
**데이터의 축적과 활용을 통한
디지털 소재혁신 강화 실행계획(안)
- “DM(Digital Material)” 프로젝트 -**

2021. 3. 15.

산업통상자원부

목 차

I. 소재혁신과 데이터	1
II. 선진국의 움직임과 우리의 현실 진단 ...	3
III. 소재혁신을 위한 데이터 전략방향	6
IV. 주요 추진 과제	7
V. 추진 일정	19

요 약

DIGITAL MATERIAL 프로젝트

데이터의 축적과 활용을 통한 디지털 소재혁신 강화 실행계획

소부장 미래 경쟁력 선점 + 글로벌 소재 경쟁력 우위 확보

- 핵심 신소재 분야의 8대 프로젝트 중심 데이터 개발 플랫폼
- 데이터의 축적과 활용을 위한 산학연 “연대와 협력”
- 소재 개발 소요 비용과 시간의 획기적 단축

추진전략

01 8개 핵심소재 프로젝트 선정 및 10만건 이상 데이터 확보(~'22년)

02 데이터표준화 중점추진 + 공공부문 우선 → 민간확산 등 단계적 추진

03 관련 기관간 강력한 협력 추진체계를 가동하고 주기적으로 점검

주요 추진과제

 8대 선도 프로젝트 추진 3대 분야 8대 선도 프로젝트 선정·추진	<ul style="list-style-type: none">• 4대 소재 기관, 산학연 전문가 검토• 공공연구기관 중심 컨소시엄 구성
 데이터 표준화 표준입력 템플릿 확보	<ul style="list-style-type: none">• 표준화 위원회 가동(용어, 단위 표준화 포함)• DB구조 표준화로 연구기관 간 데이터 연계, 호환
 데이터 축적과 활용 데이터 수집·저장, AI모델 개발	<ul style="list-style-type: none">• 소재정보은행 재규격화, 통합 데이터 관리시스템 구축• 웹기반 AI 서비스 제공(소재물성예측, 역설계)
 데이터 기반 확충 선순환 구조의 데이터 활용 제도 마련	<ul style="list-style-type: none">• 바우처 등 인센티브제 시행, 데이터 관련 R&D제도 개선 등• 공공-민간 오픈 플랫폼 구성, 산업지능화 펀드 지원
 소재데이터 Governance 가동 소재부품장비 정보협의회 구성·운영	<ul style="list-style-type: none">• 중장기 전략과 연차별 추진 계획 수립• 분과협의회를 통한 구체적 이행방안 마련

I. 소재혁신과 데이터

◇ 4차 산업혁명에 대응하고 친환경·고부가 산업으로 전환하기 위해, 디지털데이터를 활용한 혁신적 소재부품 기술 확보 방식 필요

◇ 소재개발 사이클을 단축시킬 수 있는 데이터 기반 개발 방식 부각

- 제품의 수명주기가 짧아지고, 개발 속도 경쟁이 치열해짐에 따라, 주관적 경험과 암묵지(暗黙知)에 기초한 전통적 개발 방식 한계 부각
 - 선형(Linear)적 개발 방식으로 소재개발 과정이 데이터 형태로 저장되지 않고 시행착오가 반복적으로 발생
- 디지털 방식은 데이터를 활용하여 모든 개발과정의 환류가 가능한 선순환(Circular) 방식으로, 시간과 비용의 획기적 감축 가능

◇ 4차 산업혁명 등 소재 혁신을 위한 소재 디지털화 전략 필요

- 소재·부품·장비는 산업에 필수적인 '보이지 않는 기술속의 기술'로, 제조업의 회복력(Resilience), 공급망 安保와 미래 경쟁력 결정
 - 소재는 산업 밸류체인의 첫 단계에 투입되어 제품의 기능과 품질 좌우
- 4차 산업혁명 기술 활용과 저탄소·고부가 구조 전환을 위해 新 소재 경쟁이 격화되고, 디지털화를 통한 혁신적 개발 방식 대두
 - * 산업 디지털 전환 촉진법 제정 등 데이터·AI를 활용한 산업데이터 전략과도 부합

◇ 선진국 중심으로 디지털 소재개발 플랫폼 확보 경쟁 본격화

- 美·日 등 소재분야 선도국들은 미래산업 주도권 유지와 강화를 위해 디지털 소재개발 프로그램을 전략적으로 강력 추진 중
 - * (美) Material Genome Initiative('11), (日) 新원소전략('12), 머리티얼 혁신 강화전략('20)

☞ 디지털데이터 기반의 '디지털 소재혁신 강화' 본격 가동
→ 소재 분야의 디지털 대전환을 위해 "Digital Material" 프로젝트 추진

- (개념) ①기초소재, 혼합소재, ②원료조성, 공정방식, 측정조건 등에 따라 다양하게 조합·생성되는 데이터를 디지털化 → 신소재 개발에 활용

< 전통적 소재개발과의 차이점 >

	As-Is		To-Be
데이터	암묵지(暗黙知)	⇒	객관적
개발방식	실험반복	⇒	데이터 기반 물성 예측
특징	선형(Linear) 모델	⇒	순순환(Circular) 모델
Cost	장시간·비용多	⇒	개발 Cycle 단축

- (기본 구조) 디지털 데이터, 컴퓨팅 처리기반, 제도·체계로 구성

3大 구성요소

세부 구성요소

Digital Data	<ul style="list-style-type: none"> - 데이터 표준화 : 구성항목(X변수, Y변수) - 데이터 축적 : 수집방식(실험데이터, 계산과학) - 데이터 활용 : 인공지능 알고리즘(머신러닝, 딥러닝)
컴퓨팅 처리기반	<ul style="list-style-type: none"> - H/W : 대규모 데이터 처리 및 연산능력 (GPU 등) - S/W : PyTorch(페이스북), TensorFlow(구글) 등
제도·체계	<ul style="list-style-type: none"> - 데이터 축적 활성화 인센티브 제도 - 데이터 보호를 위한 개방성, 보안체계 등

