

# 인터넷 데이터 집중화 문제와 대응연구 동향

## Survey on Data Consolidation Issue in the Internet and Corresponding Research

정희영 (H.Y. Jung, hyjung@etri.re.kr)

데이터중심네트워크연구실 책임연구원

고남석 (N.S. Koh, nsko@etri.re.kr)

데이터중심네트워크연구실 책임연구원/실장

### ABSTRACT

The Internet is evolving into a fundamental social infrastructure, similar to electricity and water. In particular, the role of the Internet is becoming more critical as the global digital transformation infrastructure of the 4th industrial revolution. Based on this trend, we note that the Internet is monopolized by a few large organizations even though it was originally designed as a purely distributed system. In this paper, we have reviewed the domination trends, especially in terms of data consolidation. In addition, we have summarized the corresponding research activities from the Internet Society, EU Horizon-2020, and Decentralized web. Further, this paper proposes an approach to address the data consolidation issue from the Korean perspective.

**KEYWORDS** 인터넷 집중화, EU NGI, 분산 웹

## 1. 서론

인터넷은 전기, 수도와 같이 우리가 일상생활을 살아가기 위한 필수적인 사회 인프라로 자리매김하고 있다. ITU 통계에 따르면 2015년경에 인터넷 사용자는 32억 명 이상으로 30억 명의 수도 인프라 사용자 수를 이미 넘어섰으며[1], 2019년 현재 전 세계 인터넷 사용자 수는 44억여 명으로 전 세계 인구의 57.3%가 인터넷을 이용하고 있다[2].

또한 4차 산업혁명에 따른 디지털 트랜스포메이션(Digital Transformation) 추세에 따라 이를 실현하기 위한 글로벌 인프라로서 그 중요성이 더욱더 커지고 있다.

이와 같이 인터넷이 단순한 통신기술을 넘어서서 필수적이고 보편적인 사회인프라로 진화함에 따라 여러 가지 예상하지 못했던 문제점들도 나타나고 있다. 그 대표적인 예가 인터넷에서의 집중화(Consolidation) 현상이다. 2018년 인터넷 보고서

\* DOI: <https://doi.org/10.22648/ETRI.2019.J.340611>

\* 본 연구는 과학기술정보통신부 및 정보통신기술진흥센터의 정보통신·방송 연구개발 사업의 일환으로 수행하였음[2017-0-00045, 초연결 지능 인프라 원천기술 연구개발].



본 저작물은 공공누리 제4유형

출처표시+상업적이용금지+변경금지 조건에 따라 이용할 수 있습니다.

©2019 한국전자통신연구원

[3]에 따르면 소수의 대형 기업들이 온라인 세계를 지배하고 있다. 예를 들면 인터넷 검색의 경우 구글이 차지하는 비중이 90%를 넘고 있으며, 인터넷 브라우즈의 경우 구글 크롬의 점유율이 데스크톱의 경우 64.7%, 이동단말기의 경우 47.8%에 달하고 있다. 페이스북은 페이스북 메신저, 왓챗(WhatsApp), 인스타그램(Instagram)과 같은 자사 서비스를 포함하면 전 세계 상위 6개의 소셜 미디어 플랫폼 중 4개를 차지하고 있다[4]. 또 하나 주목할만한 부분은 아시아태평양 지역에서 구글 서비스에 관련된 트래픽이 전체 트래픽 중 40%가 넘는 독점현상을 보인다는 것이다[5]. 이러한 독점에 따라 페이스북과 구글은 중국을 제외한 글로벌 디지털 광고 시장의 84%를 차지하고 있으며, 아마존은 미국의 소매 시장의 49%, 알리바바는 중국 이커머스(e-commerce) 시장의 60%를 차지하고 있다[4].

이러한 소수 기업에 의한 집중화는 자연스럽게 사용자 데이터의 집중화를 가져온다. 데이터는 빅데이터와 인공지능을 기반으로 하는 4차 산업시대의 가장 중요한 자원이다. 따라서 빅데이터 시대에서는 현재 산업의 기반이 되는 석유와 같은 역할을 데이터가 할 것이라고 예상되며, 이에 따라 데이터를 차지하기 위한 치열한 경쟁들이 발생할 것으로 전망되고 있다[6].

