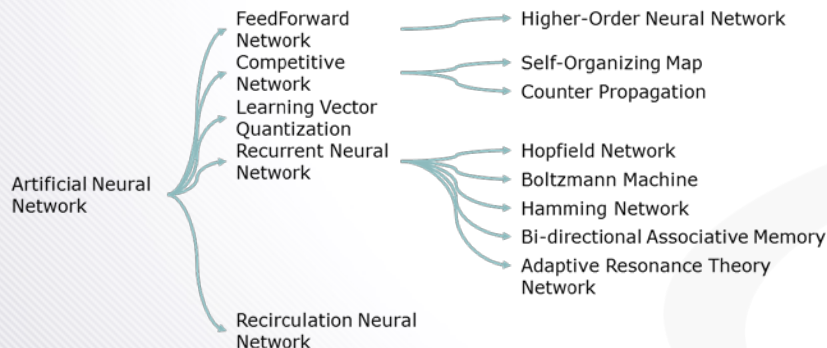
» 연구개발의 최종목표

구분	내용
최종 목표	<ul style="list-style-type: none"> 원전 적용 인공지능 소프트웨어 안전성 확보를 위한 확인 및 검증 방법론 개발
세부 목표	<ul style="list-style-type: none"> 인공지능 소프트웨어 확인 및 검증 적용 가능성 분석 <ul style="list-style-type: none"> 원전 안전성 향상 목적으로 원전에 적용 가능한 인공지능 알고리즘 분류 및 특성 분석 기존 디지털 기반 원전 소프트웨어 확인 및 검증 방법론의 수정 적용 가능성 분석 원전 인공지능 소프트웨어 확인 및 검증 방법론 수립 <ul style="list-style-type: none"> 원전 적용 인공지능 소프트웨어에 특화된 확인 및 검증 요소 도출 원전 적용 인공지능 소프트웨어 데이터 건전성 확인 및 검증 방안 수립 원전 적용 인공지능시스템 소프트웨어에 대한 통합 확인 및 검증 절차 개발 원전 인공지능 소프트웨어 시험 방법론 개발 <ul style="list-style-type: none"> 인공지능 소프트웨어 검증을 위한 단계별 소프트웨어 시험 설계 기술 및 Gray-Box 테스트 기법 개발 인공지능 소프트웨어 검증을 위한 시험 설계 기술 및 Gray-Box 테스트 기법의 적용성 검토 원전에 적용 가능한 인공지능시스템에 개발된 확인 및 검증 방법론 적용

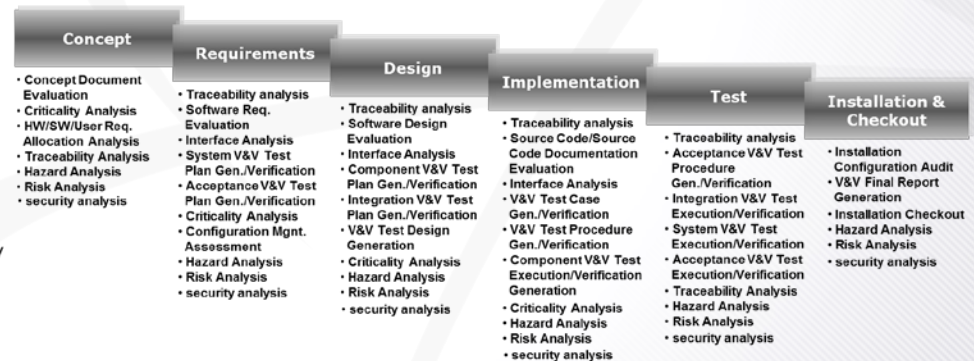
03 연구 개발 목표 및 내용

» 인공지능 소프트웨어 확인 및 검증 적용 가능성 분석

- 인공 신경망 속성 및 응용에 따른 분류 및 AI 알고리즘 별 특성 분석
- 원전에 적용 가능한 AI 기술 및 알고리즘 분류
- 기존 디지털 기반 원전 소프트웨어 확인 및 검증 방법론 개선 또는 수정사항
도출



출처: Methods and Procedures for the Verification and Validation of Artificial Neural Networks (NASA, 2006)