- (데이터 표준화) 데이터 입력값, 산출값 등 구성항목을 표준화
- (데이터 축적) 직접 실험과 계산과학*을 통해 데이터를 확보할 수 있으며, 실험이 불가능한 영역에 계산과학 데이터를 보완
 - * 과학적 이론식을 기반으로 수학적 컴퓨팅을 이용, 분자나 원자들의 물성을 예측 방법
- (데이터 활용) DB의 고도화 수준에 따라 패턴 분석 → 기계학습 (Machine Learning), 딥러닝(Deep Learning, Data-driven AI) 등으로 진화
 - * 머신러닝 : 입력값(X)과 산출값(Y)사이의 함수 관계를 기계가 스스로 파악하는 기술
 - ** 딥러닝 : 다양한 머신러닝 알고리즘(결정트리, 선형회귀, 베이지안, 딥러닝 등)중 하나로, 인간의 신경망과 유사한 형태를 가진 기술

II. 선진국의 움직임과 우리의 현실진단

가. 선진국의 움직임

◇ 양질의 디지털 데이터기반 소재개발로 빠르게 전환 → **초격차 확보**



미국

구체적 프로젝트 중심 소재개발, 계산과학 방식 데이터 수집

- 소재개발의 디지털화를 통한 소재개발 기간·비용을 최소 2배 단축시키기 위한 **MGI(Material Genome Initiative) 전략** 수립('11~, 2.5억달러)
 - ①R&D 방식 변화(데이터 중심), ②계산과학 기반으로 실험데이터 수집, ③소재 DB 구축(290만건), ④인력양성 등 4대 분야에 집중
 - 경량소재, 고분자 복합체, 에너지저장 및 전자·광학 소재 등 **핵심 프로젝트 중심**으로 디지털 기반 소재개발 진행



일본

데이터 관련 제도 확충 선행, 계산과학과 실험데이터 중시

- 잃어버린 10년을 회복하고 자국의 강점인 소재부품의 글로벌 경쟁력을 강화하기 위한 「**新원소전략**」 등 국가전략 수립('12~, 700억엔)
 - 「**머티리얼 혁신 강화전략**」을 통해 디지털 소재개발을 위한 제도적 기반과 세부 실행 계획 발표 예정('21년 예정)
 - 소재분야 실험·계산 데이터와 **모노즈쿠리 공정기술을 수치화**하여 기술력의 계승·전승을 목적으로 데이터베이스 구축 中



유럽

연구자 중심 프로젝트, 계산과학 방식

- 데이터 공개·공유, 데이터베이스 구축 등 디지털 기반의 소재개발 혁신을 위한 「**Horizon 2020**」 **전략** 수립('14~, 491만 유로)
 - NOMAD 등 계산과학 중심의 DB를 구축 중(440만건)이며, 연구자의 접근성 향상과 데이터 축적에 집중

나. 우리나라 상황진단

◇ 소재데이터 디지털화를 위한 데이터 표준화, 축적규모, 관련 제도 등 전반적 기반이 충분히 갖춰지지 않은 초기 단계

① (데이터 표준화) 데이터 수집을 위한 표준화, 기관간 협력 등 시스템 부족

- 데이터축적을 위한 핵심항목 등 표준화에 대한 논의가 부족하고, R&D 과정에서 발생하는 데이터가 사장(死藏)되고 있는 상황
- 2007년부터 구축한 4대 소재정보은행 기관간 협력체계가 작동되지 않고, 개별기관별로 축적되어 데이터 균질성과 효율 저하
 - * 연구기관간(화연, 재료연, 세기원, 다이텍 등) 표준화 체계없이 개별 운영하고 있어 상호참조, 대체소재 검색 등 곤란

② (데이터 축적) 디지털 분석과 탐색에 필요한 데이터 규모와 수준 미흡

- 소재정보은행에 금속, 화학 등 160만건의 물성정보가 축적되었으나, 유럽 440만건, 미국 290만건에 비해 양적 규모(Critical Mass) 부족
- ①기초소재, 혼합소재, ②원료조성, 공정방식, 측정조건 등 단계에 따라 완결형의 쏠주기 디지털데이터의 축적수준 미진

데이터 축적분야

분야	원료	조성	공정	물성	건수(만건)
금속	철강, 알루미늄합금, 마그네슘합금, 구리소재 등 (인장강도, 연신율, 경도, 탄성계수, 열전도도 등)				28
화학	플라스틱, 고무, 첨가제, 필름, 접착, 도료, 복합재 등 (점도, 전기전도도, 스펙트럼, 스크래치 특성 등)				84
세라믹	반도체/센서소재, 기계구조재료, 생체소재 등 (밀도, 절연저항, 공진주파수, 유전율, 굴절률 등)				32
섬유	PET, 나일론, 아라미드, 탄소섬유, 토목소재 등 (강도, 녹는점, 흡수율, pH, 기공크기, 기포량 등)				13

③ (데이터 활용) 데이터 활용 촉진을 위한 전략방향 설정 필요

- 단계적 데이터 확산을 위한 우선순위 설정, 기관간 역할 분담 및 관련제도 확충 등에 대한 전략설정이 부족
 - * 사회적 가치가 높은 데이터 특성상, 공공부문 → 민간확산 등 우선순위 설정 등 전략중요

- 민간의 디지털 데이터 축적이 가시화되고 있으나, 역량이 미흡한 중소기업의 소재혁신 데이터 공유와 확산을 위한 노력 필요

< 국내 데이터기반 소재개발 사례 >

구분	개발소재	주요 내용
삼성전자('18)	OLED 소재	· AI를 활용하여 36,000종의 화학구조 중 딥블루 OLED 소재에 적합한 2종의 신규 화학구조 발견
LG사이언스파크('20)	친환경축매, 차세대광학소재	· 토탈(프), 토론토대(캐), 맥마스터대(캐) 등과 함께 인공지능 기술을 활용한 화학소재를 개발하는 컨소시엄 결성

