

## 맞춤형 정밀의료 시대를 위한 의료 솔루션 디지털 병리 시스템 및 합성 의료데이터 기술동향

### 제 1 장 디지털 헬스와 디지털 병리학(digital pathology)의 기술 개요

#### 1. 새로운 패러다임 디지털 헬스

##### 1-1. 디지털 헬스케어의 진화

###### 1-1-1. 디지털 헬스케어와 데이터

###### (1) 디지털 헬스케어와 데이터의 상관관계

가. 새로운 의료 패러다임 디지털 헬스케어의 진화

나. 데이터와의 상관관계

###### (2) 데이터 3 법

###### (3) 국내 의료 데이터의 활용성

###### 1-1-2. IT 통합

###### 1-1-3. 치료 방식의 변화 유전체 시퀀싱(Sequencing)

###### (1) 유전자 분석

###### (2) 차세대 유전체 분석 기술

가. 전장엑솜시퀀싱(Whole Exome Sequencing, WES)

나. 전장 유전체 시퀀싱(Whole Genome Sequencing, WGS)

다. 목표 유전자 시퀀싱(Targeted Gene Sequencing, TGS)

라. 유전체 분석 기술과 디지털 헬스케어

###### (3) DTC 유전자 검사

가. DTC 유전자 검사 개요

나. 국내외 DTC 유전자 검사 동향

다. DTC 유전자 검사시 유의사항

###### 1-1-4. 디지털 병리학(Digital Pathology)

###### (1) 디지털 병리학 개요

###### (2) 디지털 병리학 기반 생산성 향상

###### 1-1-5. 정밀의학(precision medicine)의 전환

##### 1-2. 원격의료와 진화를 통한 디지털 헬스케어의 발전 전략

###### 1-2-1. 기존 의료시스템과의 결합

###### 1-2-2. 디지털 헬스케어의 발전 전략

## 2. 디지털 병리학(digital pathology) 기술 동향

### 2-1. 디지털 병리학의 등장 배경

#### 2-1-1. 병리학(Pathology)

(1) 질병과 병리학

(2) 병리 진단

#### 2-1-2. 디지털 병리학의 필요성

### 2-2. 디지털 병리 시스템

#### 2-2-1. 병리 검사과정

(1) 육안조직병리 관찰

(2) 조직슬라이드 제작

(3) 현미경 관찰

#### 2-2-2. 디지털 병리 과정

#### 2-2-3. 디지털 병리 솔루션

#### 2-2-4. 병리 데이터

### 2-3. 디지털 병리 활용을 위한 생태계 조성 및 이점

#### 2-3-1. 디지털 병리학 생태계 조성

#### 2-3-2. 디지털 병리학 장점

### 2-4. 디지털 병리학 시장 전망

#### 2-4-1. 시사점

#### 2-4-2. 디지털 병리학의 시장 전망

## 제 2 장 합성 데이터(Synthetic Data)와 디지털 치료제(Digital Therapeutics) 기술 동향

### 1. 합성 데이터 기술 동향

#### 1-1. AI 와 의료 기술

#### 1-2. AI 기반 합성 데이터(Synthetic Data)

##### 1-2-1. 합성 데이터(Synthetic Data) 개념

(1) 합성 데이터(Synthetic Data) 등장 배경

(2) 합성 데이터 개념

##### 1-2-2. 합성 데이터 목적

(1) 합성 데이터와 비식별화

(2) 합성 데이터의 특징

가. 합성 데이터의 이점

나. 합성 데이터의 단점

##### 1-2-3. 합성 데이터와 실제 데이터 비교

(1) 합성 데이터 생성 기술

- (2) 합성 데이터 대 실제 데이터
- (3) 합성 데이터 생성 방법
- (4) 합성 데이터 생성 과정
- 1-3. 합성 데이터 생성 방법의 종류
  - 1-3-1. 완전 합성 데이터(Fully Synthetic Data)
  - 1-3-2. 부분 합성 데이터(Partially Synthetic Data)
  - 1-3-3. 복합 합성 데이터(Hybrid Synthetic Data)
- 1-4. 합성 데이터 생성 프로세스
  - 1-4-1. CART 기법을 적용한 재현 데이터 생성
  - 1-4-2. 딥러닝 기반 합성 데이터 생성 방법
    - (1) GAN(Generative Adversarial Network)을 활용한 합성 데이터 생성
      - 가. GAN 기술 개요
      - 나. GAN 구조
        - ① 학습데이터
        - ② 생성자(generator) 네트워크
        - ③ 판별자(discriminator) 네트워크
      - 다. 적대적 학습 방법
    - (2) VAE(Variational Autoencoder)로 합성 데이터를 생성하는 과정
    - (3) 베이지안 네트워크(Bayesian Network)를 활용한 합성 데이터 생성