출처: IEEE Standard for Software Verification and Validation

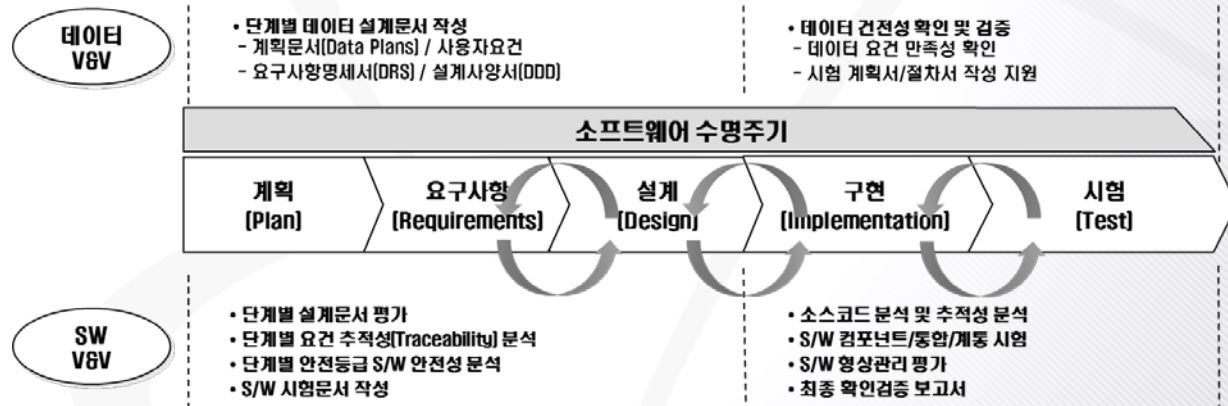
03 연구 개발 목표 및 내용

» 원전 인공지능 소프트웨어 확인 및 검증 방법론 수립

- 원전 적용 인공지능 소프트웨어 대한 코드 요건 기반 검증 속성 도출
- 인공지능 **학습용 데이터 요건/설계 사양 도출** 및 V&V 방안 수립
- 소프트웨어 수명주기 별 원전 적용 인공지능 소프트웨어 통합 V&V 상세 절차 수립

요건/설계 문서 검증				
인허가 적합성 검토 (NUREG-0800)		상세 검증 (IEEE 1012)		
기능 특성	공정 특성	추적성	명세 평가	Interface 분석
정확도	완전성	완전성	정확도	정확도
신뢰성	일관성	일관성	완전성	완전성
강인성	정확성	정확성	일관성	일관성
안전성	스타일		정확성	정확성
보안성	추적성		판독성	시험성
타이밍	검증성		시험성	

➔ 검증 특성 별 검증항목(Checklist)을 만족하는지 확인
➔ 정방향 추적성과 역방향 추적성을 모두 확인



03연구 개발 목표 및 내용

» 원전 인공지능 소프트웨어 시험 방법론 개발

- 인공지능 소프트웨어 단계별 시험 분류 및 Gray-Box 테스트 기반 시험 절차 개발
- AI 알고리즘을 이용한 소프트웨어 테스트 케이스 자동생성 기법 개발
- 소프트웨어 테스트 결과 확률론적 평가 및 시험 방법론 적용성 검토