다. 시사점

① 양질의 데이터 표준화, 축적을 위한 구체적 전략과 프로젝트 필요

- 데이터 표준화, 축적 → 활용(인공지능 모델 개발) 등 소재혁신 진화단계에 맞게, 양질의 데이터 축적을 위한 단계적 전략 접근
- 단기적으로 모든 소재영역을 대상으로 하기보다는 구체적 타겟 프로젝트를 설정하여 사실상(De Facto) 표준모델화 → 확산 추진

② 소재데이터 활용 및 확산을 위한 제도적 기반 구축 강화

- 데이터의 특성상 초기단계에서 공공부문을 중심으로 추진하고, 중소기업 등 민간주체에 개방을 확대하는 소재혁신 플랫폼화
- 자발적 데이터 축적을 위한 인센티브, 보안 등 제도적 틀 마련

③ 지속적인 추진을 위한 「디지털 소재데이터 Governance」 마련

- 지속적이고 예측가능한 소재혁신정책을 위해 공공·민간이 모두 참여하는 「디지털 소재데이터 Governance」 체계 필요
- 디지털화 정책이슈 발굴 및 대안마련 등을 위해, 소재·공정 도메인 (Domain Knowledge) 전문가와 데이터 전문가의 협력체계 마련

III. 소재혁신을 위한 데이터 전략방향

< 기본 방향 >

◇ 데이터기반 소재개발로 소부장 미래 경쟁력 선점

* 개발 필요성이 높은 수소경제, 탄소중립 등 핵심 新소재개발 분야 우선 추진

◇ 디지털 소재개발 플랫폼 구축으로 글로벌 경쟁우위 확보

추진 전략

① 8대 핵심소재 프로젝트 선정 및 10만건 이상 데이터 확보(~'22년)

② 데이터표준화 중점추진 + 공공부문 우선 → 민간확산 등 단계적 추진

* '20~24년까지 700억원 집중 투입

③ 관련 기관간 강력한 협력 추진체계를 가동하고 주기적으로 점검

주요 추진과제

1. 8대 선도 프로젝트 추진	3대 분야 8대 선도 프로젝트 선정·추진
2. 데이터 표준화	표준화 체계·표준 템플릿 마련
3. 데이터 축적과 활용	데이터 수집·저장, 인공지능 모델개발
4. 데이터 기반 확충	선순환 구조의 데이터 활용 제도 마련
5. 소재데이터 Governance 가동	정보협의회·분과위원회 운영

IV. 주요 추진 과제

< 주요 추진 골격 >

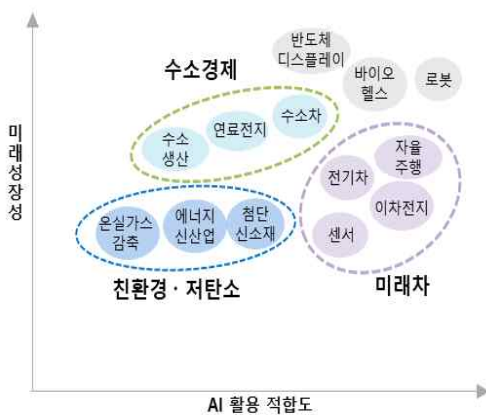


1 8大 선도 프로젝트 추진

◇ **친환경·저탄소 및 성장 유망성 등을 고려, 화학, 금속, 세라믹, 섬유 소재 중심의 프로젝트를 선정·추진**

- **(8大 선도 프로젝트)** 광범위한 분야를 모두 대상으로 하기보다는 구체적인 프로젝트 중심으로 추진하여 프로토타입화
 - (3대 분야) 친환경·저탄소, 성장 유망성 및 데이터 활용 개발가능성 등을 종합적으로 고려하여 선정
 - 데이터 기반 핵심 소재 개발로 소부장 미래 경쟁력을 선점할 '수소경제', '탄소중립', '미래 모빌리티' 등 3대 분야

< 프로젝트 분야 >



3대 분야	세부 프로젝트
수소경제	① 고효율 그린수소 생산용 금속 촉매소재 ② 양방향 연료전지용 All-Ceramic 전극소재
탄소중립	③ 대체원료 활용 올레핀 생산용 촉매소재 ④ 생분해성 고강도 섬유소재 ⑤ 가스터빈 부품용 고엔트로피 합금소재
미래 모빌리티	⑥ 미래모빌리티 내외장용 경량복합수지 ⑦ 미래수송기기용 고내광성 친환경 내장재 ⑧ 미래차 전장용 고신뢰성 커패시터 소재

* 출처: Deloitte, McKinsey, 전문가 서베이 재가공

- **(4대 소재) 화학·금속·세라믹·섬유 등 4대 소재 분야의 공공 연구기관 중심으로 선도 프로젝트 컨소시엄 구성(~'21.4月)**
 - * 기존 데이터를 보유하고 있는 4개 소재기관(화연, 재료연, 세기원, 다이텍) 중심으로 학계, 연구기관, 소부장 테스트베드 기관 등이 참여
 - 특히, 기술융합적 성격을 고려, 기관간 협업 분야를 구체화하여 프로젝트간 연계와 효과성을 제고
 - 향후 진전단계에 따라 민간기업 등으로 참여를 확대하는 방안 검토

⇒ **소부장 인프라 예산을 활용, '22년까지 프로젝트별로 40억원 내외의 예산을 집중 투입(3년간 총 300억원 규모)**

수소경제

① 그린수소 생산용 촉매 고효율 그린수소 생산용 금속 촉매소재

- 현재 그린수소 생산용 물전기분해 시스템은 코발트 계열의 2원계 촉매를 사용하고 있으나, 높은 전류 투입시 열화되어 효율이 낮음
- 고전압·대전류에 강하고 높은 효율(80%이상)로 수소생산이 가능한 코발트-니켈, 코발트-철 등 **3원계 합금 산화물 촉매 소재** 개발
 - * 2원계 코발트 산화물(Co_3O_4) 촉매 → 3원계 합금 산화물(NiCo_2O_4 , FeCo_2O_4) 촉매

