

# 의료기기에서 생성되는 사운드 생체신호 분석을 위한 빅데이터 플랫폼 설계

고광만\*, 김성진\*\*, 신정훈\*\*\*, 윤희선\*\*\*

\*상지대학교 컴퓨터정보공학부

\*\* (주)테스티안

\*\*\* (주)리아컴즈

e-mail:kkman@sangji.ac.kr

## Design of Big Data Platform for Sound Bio-Signal Analysis from Medical Devices

Kwang-Man Ko\*, Kim Seongjin\*\*, Jung-Hoon Shin\*\*\*, Hee-Sun Youn\*\*\*

\*School of Computer Information Engineering, Sang-Ji University

\*\* (Ltd)TESTIAN, \*\*\* (Ltd)RiaComs

### 요 약

최근에는 의료 빅데이터 분야에서 의료기기, 의료전문가로부터 생성 또는 감지되는 사운드 생체신호 (심장박동, 호흡, 맥박, 진맥) 데이터의 특징을 디지털 데이터로 추출하여 패턴 데이터로 변환한 후, 이를 빅데이터 분석 플랫폼 기반으로 분석하여 진료, 처방, 예방 등에 유용한 정보를 생성하는 모델 구축 연구가 활성화되고 있다. 본 논문에서는 사운드 생체신호 특징을 디지털 데이터로 추출하여 ㈜리아컴즈 NeoQubit 빅데이터 플랫폼을 기반으로 패턴 데이터를 분석하고 예측할 수 있는 모델을 제시한다.

### 1. 서론

빅데이터 분석 플랫폼은 수집된 정보를 분석하여 비즈니스 의사결정을 효과적으로 할 수 있도록 지원하고, 생산성 향상에 기여할 수 있는 의미있는 데이터를 제공한다. 빅데이터 분석 기술을 활용하여 적용한 아마존이나 월마트의 사례에서 보듯 매출 신장에 직접적인 영향을 미치므로 경쟁력을 확보하려는 수많은 기업체들이 필요로 하고 있는 상황이며 여러 가지 형태로 적용하기 위한 시도가 나타나고 있다.