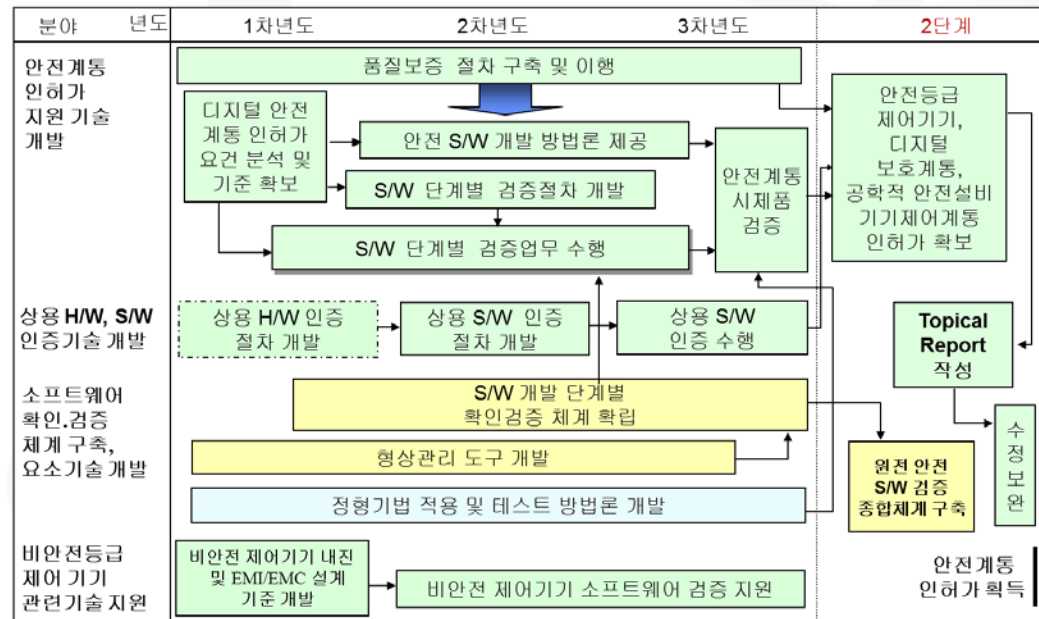
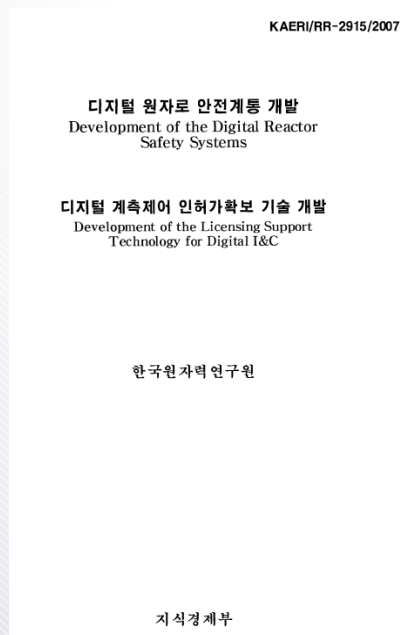


04 연구 개발 추진체계

04 연구 개발 추진체계

» KNICS 국산화 경험 – 디지털 계측제어 인허가 확보 기술 개발

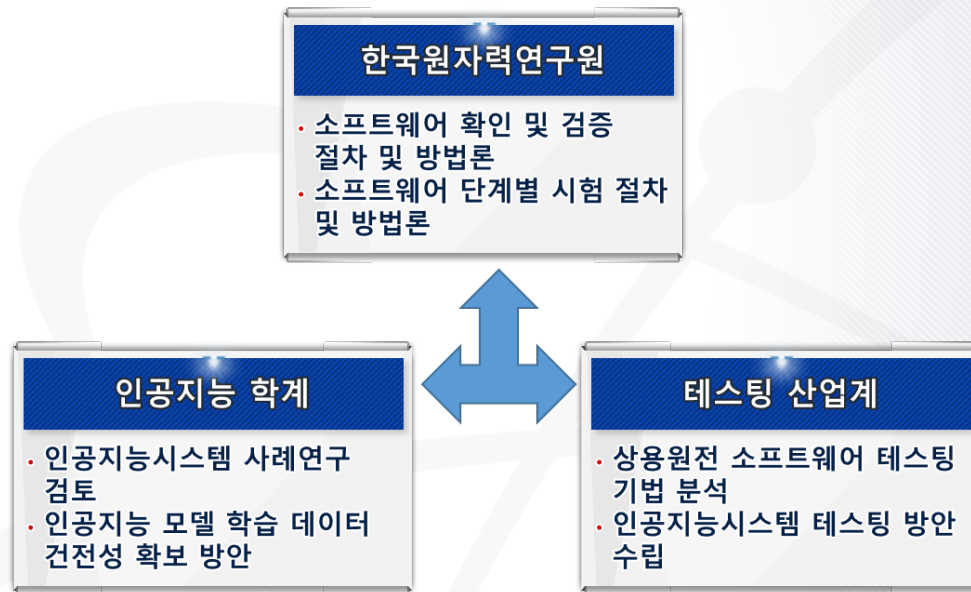
- 2001년부터 7년간 KNICS(원전계측제어시스템) 사업을 통해 **디지털 제어기기 및 보호 계통의 인허가 확보 기술 개발 경험**을 보유
- KNICS 사업에서 개발된 인허가 확보 기술은 신한울1,2호기 및 신고리5,6호기 **적용 및 입증**