② 차세대 연료전지 전극소재 고체산화물 연료전지용 전극소재

- 양방향 SOFC*는 전기 생산(순방향)과 수소 생산(역방향)이 가능하나, 기존 금속·세라믹 혼합 전극은 탄소침착으로 수명과 성능이 저하
 - * SOFC(Solid Oxide Fuel Cell)는 산소와 연료(수소, 메탄 등)를 투입하여 전력 생산
- 금속·세라믹 혼합 전극을 대체할 수 있는 고전류밀도·고안전성의 페로브스카이트계 **All-Ceramic 전극 소재** 개발
 - * 금속·세라믹 혼합전극 → 티탄산란타노이드(LnTiO_3 , Ln=La, Sr, Pr, Nd) 계열

탄소중립

③ 올레핀 생산용 촉매 나프타 대체원료 활용 올레핀 생산용 촉매소재

- 플라스틱 원료인 올레핀은 나프타를 고온($850\sim 1000^\circ\text{C}$) 열분해하여 생산하나 다량의 에너지 소모 및 CO_2 발생(부생가스 메탄 사용)
- 나프타 원료 대체(CO_2 , 수소, 저활용 석유유분 등) 및 나프타 저온 분해 등을 위한 다양한 **촉매소재 및 공정최적화 기술** 개발
 - * 고온 열분해(나프타) → 제올라이트, Metal Carbide 등 촉매 활용 전환(대체원료)

④ 생분해성 플라스틱 생활·산업용 바이오기반 고강도 섬유소재

- 생분해되는 PLA 등은 의류, 일회용품, 장난감 소재 등으로 사용 중이나, 섬유화 시 물성(강도, 내열성) 부족으로 용도가 제한
- 포장재, 보온재, 건축용 자재 등 생활·산업 자재용으로 활용을 확대하기 위해 **강도를 2배이상 높인 생분해 섬유소재 개발**
 - * PLA(옥수수 추출 수지) → PLA + 생분해원료(PBAT), 가소제, 안정제, 결정증진제 등 첨가

⑤ 가스터빈 부품용 합금소재 고내열성 고엔트로피 합금 신소재

- LNG용 가스터빈 부품은 고가의 니켈계 초내열합금을 사용 중이나 터빈 대형화에 따라 저렴하고 내열성이 높은 대체합금 개발 필요
 - * 4~5개 원소가 비슷한 비율(10~20%)로 구성된 금속, 일반적인 금속소재는 주원소가 60% 이상
- 가스터빈의 압축기 블레이드 등 고온에 적용할 수 있는 **저가 원소 조합의 내열 고엔트로피 합금소재** 개발
 - * 고가 합금(니켈, 크롬 70% 이상) → Fe, Al, Cu 등 저가 4~5개 원소 10% 이상 비율

미래 모빌리티

⑥ 미래모빌리티용 경량복합재 항공, 드론, 미래차 등 내외장용 경량복합수지

- 항공기, 드론, 전기차 등 미래모빌리티의 배터리 장착시 무게 증가로 주행·비행 거리 감소 등 효율이 낮아지는 문제 발생
- 기존 무게 대비 10% 이상 경량화와 기계적 물성 20% 이상 향상시킬 수 있는 **고강성 복합수지 소재** 기술 개발
 - * 경험기반 단순 원료 배합 → 필러, 수지, 첨가제 및 고분자 배합과 공정 조건 최적화

⑦ 친환경 내장재 친환경 염색가공 공정 적용 고내광성 내장재

- 친환경 소재를 적용한 자동차 내장재 수요가 확대되고 있으나, 리사이클 PET 등 친환경 섬유소재는 물성과 성능이 미흡
- 환경 규제 대응과 완성차 업체의 요구 조건(내광성 등)을 동시에 만족하는 20%이상 성능이 향상된 **리사이클 섬유소재** 개발
 - * 리사이클PET + 염료, 아세트산 → 리사이클PET + UV-흡수제, 침투제, 분산제 활용 염색조성 최적화

⑧ 미래차용 전자소재 미래차 전장용 고신뢰성 커패시터 소재

- 미래차 모터·배터리 제어, 인공지능 모듈 등 전자장치 증가로 차량용 커패시터 사용이 급증(2만개/대, 테슬라 모델3)
 - * 커패시터는 에너지를 축적하여 전기전자회로에 안정적으로 전기를 공급해주는 '담' 역할
- 전자파 간섭 억제, 배터리 공급 전압 안정 등 미래차 주행 안전성 확보를 위해 **고신뢰성, 고성능 커패시터 적용 유전소재** 개발
 - * 티탄산바륨(BaTiO₃) + Mg, Mn → 티탄산바륨 + Si, Y, Dy 등 원소 추가 및 입도 20% 미세화

2 데이터 표준화

◇ 8대 선도 프로젝트별 데이터 항목·구조·방식 등 표준입력 템플릿 확보

① **(표준화 항목 도출)** 4대 소재의 소재개발 단계별(원료→조성→공정→물성) 포함시켜야 할 표준화 공통 구성항목 도출 (4대 소재기관)

○ 표준화 공통항목을 토대로 4대 소재별 특성에 따라 필요한 세부 항목을 구체화 추진(~'21.6月)