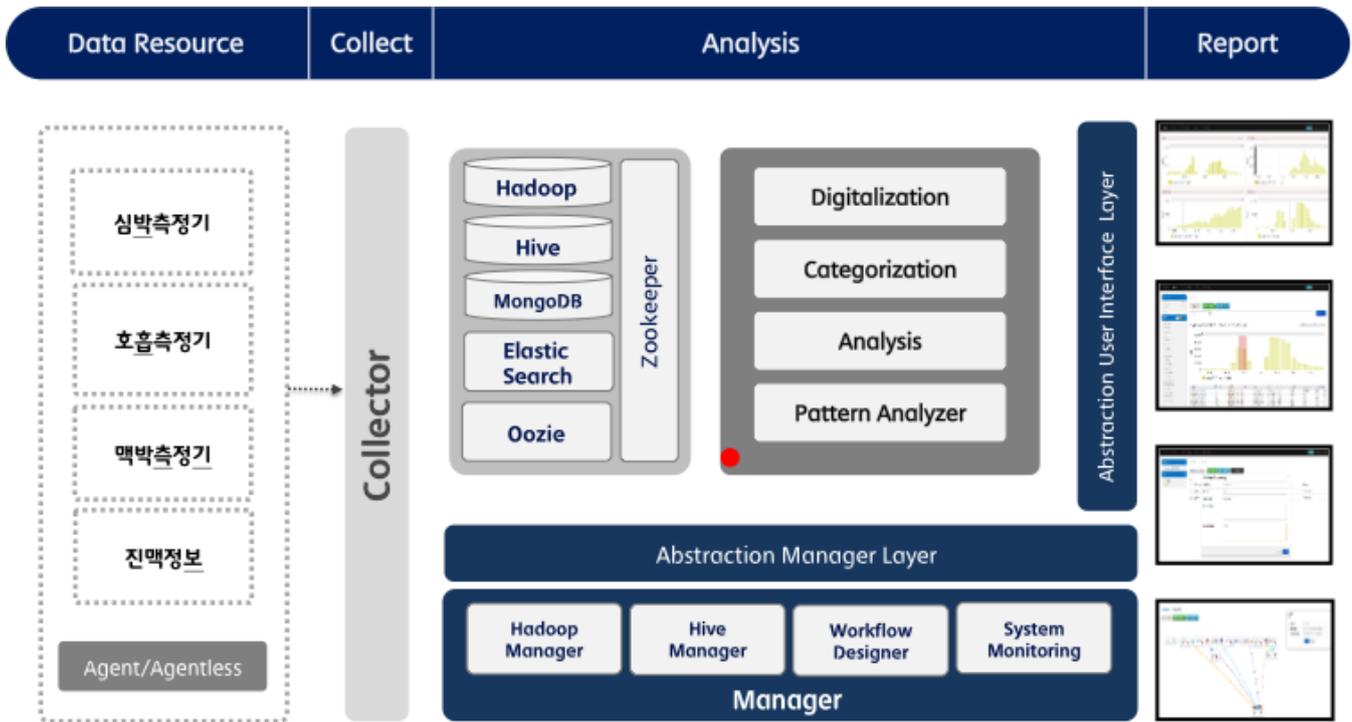
인구 고령화와 만성질환 유병률의 증가로 의료비 문제와 의료서비스의 접근성 및 질에 관한 문제가 논의되면서 많은 국가에서 IT와 의료기술을 접목한 u-Health 도입을 추진하고 있고, 의료비 절감 등의 사회경제적 비용감소에 관심이 집중되어 있으며, 최근 질병을 조기에 진단하지 못하거나 의사의 오진으로 인한 환자의 피해가 지속적으로 발생하고 있으며, 이상적 환경의 임상시험과 실제 진료 현장과의 차이로 표준적 치료법의 보급이 늦어지거나 현장에서는 잘 맞지 않는 경우가 많이 있다. 의료계는 이러한 문제점의 개선을 위해 진단기술 향상을 위한 혁신적 노력이 필요하다. 의료정보 빅데이터 분석 기법을 활용하여 의료의 오진율을 줄이는 방법으로 의료기기에서 발생하는 사운드 생체신호(bio-signal)를 분석함으로써 진료의 판단에 보조적인 역할을 하고자 한다. 이를 위해, 보다 많은 데이터를 패턴화하고 분석함으로써 정확도 측면에서 향상될 것이 예상되며 더욱더 오진율을 낮출 수 있어 더 많은 생명을 구할 수 있으리라 판단된다.

(주)리아컴즈의 'NeoQubit'은 하둡 기반의 분산처리 구조를 적용하여 정형, 반정형 및 비정형 데이터를 효과적이고 빠르게 처리 할 수 있도록 개발된 빅데이터 분석 플랫폼이다. 특히, 페타바이트급 이상의 대용량 데이터 처리 및 분석이 가능하고, 비즈니스 요구 사항에 맞는 정보 추출이 용이하며, 데이터의 실시간 분석, 배치 분석 및 상관 분석을 통하여 최적의 고객 Needs에 맞는 정보를 제공할 수 있는 빅 데이터 분석 플랫폼을 지향하고 있다.

본 논문에서는 (주)리아컴즈 NeoQubit을 핵심 기반으로 사운드 생체신호 데이터를 분석하고 진단에 필요한 정보를 예측하기 위한 의료정보 빅데이터 분석 플랫폼 모델을 제시한다.

### 2. 관련연구: (주)리아컴즈 NeoQubit

(주)리아컴즈는[1] 2012년에 데이터 분석 플랫폼 개발하기 위해 수집, 저장·색인, 분석의 핵심 오픈소스인 하둡과 에코시스템을 활용하여 2013년 7월에 NeoQubit라고 하는 빅데이터 분석 플랫폼을 상용화하였다. 하둡 코어[2]인 HDFS와 Map/Reduce[3]를 기본으로 데이터 수집을 위한 오픈소스인 Flume 기반으로 데이터를 Agent 와 Collector를 개발하였고, 실시간 데이터 스트림에 대한 처리를 위하여 CEP 엔진인 Esper를 활용하여 실시간 데이터에 대한 모니터링이 가능한 엔진을 플랫폼에 추가하였다. 또한 Batch 분석을 위해 Hive를 탑재하여 사용자의 다양한 분석을 처리할 수 있는 기능을 추가하였으며, 분석 결과에 대한 저장을 위하여 분산 병렬 쿼리가 가능한 MongoDB