04 연구 개발 추진체계

» 기술협력 기반 체계 구축 – 현실성 있는 방법론 수립

- KNICS 사업부터 APR1400 건설 원전 MMIS 소프트웨어 안전성 분석 및 확인 검증 수행 경험을 보유하고 있는 **산/학/연 기술협력 연계 체계** 구축
- 원전 인공지능 소프트웨어 확인 및 검증 방법론의 구현 및 인허가 가능성을 제고하고 지원 **도구의 상용화를 위한 전문기업의 기술 협력**



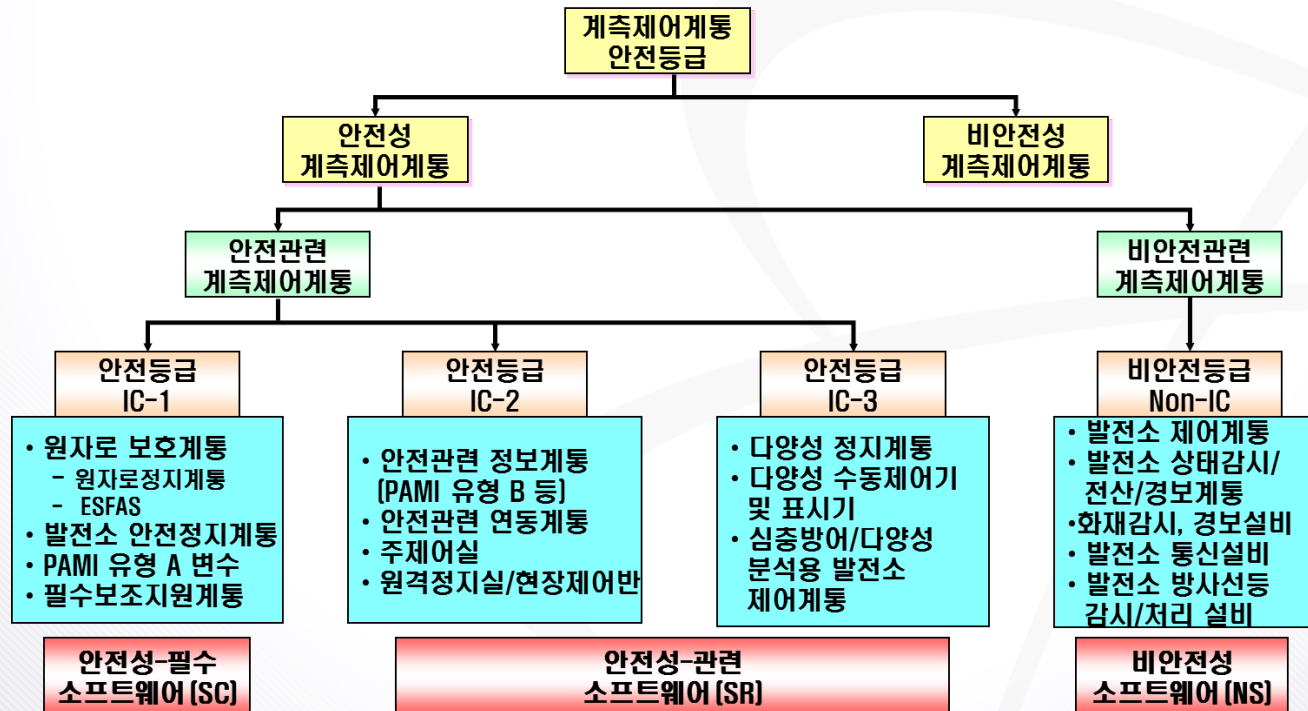
05

활용방안 및 기대효과

05 활용방안 및 기대효과

» 가동원전 안전성 향상을 위한 계측제어계통 점진적 적용

- 인허가 사항과 직접적으로 연관되지 않는 **비안전성 계측제어계통의 적용을 시작으로 안전성 계측제어계통으로 적용을 진행함**
- 안전성 계측제어계통 적용을 위해서 관련된 **인허가 규제 기술 개발을 지원함**



05 활용방안 및 기대효과





THANK YOU