* (例)소재에 따른 공정 차이 : (금속) 단조, 용접 / (화학) 사출, 압출 / (세라믹) 용융, 소결 등

< 소재개발 단계별 표준화 항목 (예시) >

개발단계	공통 구성항목
원료	원료명(원소명), 소재분류, 제조사, 분자식, 화학구조, 스펙트럼, 구입날짜, 작성자, 원료 물성 정보 등
조성	원료(원료명, 소재분류), 투입량, 단위, 기관, 날짜, 작성자 등
공정	공정장비(공정장비사양), 공정설정(공정조건 : 온도, 압력 등 공정 설정값), 공정순서(연속공정 순서정보) 등
물성	시편형태, 측정장비, 측정조건, 측정물성, 평균물성, 거동물성

② **(표준템플릿 구축)** 항목 표준화를 바탕으로 선도 프로젝트별로 표준 템플릿 마련 (표준화 위원회)

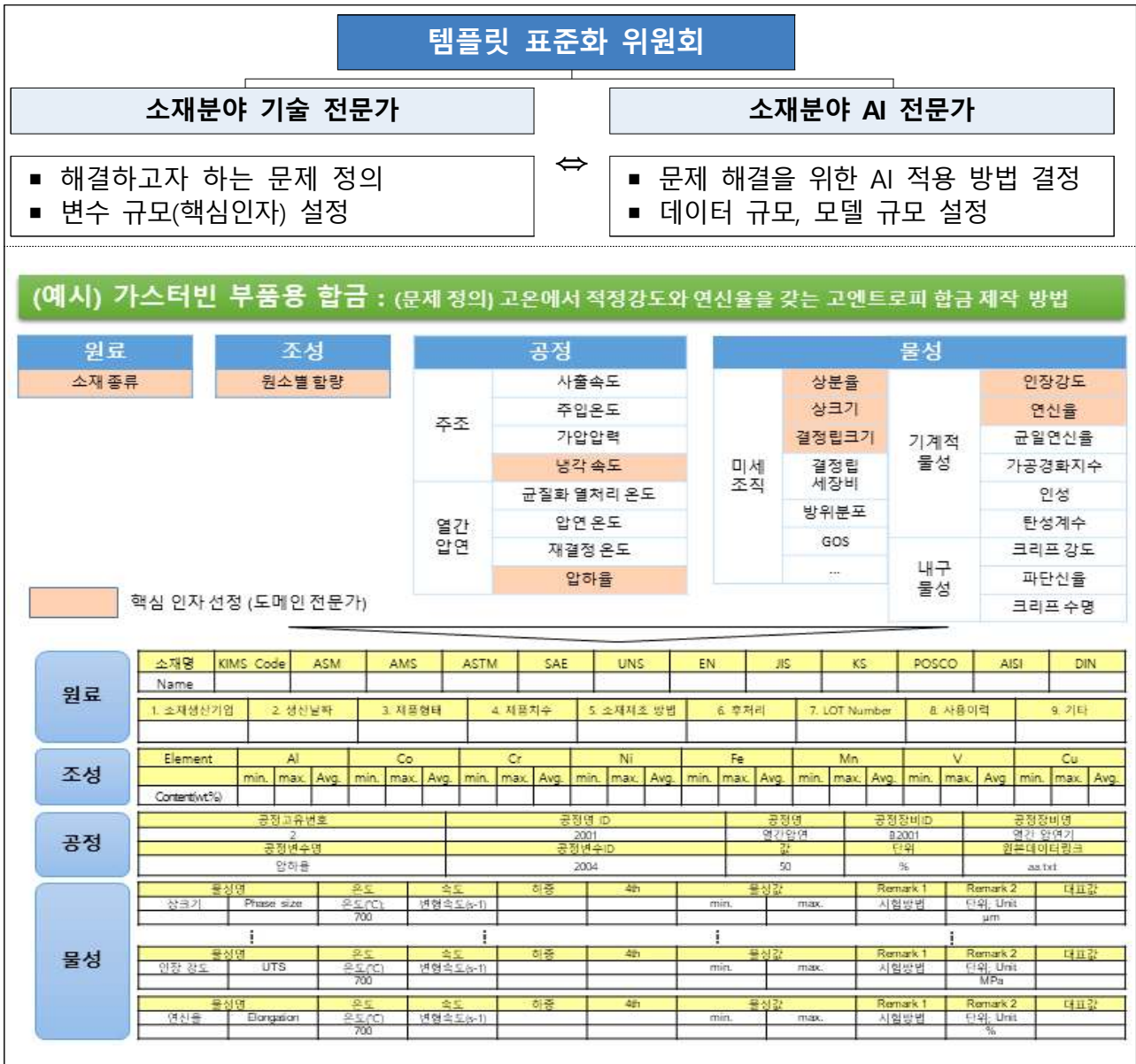
○ 실제 데이터를 생산할 수 있는 능력을 갖춘 해당 도메인 전문가 및 소재 AI 전문가 등 '템플릿 표준화 위원회' 가동('21.7月~)

- 해외 표준화 항목, D/B구조 등을 비교분석하고, 글로벌 수준에 맞는 템플릿 구성

* 소재별 데이터 수준(Level) 보정, 소재특성의 차이 반영 등 방대한 작업으로 전세계적으로도 상세한 소재 데이터 표준화는 미진

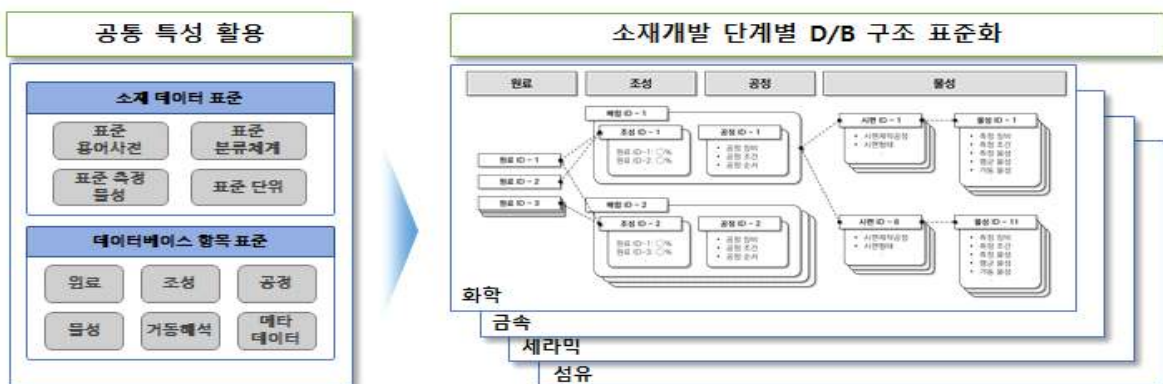
- 기존 데이터를 최대한 변환시켜 활용도를 높이는 한편, 용어, 분류체계, 측정단위 등에 대한 표준화 방안도 함께 마련