구글과 페이스북은 자신들이 수집한 데이터들을 목표(Target) 광고를 위한 용도로 주로 사용하였으나 최근에는 데이터들을 적절히 처리한다면 다양한 종류의 인공지능이나 인지 서비스에 사용할 수 있다는 것을 발견하였으며, 이를 통해 번역, 시각 인식, 개인 정보 획득과 같은 새로운 비즈니스 모델을 만들어 내고 있다. 또한 이런 데이터들은 서비스를 향상시키고 사용자를 유인하기 위한 데이터 네트워크 효과(Data-network effect)라고 불리는 경제적 수단으로도 활용되고 있다. 역사학자인

Yuval Noah Harari는 2018년 다보스 포럼 강연에서 인류 사회의 기반이 되는 자산이 고대 사회에서는 토지였고, 산업혁명 이후는 기계가 그 역할을 하였으나, 이제는 데이터가 21세기의 모든 생산품을 만드는 기반이 되고 있음에 주목하고, 그 데이터가 소수에게 집중되어 소수가 세상을 지배하는 어두운 미래를 전망하고 있다[7].

우리의 일상생활 경험에서 알 수 있듯이 거대 기업에 의해 수집되는 사용자 데이터는 대부분 인터넷 사용자들이 서비스를 무료로 사용하는 대가로 자신도 인지하지 못하고 제공하는 것이다. 이렇게 수집된 데이터를 통해 개인의 프라이버시 침해와 인터넷 서비스 산업이 몇몇 기업에 의해서 독과점되는 문제가 발생한다. 인터넷 서비스를 통한 개인 정보 오용의 위험성에 법적으로 대응하기 위해 최근에 실행되는 것이 유럽에서 2018년부터 실행되고 있는 GDPR(General Data Protection Regulation)이다[8]. GDPR은 모든 회원국이 의무적으로 준수해야 하는 강제 규정으로 개인 정보 보호에 대한 기업의 책임과 사용자 데이터의 자유로운 이동 보장을 규정하며 위반 시 과징금을 부과한다. 2018년 프랑스는 GDPR을 위반한 구글에 5,000만 유로의 벌금을 부과하였다. GDPR과 더불어 국가의 정보주권 보호를 위해 기업이 데이터를 수집한 국가 내에서 관련 데이터를 저장 및 처리할 것을 요구하는 데이터 지역화(Data Localization)를 법제화하는 국가도 점차 늘어나고 있다.

이러한 인터넷 데이터 집중화에 대한 법적인 대응과 더불어 기술적으로 인터넷 데이터 집중화 문제를 해결하고자 하는 연구가 미국, 유럽 등을 중심으로 최근 시작되고 있다. 본 고에서는 인터넷 기술과 정책을 주도하는 ISOC(Internet Society)와 유럽 Horizon-2020 NGI(Next Generation Internet) 프로젝트 및 웹의 창시자인 Tim Berners

Lee에 의해서 주도되고 있는 분산 웹(Decentralized Web) 등에서의 최근 연구동향 분석을 통해 인터넷 데이터 집중화에 대응하고 있는 국제 동향을 살펴보고 연구개발 측면에서의 한국의 대응방안을 논하고자 한다.

## II. 대응 연구 동향

현재 인터넷에서의 데이터 집중화 현상에 대해 인터넷 커뮤니티와 유럽 R&D 프로젝트를 중심으로 대응 연구가 활발히 이루어지고 있으며 주요 연구 현황은 다음과 같다.

### 1. 인터넷 소사이어티(ISOC)

ISOC는 글로벌 인터넷에 대한 표준, 교육 및 정책을 담당하는 국제 비영리 기관으로, 전 세계 기관 및 개인에 대한 인터넷 사용 환경을 향상시키고 지원함으로써 인터넷을 개방형으로 지속해서 발전시키는 업무를 담당한다. 이를 위해 인터넷 기술에 대한 표준을 담당하는 IETF(Internet Engineering Task Force)와 IRTF(Internet Research Task)를 관장하

는 IAB(Internet Architecture Board)와 인터넷 주소 할당을 담당하는 IANA/ICANN(Internet Assigned Numbers Authority/Internet Corporation for Assigned Names and Numbers)을 산하 기관으로 두고 있다.