## 2. 합성 데이터 활용 사례 및 기술 동향

- 2-1. 합성 데이터의 활용 사례
  - 2-1-1. 빅데이터 분석을 위한 합성 데이터
  - 2-1-2. 기계학습 알고리즘 교육을 위한 합성 데이터
- 2-2. 합성 데이터 활용 분야
  - 2-2-1. 합성 의료 데이터
    - (1) 보건의료산업과 데이터
    - (2) 의료 분야에서 합성 데이터 활용
      - 가. EHR 기반 합성 의료 데이터 생성
      - 나. 임상데이터에 합성 의료 데이터에 활용
      - 다. 합성 데이터를 활용한 신약 개발
  - 2-2-2. 자율주행 자동차 구현을 위한 합성 데이터
    - (1) 합성 데이터 사용
    - (2) 합성 데이터 유용성을 위한 요구 사항
      - 가. 정확한 센서 데이터
      - 나. 환경에 따른 데이터 변형
      - 다. 정확성을 위한 주석

2-2-3. 제조업에서의 합성 데이터 활용

2-2-4. 금융산업에서의 합성 데이터 활용

(1) 금융 데이터

(2) 금융분야의 AI 도입

가. RPA 시스템

나. 영업 및 마케팅 분야

다. 챗봇 활용

라. 투자 관리 및 트레이딩(Trading)

마. 사기 및 부정방지 기능

바. AI 기반 신용분석시스템

(3) 금융기관의 AI 도입을 저해하는 요소

(4) 금융분야의 합성 데이터 활용

2-3. 국내외 합성 데이터 기술 동향

2-3-1. 해외 합성 데이터 기술 동향

(1) 미국 SSB

(2) 미국 SynLBD

(3) 독일 노동청의 사업장 패널(IAB Establishment Panel)

(4) 영국 SYLLS(Synthetic Longitudinal Studies Data)

2-3-3. 국내 합성 데이터 기술동향

(1) 통계청

(2) 한국신용정보원 개인신용정보 재현자료

(3) 한국정보화진흥원의 코리아크레딧뷰로(KCB)

2-4. 합성 데이터의 해결 과제

## 참고문헌

### 그림 목차

[그림 1] 데이터 흐름 중심의 디지털 헬스케어 산업 구조

[그림 2] 의료 및 생명과학을 위한 비정형 데이터 관리

[그림 3] 헬스케어 서비스 모델의 미래

[그림 4] 의료 데이터 공유 및 원격 백업

[그림 5] 데이터 3 법 주요 내용 및 향후 과제

[그림 6] 마이 헬스웨이 의료 데이터 수집 체계

[그림 7] 병원·헬스케어 영역의 일반적인 스토리지

[그림 8] 임상유전체 연구를 위한 NGS의 활용

[그림 9] 표적 유전자 패널, 전체 엑솜 시퀀싱 및 전체 유전자 시퀀싱 접근법의 비교

[그림 10] 엑솜시퀀싱(Exome Sequencing)

- [그림 11] 전장 유전체 시퀀싱의 주기적인 재분석 과정 모식도
- [그림 12] Targeted sequencing flow
- [그림 13] DTC 유전자 검사 과정
- [그림 14] 공개 유전자 데이터베이스의 범위를 보여주는 가상 가계도
- [그림 15] DTC 유전자 검사 대 소비자 소유 유전자 데이터
- [그림 16] 디지털 병리학 기반 품질 향상
- [그림 17] 디지털 병리학 기반 생산성 향상
- [그림 18] 디지털 병리학의 분석 단계
- [그림 19] 의료분야의 중요 기술과 의료 데이터의 종류 및 동향
- [그림 20] 바이오마커를 이용한 맞춤형 치료제 개발
- [그림 21] 원격의료를 활용한 통합 의료 사이클
- [그림 22] 원격의료 진료 과목 비중
- [그림 23] 2021 년 글로벌 헬스케어 부문이 해결해야 할 과제
- [그림 24] 진단 및 정밀 종양학을 위한 AI 기반 디지털 병리학
- [그림 25] 인체의 항상성
- [그림 26] 디지털 병리 워크플로우
- [그림 27] 가상 현미경(virtual microscopy)
- [그림 28] 디지털 병리 이미지 분석
- [그림 29] 병리조직 검사 과정
- [그림 30] 아날로그 방식의 병리학
- [그림 31] 전통적 병리학과 디지털 병리학의 작업 흐름 차이
- [그림 32] 디지털 병리 과정
- [그림 33] 디지털 병리학 워크플로우
- [그림 34] 병리학 PACS 아키텍처의 정보 흐름
- [그림 35] 디지털 병리학의 분석 단계
- [그림 36] 디지털 병리학 시스템
- [그림 37] 디지털 병리학의 7 단계
- [그림 38] 디지털 병리학 시장의 세분화
- [그림 39] AI 의료 기술의 가치 사슬
- [그림 40] 인공지능(AI) 및 빅데이터
- [그림 41] 합성 데이터 생성을 위한 워크플로우
- [그림 42] 합성 데이터 프로세스
- [그림 43] 데이터 익명화
- [그림 44] 합성 의료 데이터의 생성 및 활용 과정
- [그림 45] 데이터 창출 과정
- [그림 46] 합성 데이터 라이브러리 구조
- [그림 47] 데이터 패러다임 변화 분석