< 템플릿 표준화 위원회와 표준 템플릿 구성 예시 >



③ **(D/B 구조 표준화)** D/B구조를 소재개발 단계별로 표준화하여, 연구기관별 개별 데이터의 상호 호환·연계 추진(4대 소재기관, ~'21.9월)

< D/B 구조 표준화 >



3 데이터 축적과 활용

◇ 표준입력 템플릿에 따라 '22년까지 10만건 이상 양질의 데이터를 확보하고 소재개발 AI 예측 모델 개발

가. 데이터 축적

- ① **(공공 데이터 확보)** 데이터 기반 소재개발에 필요한 데이터 규모·수준(Critical Mass)을 위해 3가지 방식으로 추진 (4대 소재기관)
 - ① **(기존 데이터)** 해당 프로젝트 관련된 기존 소재정보은행 데이터를 재규격화 하고, 문헌 자료 및 추가 실험 등을 통해 보완(~'21.12月)
 - * (재규격화) 데이터를 선별·통합·정제·가공하여 전주기 데이터셋으로 개선
 - ② **(실험 데이터)** 프로젝트별 컨소시엄이 직접 실험을 통해 고품질 기계학습이 가능한 대량의 신규 데이터 생성('21.9月~)
 - 4개 소재기관 실험 장비, 32개 소부장 융합 혁신지원단 테스트 베드 등을 활용하여 표준입력 템플릿에 따라 데이터 수집
 - * 센서, IoT 등 자동 계측 장치를 활용한 실시간 데이터 수집체계 구축
 - 연구기관 간 데이터의 상호 교차검증을 통해 데이터 신뢰성 확보
 - * (例) 섬유 : 다이텍연구원(4대 소재기관) - 한국섬유소재연구원(테스트베드 구축기관)
 - 필요시 해당 프로젝트 관련 국내외 전문 연구기관을 탐색하여 실험 데이터 생산을 의뢰하고 구매하는 방식도 추진
 - ③ **(계산 데이터)** 실험조건에서 구현하기 어려운 '공백 데이터'를 확보하기 위해 계산 프로그램을 활용하여 추가적인 데이터 생성('22~)
 - VASP(제일원리*) 등 상용 프로그램 사용, 다량의 정제된 데이터 생성
 - * 제일원리(First Principle) : 양자역학 기반 기본원리를 바탕으로 구조적 성질을 계산
 - 다만, 실험 데이터와 혼용되지 않도록 별도 카테고리를 구성하여 관리하고, 계산 데이터 활용여부는 연구자에게 위임

② (데이터 저장 및 분석) 통합 저장 및 처리기법 개발 (4대 소재기관)

- (저장) 기존 소재정보은행은 소재 기관별 서버에 데이터를 저장하였으나, 데이터 통합 관리를 위해 통합서버 신규 구축('21.末)
 - 향후에는 컴퓨팅 환경을 연계한 '(가칭)산업소재 데이터센터'*를 지정하여 데이터 보안, 개방 등 종합관리 추진
- * (1단계) 7개의 서버로 구성, CPU 354 core, GPU 47개, 저장 용량 622TB
(2단계) 산업소재전용 초고성능 컴퓨팅(HPC) 구축 및 과기부 시스템과 연계 등
- (분석) 템플릿으로 구현된 원료, 조성, 공정, 물성 변수를 조정하여 중요특성 발굴 등 데이터 분석기법 개발(~'21.9月)
 - 화학, 금속 등 해당 분야 전문지식 및 피쳐(feature) 엔지니어링 기법*을 활용하여 불필요한 변수 제거 또는 생성
- * 소재 전문가가 AI 활용에 필요한 핵심 인자(원료, 조성, 공정, 물성 등)를 도출하는 기법

나. 데이터 활용

① (AI 표준모델 개발) 축적 데이터와 인공지능 상용 패키지를 활용, 8개 프로젝트별 특성에 맞는 AI 표준모델*을 단계적으로 개발(4대 소재기관)

- * 원료조성·공정요인(X변수)과 소재물성(Y변수)간 인과관계를 파악

< 인공지능 패키지 종류 >



- (1단계) 원료조성·공정요인(X변수)을 활용하여 소재 물성(Y변수)을 예측할 수 있는 모델 개발(~'23.上)
 - 지속적인 학습을 통해 예측 알고리즘의 성능향상을 도모
- (2단계) 소재 물성목표(Y변수)를 달성할 수 있는 원료조성·공정 설계(X변수)를 예측할 수 있는 역설계 모델 개발(~'24.下)
- * 후보물질 고속 스크리닝 등을 통해 소재개발 기간을 획기적으로 단축 가능

② **(활용목적별 서비스 제공)** 사용자 친화적인 웹 기반을 구축하여 소재 데이터 및 표준모델을 기업 및 연구자에게 제공 (4대 소재기관)

○ (데이터 제공) 소재개발에 필요한 독자적인 인공지능 모델을 개발하려고 하는 기업 및 연구자 등에게 데이터 제공 ('22.末)