ISOC는 글로벌 사회 인프라로서의 인터넷이 사용자의 사용 형태 변화에 따라 다양한 방향으로 진화하고 있음에 주목하고 인터넷의 현황 및 진화와 관련된 상황을 보고서 형태로 지속해서 발간하고 있다. 2017년 보고서는 미래의 고도로 연결된 인터넷 기반 경제를 전망하고 있으며, 인터넷 기술이 이러한 인터넷 기반 경제에 어떤 영향을 미칠 것인지에 대한 논의를 다루고 있다. 특히 올해 2019년에 발간된 보고서에서는 인터넷에서의 집중화 문제를 현재 인터넷이 당면한 가장 중요한 이슈로 제기하였다[9].

이 보고서의 주요 내용은 인터넷 기반 경제의 집중화 현상과 이 집중화가 함의하는 바가 어떤 것인지에 대한 ISOC 관점에서의 분석이다. 보고서에 따르면 인터넷은 현재 40억 명이 넘는 사용자가 사용하는 글로벌 사회 경제적 인프라이며 전 세계인들에 대해 지속적인 혁신과 기회를 제공하는 플랫폼의 역할을 하고 있다. 인터넷의 역사를 돌아보면 인터넷은 항상 더 나은 아이디어가 기존의 아이디어를 대체할 수 있는 지속적인 혁신을 지원하는 플랫폼의 역할을 하여왔으며, 이러한 성격이 인터넷의 자연스러운 진화의 특징임을 설명하고 있다. 인터넷 기반 경제를 분석하기 위해 이 보고서는 인터넷 경제를 크게 인터넷 응용(Internet Applications), 인터넷 접속(Access Provision) 및 서비스 인프라(Service Infrastructure)의 세 부분으로 나누고, 이 세 가지 분야에서 일어나고 있는 집중화 현상에 대한 분석을 제공한다.

먼저 인터넷 응용은 사용자들이 인터넷상에서 어떻게 서로 간에 정보를 나누며 혁신적인 아이디

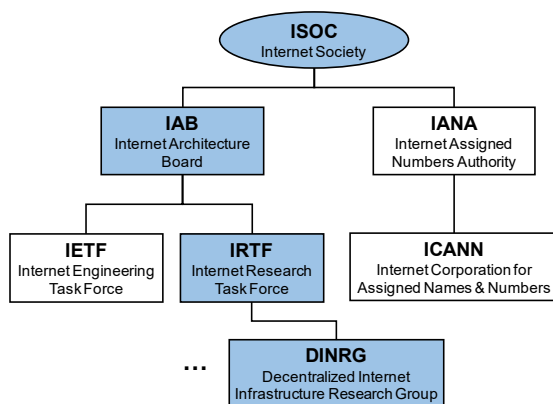


그림 1 ISOC 구조와 데이터 집중화 관련 그룹

어를 구현하고 있는가에 관련된 부분이다. 보고서에서는 현재 상위 5개의 업체(Alphabet, Amazon, Tencent, Facebook, Alibaba)가 이 분야의 시장을 대부분 점유하고 있는 현상을 지적하고 있다. 뿐만 아니라 아마존의 클라우드 서비스 제공이나 구글의 모기업인 알파벳이 다양한 인터넷 서비스 회사를 인수·합병하고 있는 것과 같이 각 업체는 기존의 서비스 범위를 다른 서비스 분야로까지 지속해서 확장하고 있음을 주목하고 있다. 이에 따라 전체적으로 증가하는 사용자와 트래픽이 소수의 지배적 단일 기업(One-stop players)에 집중되고 있다고 분석한다.

인터넷에 대한 접속은 인터넷 사용자에게 대한 글로벌한 연결을 제공하는 분야이다. 이 분야는 국가에 따라 상황이 서로 달라 집중화의 추세를 파악하기는 쉽지 않다고 분석하고 있다. 그러나 기존의 접속 제공 사업자들이 대체로 모바일의 증가, 데이터 수요의 증가, 이익 감소 등에 따라 변화의 압력을 받고 있으며, 이에 대응하기 위해 인프라를 공유하거나 가용 스펙트럼의 효율화 방안 개발 및 전략적 제휴와 인수합병을 통해 수익을 높이고 시도하고 있다고 분석하고 있다.