[그림 1] NeoQubit 기반 사운드 생체-신호 분석을 위한 빅데이터 분석 플랫폼

를 플랫폼에 탑재하였다. 특히, 대중적인 분석 툴인 R 과 인터페이스를 통하여 연동이 가능하며, 다양한 형태의 그래프로 이루어진 대시보드를 사용자에게 제공할 수 있다.

NeoQubit의 특징은 정형, 반정형, 비정형 등 다양한 형태의 데이터 구조를 저장, 처리, 분석 및 리포트할 수 있는 빅데이터 분석솔루션으로써, 기계 데이터 분석에도 탁월한 경쟁력을 가진 솔루션이다. 마찬가지로 기계 데이터 분석 및 배치분석에 강점을 가지고 있으며, 솔루션 비즈니스에 강점을 가진 글로벌 기업 Splunk 빅데이터 분석 솔루션과의 기능적, 성능적 비교가 매우 중요한 의미를 지닌다고 볼 수 있다. 최근에는, 기 개발된 NeoQubit을 기반으로 기존의 기계 데이터 분석에서 더 나아가 사운드를 패턴화하여 이를 분석하는 기법을 적용하여 의료분야 빅데이터 분야로 비즈니스 모델의 확장을 도모하고 있다.

### 3. 사운드 생체신호 분석을 위한 빅데이터 플랫폼 설계

#### 3.1 전체 시스템 설계

사운드 생체신호에 대한 효율적인 분석결과를 도출하기 위해 Neoqubit 빅데이터 분석 플랫폼을 기반으로 [그림 1] 과 같은 시스템을 개발한다. 데이터 리소스 부분에서는 심박측정기, 호흡측정기, 맥박측정기, 진맥정보와 같은 사운드 생체신호를 입력으로 받아 발생하는 사운드의 특징을 판별할 수 있는 패턴 개발, 사운드 생체신호의 특징을 디지털 수치로 표현, 디지털 수치로부터 데이터를 추출할 수 있는 기법을 적용한다.

#### 3.2 사운드 생체-신호 빅데이터 분석 검증

의료기기, 의료전문가로부터 생성 또는 감지되는 사운드 생체신호(심장박동, 호흡, 맥박, 진맥) 데이터의 특징을 디지털 데이터로 추출하여 패턴 데이터로 변환한 후, 이를 (주)리아컴즈 NeoQubit 빅데이터 분석 플랫폼 기반으로 분석하여 진료, 처방, 예방 등에 유용한 정보를 생성하는 비즈니스 모델을 검증하기 위해부터 첫째, 의료기기 또는 의료전문가로부터 생성 또는 감지되는 사운드 생체신호에 대한 200건 이상의 디지털 데이터베이스로 구축하였으며, 둘째, 사운드 생체신호의 특징을 디지털 데이터로 추출하여 디지털 데이터로 변환하고 이를 패턴화된 디지털 데이터 변환 기술을 개발중이다. 또한 패턴화된 디지털 데이터와 비교 분석할 수 있는 기술을 접목하여 사운드 생체-신호 빅데이터 모델을 검증하였다.

#### 4. 결론

사운드 생체신호에 대한 효율적인 분석결과를 도출하기 위해 Neoqubit을 기반으로 빅데이터 분석 시스템을 설계한 후 이를 위한사운드 생체신호 변환 기법 및 검증 방법을 제시하였다.

#### 참고문헌

- [1] ㈜리아컴즈; <http://www.riacoms.com>
- [2] 톰 화이트, 하둡 완벽가이드, 한빛미디어, 2013.
- [3] 스나리스 페레라, 하둡 맵리듀스 프로그래밍, 에이콘출판사(주), 2013.
- [4] 라스 조지, HBase 완벽가이드, 한빛미디어, 2013.