* 공공기관에 구축한 데이터를 단계적으로 개방 확대

○ (표준모델 제공) 자체적인 모델 개발이 어려운 중소기업에게는 AI 표준모델 소스코드를 개방하여 활용 지원 ('23.末)

- 표준모델을 그대로 활용하거나 해당 기업 데이터를 적용하여 응용 가능

* 다만, 응용된 모델은 공공기관에게 다시 제공하여 모델의 성능향상 도모

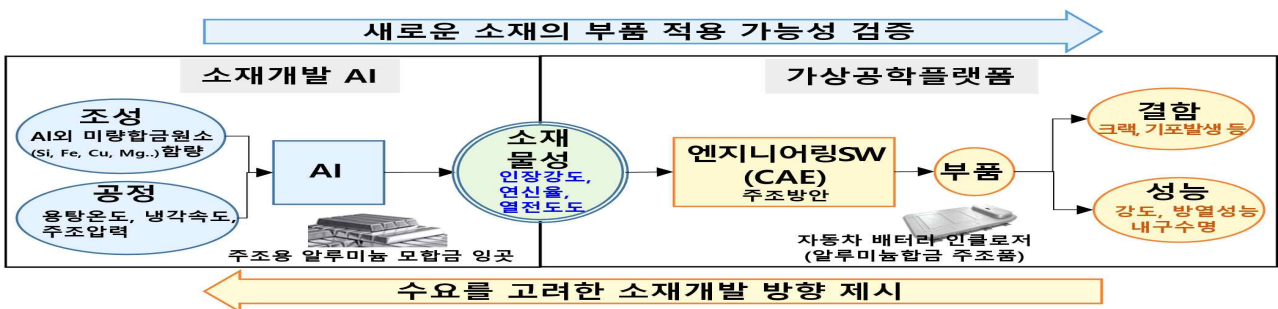
③ **(소재⇨부품화 가능성 검증)** 가상공학 플랫폼*과의 연계를 확대하여, 데이터기반 개발 소재의 부품화 가능성을 검증(4대 소재기관)

* 가상공학 플랫폼 : 시뮬레이션 SW를 활용, 선택된 소재가 부품화되는 공정을 분석함으로써 부품의 형상, 성능, 결함 등을 예측

○ 새로운 소재 정보를 엔지니어링 SW(CAE)에 입력하여 부품화에 따른 성능과 결함 등에 대한 예측 확대('22.上~)

○ 가상공간에서 소재설계에서 부품화까지 전주기 개발 과정을 수행하여 상용화 시기 단축 및 비용절감

<기계학습 모델 - 가상공학플랫폼 연계>



④ **(전문인력 양성)** 소재 및 인공지능 전문성을 보유한 융복합 인력양성

○ 실무 기술인력 대상 AI 융합 교육 과정을 신규 운영 ('21, 65억원)

* AI 기초교육(머신러닝, 딥러닝 등), AI 산업융합 심화 교육(임베디드 딥러닝 등) 등

○ 소재 기업, 연구자 대상으로 AI 플랫폼 활용 과정 및 산업인공지능(industrial-AI) 전문가 과정 신설·운영 ('22, 4대 소재기관)

* 8대 선도 프로젝트 內 시범 추진(ETRI 등 유관 기관과 협약 등)

4 데이터 기반 확충

◇ **데이터 표준화·축적·활용을 선순환 구조로 확산하기 위해 표준 제정, 인센티브 부여, 관리 시스템 구축 등 제도·체계 확충**

- **(데이터 표준화)** 데이터 기반 소재개발 활성화를 위한 표준 마련
 - 他소재·산업과의 데이터 상호호환성, 글로벌 경쟁력 확보를 위해 **표준 템플릿을 국가 표준으로 제정** 추진(국표원, '22년~)
 - 프로젝트별 표준 템플릿 기반으로 **다양한 소재분야별 템플릿 마련**
 - * 유사 소재의 경우 기존 템플릿 활용 등 **최소비용으로 다양한 템플릿 구성 가능**
(예시 : 미래차용 경량소재 → 드론용 경량소재)

- **(데이터 축적)** 인센티브 부여, 자금지원 등을 통해 민간의 자발적 데이터 축적을 지원하고, 데이터 지속 확충을 위한 **가이드 마련***(산업부, '21.末)
 - * 『산업 디지털 전환·지능화 촉진법(안)』 제10조②항 등 관련 조항 근거 활용
 - **(인센티브) 자발적인 데이터 축적을 위한 시스템적인 접근 강화**
 - 소재 R&D 수행기관의 **데이터 제출 절차***를 마련하고, 입력 실적을 **지표화**하여 연구 수행 기간 중 **지속 관리**
 - * 데이터 관리계획(Data Management Plan) 제출 ⇨ 표준템플릿 활용 연구결과 입력
 - 데이터 제출 규모에 따라 **바우처(데이터 마일리지)를 제공**하여 DB 우선 활용, DB 전문가 컨설팅 지원 등 추진
 - 데이터를 제공받은 **他 연구자가 해당 데이터를 활용하여 수익을 창출**할 경우, 데이터 공여자와 **수익을 공유**할 수 있는 체계 마련(데이터실명제 등)
 - * 데이터 생산자, 활용자 등 데이터 활용 표준계약서 마련 등
 - 산업 R&D 사업비 산정시 **데이터 구매, 컨설팅관련 비용 항목 신설** 및 데이터 전문 인력의 **참여율 추가 인정***
 - * (現) 100% → (改) 130%까지 확대

○ (자금지원) 데이터 개방·공여 등 소재데이터 활용·선도기업을 선정하여 산업지능화 펀드 지원(산업부, '21년~)

* 산업지능화 펀드 : AICBM기술(AI, IoT, Cloud, Big data, Mobile)을 활용해 공정·제품·서비스 혁신을 추진하는 산업 디지털 혁신 선도기업에 투자(총 4,000억원 규모)