서비스 인프라는 인터넷을 구성하고 유지하기 위한 서비스와 비즈니스를 의미하며, 인터넷 네임과 주소의 관리, 콘텐츠의 호스팅과 분배, 네트워크 간의 상호연결 등을 포함한다. 이 분야에서는 아마존이 클라우드 서비스와 CDN(Content Delivery Network)에 투자하거나, 구글이 해저 케이블을 설치하는 것과 같이 대형 응용 사업자들이 인프라 계층으로 영역을 확장하고 있다. 또한 인터넷 1계층 네트워크상의 국제 간 데이터 전송에 따른 비용을 줄이기 위해 CDN의 사용이 증가하고 있으며, 대형 플랫폼 업체는 자신들만의 인프라를 구축하는 집중화 현상이 발생하고 있음을 지적하고 있다.

이 보고서에서는 인터넷이 현재와 같이 글로벌 플랫폼으로 성공하기까지 가장 핵심이 되는 요소가 특정 기업이나 기술에 의한 지속적인 영향을 받지 않는 특성에 있었으며, 이러한 소수 기업에 대한 의존성의 심화는 인터넷에 대한 위협 요소가 되고 있음을 지적하고 있다. 보고서에서는 집중화 문제를 해결하고자 하는 움직임이 정부 차원 및 인터넷 이해당사자들로부터 일어나고 있음을 설명하고 있다. 정부 차원의 대표적인 대응으로는 EU의 GDPR, 중국, 러시아 등의 데이터 지역화 관련 법규 등을 들고 있으며, 인터넷 이해당사자들 간의 대응으로는 IAB에서 이루어지고 있는 관련 논의나 웹의 창시자인 Tim Berners Lee가 수행 중인 솔리드(Solid) 프로젝트 등을 그 대표적인 예로 들고 있다. 이 보고서에 기반하여 ISOC는 인터넷 집중화를 2019년의 중점 과제로 중 하나로 선정하였다 [4].

ISOC의 문제 제기에 따라 인터넷에 대한 장기적인 기술 방향성을 담당하는 IAB에서 인터넷 집중화 문제에 대한 논의가 시작되었으며, 그 논의 결과를 인터넷 문서로 발간하였다[4]. 이 인터넷 문서는 인터넷 집중화 문제에 대한 다양한 경제적·기술적 측면의 분석을 그 주 내용으로 삼고 있다. 이 문서에서는 만일 인터넷의 사용 환경이 변화한다면 인터넷의 구조와 기술도 이 환경에 대응하여 변화하여야 하며, 집중화라는 환경 변화에서 새롭게 진입하거나 작은 규모의 경제 주체들을 지원하기 위한 기술 개발의 필요성을 지적하고 있다. 이 문서에서는 이러한 분석을 기반으로 집중화 문제에 대해 향후 인터넷 커뮤니티에서 연구가 필요한 세 가지 이슈를 다음과 같이 제안하고 있다.

- 집중화에 문제에 대한 측정: 인터넷의 집중화 정도와 측면, 인터넷 트래픽의 집중 장소 및 시간에 따른 변화 추세

- 집중화와 인터넷 기술과의 관련성: 현 인터넷 기술이 집중화에 끼치는 영향, 집중화가 향후 인터넷 인프라와 구조에 미칠 영향
- 인터넷의 다양성과 분산화를 증가시키는 기술에 대한 연구: 분산화 기술 및 분산 인터넷 구조 등