[그림 48] 기계학습 및 합성 데이터

[그림 49] 재현 데이터 생성 방법

[그림 50] 완전 합성 데이터와 부분 합성 데이터 비교

[그림 51] 합성 데이터를 활용한 개인정보 보호 사례의 예

[그림 52] 변수간 상관관계 비교

[그림 53] CART 결정트리 알고리즘

[그림 54] 의사결정트리 알고리즘 작동 과정

[그림 55] GAN 아키텍처 개요

[그림 56] GAN 작업 과정

[그림 57] GAN의 개념도

[그림 58] GAN의 학습 방법

[그림 59] Generative Adversarial Network

[그림 60] 생성적 적대 네트워크

[그림 61] 판별자 교육 프로세스

[그림 62] GAN의 작동 방식

[그림 63] GAN을 활용한 재현 데이터

[그림 64] VAE의 구조

[그림 65] 생성모델 개념도

[그림 66] VAE로 합성 데이터를 생성하는 방법

[그림 67] 베이지안 네트워크의 예

[그림 68] A constraint graph grouping variables

[그림 69] 재현 데이터 생성기술 기반 가명처리 서비스 Use Case

[그림 70] 합성 데이터 생성을 위해 제안된 시스템

[그림 71] 개인정보 보호를 위한 AI의 안전한 구현

[그림 72] 기계학습 알고리즘 교육을 위한 합성 데이터 개요

[그림 73] 머신러닝을 위한 합성 데이터 셋 생성 워크플로우

[그림 74] 머신러닝(ML) 워크플로우

[그림 75] 의료분야에서 인공지능 활용

[그림 76] 전자건강기록 데이터에서 예측·분석을 생성하는데 필요한 프로세스

[그림 77] 합성 데이터 생성을 위한 워크플로우

[그림 78] 임상적 증거 생성의 과정

[그림 79] 합성 의료 데이터 생성

[그림 80] 머신러닝(ML)을 활용한 의료 데이터 생성

[그림 81] 자율주행 차량의 장면 인식 및 물체 감지를 위한 딥러닝 아키텍처

[그림 82] 딥러닝을 통한 센서 융합

[그림 83] 합성 데이터와 실제 데이터 혼합

[그림 84] 자율주행 프레임워크

- [그림 85] 합성 이미지 렌더링 과정
- [그림 86] 학습용 이미지를 생성하는 Generative Model 구성도
- [그림 87] 합성 데이터를 추가하기 전과 후의 위험한 오브젝트 감지율
- [그림 88] 금융분야의 AI 도입
- [그림 89] RPA 의 이점
- [그림 90] 챗봇 구축
- [그림 91] 로보어드바이저 단계
- [그림 92] 위험 관리를 위한 인공지능
- [그림 93] AI 를 적용한 신용 프로세스
- [그림 94] GAN 을 활용한 새로운 데이터 셋
- [그림 95] 합성 생태계
- [그림 96] 데이터 보호 규정의 범위
- [그림 97] 합성 데이터 생성 과정
- [그림 98] 금융 빅데이터 개방시스템 구조도
- [그림 99] 재현 데이터 구현
- [그림 100] 비식별 조치 및 사후관리 절차

## 표 목차

- [표 1] 의료서비스 패러다임 전환에 따른 치료방식의 변화와 디지털 헬스케어의 개념
- [표 2] 모바일 헬스케어 애플리케이션 기능별 분류
- [표 3] 해외 빅데이터 활용 사례 및 의료 데이터의 진화
- [표 4] 데이터 3 법의 개정안
- [표 5] 데이터 3 법의 주요 특징 및 데이터 활용을 위한 비식별화 과정
- [표 6] 의료 데이터의 종류 및 동향과 빅데이터 기술을 활용한 치료방식의 변화
- [표 7] ACMG 가이드라인의 우연히 발견된 결과 보고 기준중 유전성암 관련 유전자
- [표 8] DTC 유전자 검사 3 차 시범사업
- [표 9] 글로벌 정밀의료 시장 전망(단위: 십억달러)
- [표 10] 병리학의 종류
- [표 11] 육안 조직 검사
- [표 12] 의료 빅데이터 활용 강화 분야
- [표 13] 합성 데이터 분야 핵심기술 및 합성 데이터 생성 과정
- [표 14] 데이터 비식별화 예시 및 비식별 조치와 사후관리 절차
- [표 15] 국내외 인공지능 윤리 관련 원칙 및 가이드 발간 현황
- [표 16] 데이터 비식별화 적용기법
- [표 17] 합성 데이터와 실제 데이터의 질 측정 결과
- [표 18] 재현 데이터 생성 도구
- [표 19] 재현 데이터의 분류

[표 20] 합성 데이터 생성 기법

[표 21] 주요 결정트리 파라미터 및 의사결정트리 분류 알고리즘

[표 22] 합성 데이터(Synthetic Data)의 용도별 분류 및 합성 데이터 오픈소스

[표 23] 합성 데이터에 사용할 수 있는 주식 유형

[표 24] 일반신용정보

[표 25] SSB 이용 절차

[표 26] 합성 데이터 (Synthetic Data) 분야 핵심기술 연구목표