□ **(데이터 활용)** 데이터 개방대상과 범위를 단계적으로 확대하고, 민관협력을 확산, 핵심 데이터 보호를 위한 보안 체계 구축

○ (단계적 데이터 개방) 초기에는 공공기관과 데이터를 제공하는 민간기업 중심으로 데이터 셋을 개방하여 데이터 축적 촉진

- 일반적인 상용 사용자에게는 가공된 데이터만 대상으로 검색 및 활용을 허용하되, 데이터 셋이 완비된 이후 전면 개방

○ (민관협력) 데이터기반 민관 공동 소재개발 프로그램 운영, 데이터 공유 등을 통해 민간 확산 추진(산업부, '22년~)

- 소재 공공기관과 주요 소재기업간 **오픈 플랫폼(Materials Data Open Platform)**을 구성, 데이터 공동생산 및 공유 추진

* (예) 일본 NIMS와 화학 4대 기업(미츠비시, 스미토모, 아사히화성, 미쓰이 화학)간 협의체를 구성, 상용 소재인 폴리올레핀 성능향상 도모

○ (보안체계) 핵심 데이터 보호를 위해 데이터 제공 시스템의 보안을 확보하고, 데이터 무단사용 방지를 위한 관리시스템 구축(4대 소재기관)

- 초기에는 온라인이 아닌 직접 기관방문 또는 폐쇄망(VPN 등) 형태*로 데이터 제공, 향후 통합서버 구축시 데이터 보안 솔루션도 함께 구축

* 기업, 연구자의 데이터 반출시 기관장 승인 절차 등 마련

- 무단 해외유출 방지 조항을 마련하고, 소재데이터 전용 식별시스템 (M-DOI*)을 구축하여 저작권 보호와 무단 사용 방지

* Materials Digital Object Identifier : 데이터 생성자, 데이터 인용이력 등 표시

5 소재데이터 Governance 가동

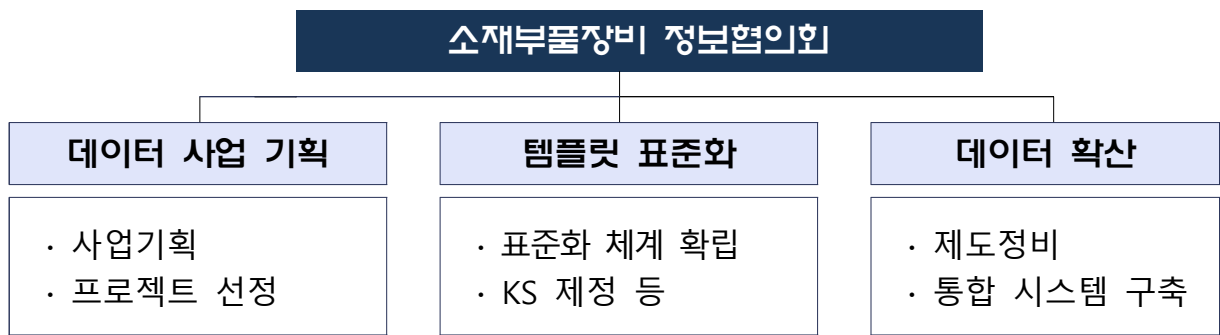
- **(구성)** 소재·부품·장비 정보협의회를 구성하고, 3개 분과별로 전문 기관이 참여하는 민관 협의체 신설 ('21.7月)

* 소재부품장비 경쟁력강화 특별법 제36조 제3항에 근거

- **(총괄 협의회)** 산업부, 4대 소재 거점기관, 협단체, 연구기관 등 20개 책임기관으로 구성하여 연차별 추진계획을 수립

* 화학연, 재료연 등 출연연 원장, KAIST, 서울대 등 대학총장, 관련 업계 등

- **(분과 협의회)** 데이터 사업기획, 템플릿 표준화, 데이터 확산 등 3개 분과를 구성하고(분과별 산학연 10인 내외), 구체적인 이행방안을 마련



- **(운영)** 분과별 협의회(월 1회), 총괄 협의회(년 1회) 개최 등을 통해 체계적인 사업추진과 실적을 점검

* 한국산업기술평가관리원(전략기획단) 또는 한국산업기술진흥원(KIAT)을 간사 기관으로 지정

- **(지원예산)** '20~24년까지 700억원 지원(소부장 R&D 예산 활용)

'20년 下	'21년	'22년	'23년	'24년
100억원	100억원	100억원	200억원	200억원

- 데이터 표준 체계 확립 및 인프라 구축(1단계, ~'22), 템플릿 확대, 민간부문으로의 확산(2단계, '23~) 등 단계적 로드맵에 따라 추진

* 구축된 인프라를 활용, 소재D/B 활용·확산을 위한 2단계 기획·로드맵 수립 (~'21.6)

V. 추진일정

추진 과제	추진 부처	추진 일정
1. 8대 선도 프로젝트 추진		
○ 선도 프로젝트 분야 선정 및 컨소시엄 구성	산업부	~'21.4
2. 데이터 표준화		
○ 소재 데이터의 표준화 공통 항목 도출	산업부	~'21.6
○ 표준 템플릿 및 표준화 D/B 구조 체계 마련	산업부	~'21.9
3. 데이터 축적과 활용		
○ 공공 데이터 확보	산업부	'21~
○ 인공지능 표준모델개발 및 서비스 개시	산업부	'22~
○ 전문인력 양성	산업부	'22~
4. 데이터 기반 확충		
○ 데이터이용 활성화 가이드라인 마련	산업부	~'21
○ R&D 규정 개정 등 보상체계·의무화 방안 마련	산업부·과기부	~'21
○ 전용식별 시스템(M-DOI) 구축	산업부	~'22
○ 표준 템플릿 KS 규격화	산업부(국표원)	'22~
5. 디지털 소재데이터 Governance 가동		
○ 소부장 정보 협의회 구성 및 분과위원회 운영	산업부	'21.7