집중화와 관련한 실질적인 기술 논의는 인터넷의 장기적 비전을 연구하는 IRTF(Internet Research Task Force) 산하에 DINRG(Decentralized Internet Infrastructure Research Group)를 통해 이루어지고 있다[10]. DINRG는 신뢰(Trust), 식별(Identity), 네임(Name), 자원(Resource) 관리 등에 관련된 비집중형 인프라 서비스에 대한 연구 이슈를 논의하기 위해 2017년 신설되었다. DINRG는 인터넷이 기본적으로 분산 시스템으로 개발되었으나, 최근 중앙 집중화된 응용과 인프라 서비스로 인해 집중화가 지배적인 동작 모델로 등장하고 있다고 지적한다. 따라서 이런 집중화 모델이 문제가 될 수 있는 유즈 케이스(Use Case)들에 대한 연구가 필요하다는 것을 주요한 연구 배경으로 설명하고 있다. DINRG에서 고려하는 작업 범위는 분산형 유즈 케이스에 대한 조사, 글로벌 인터넷 관점에서 특정 유즈 케이스에 대한 해결 방안 논의, 기술적인 해결 방안 개발 및 문서화, 규모성 이슈와 미비된 기술 요소 도출을 위한 도구 및 메트릭(Metric) 개발 및 IETF에서의 향후 표준화 이슈 도출 등이다.

현재 DINRG에서 제안하고 있는 관련 연구 이슈로는 글로벌 규모에서 분산형 인프라 서비스 제공을 위한 규모성, 비집중형 환경에서의 신뢰 관리, 프라이버시, 분산 장부(Distributed Ledger)의 적용과 여러 가지 유즈 케이스와 관련된 기술, 인터넷 인프라 서비스에 관련된 특정 시나리오에서의 합의 알고리즘, 분산형 신뢰와 컴퓨팅 등을 고려하고

있다. DINRG에서 최근까지 다양한 관련 이슈들이 논의되고 있으며, 2019년에 논의된 주요 이슈는 다음과 같다.

- 분산 장부 기술에서 멀티캐스트 및 메시지 전달
- GNU 네임 시스템
- 대규모 분산형 식별 시스템에서의 비잔틴 합의 프로토콜
- 메쉬 네트워크에서 블록체인 기반의 인터넷 액세스 플랫폼
- BGP 검증을 위한 블록체인 기반의 테스트 베드
- NDN과 블록체인 기술

## 2. EU Next Generation Internet Initiative

유럽 위원회(European Commission)는 인터넷이 오늘날 유럽 경제의 디지털화를 위한 엔진이 되고 있으며 향후 10여 년 후 모든 사회 경제에 필수적인 요소가 될 것으로 전망하고 있다. 또한 인터넷의 이러한 가치는 글로벌성과 네트워크들의 네트워크(Network of Networks)로서의 분산성, 고장에 대한 견고함, 개방성, 협력성, 투명성을 통한 제한 없는 혁신 제공과 자발적인 데이터의 교환에 기반한다고 보고 있다[11].

그러나 최근 인터넷에서는 기업에 의한 개인 정보 모니터링 및 상업화에 대해 인터넷 사용자들의 불안이 증가하고 있어, 이로 인한 보안 이슈로 인해 온라인 활동을 제한하는 사용자가 늘고 있다는 것을 문제점으로 지적하고 있다. 또한 전 세계 데이터 중 일부만이 유럽 내에 저장되고 있으며, 유럽 기업들에 의한 빅데이터 이용률도 낮다는 점을 유럽이 해결하여야 하는 문제점으로 들고 있다. 이러한 문제점들은 모두 인터넷 데이터 집중화와



밀접한 관련이 있는 문제이다. 또한 기술적 환경 변화의 관점에서는 인공지능, 사물인터넷, VR/AR, 블록체인 등의 기술 등장으로 향후 15년 이내에 인터넷은 완전히 다른 모습을 가질 것으로 전망하고 있다.

이러한 인식을 기반으로 유럽은 차세대 인터넷(NGI: Next Generation Internet) 과제를 2016년부터 범유럽 연구개발과제로 시작하였다. 유럽은 이 NGI 과제를 통해 현재의 인터넷을 인간 중심(Human-centric)으로 재설계하여, 유럽이 지향하는 목표인 개방성(Openness), 포용성(Inclusivity), 투명성(Transparency), 프라이버시, 협력(Cooperation) 및 데이터 보호를 지원하는 플랫폼으로서의 미래 인터넷 개발을 목표로 하고 있다. NGI는 견고성(Resilience), 신뢰성(Trustworthy) 및 지속성(Sustainability)의 세 가지 설계 목표하에서 개발이 이루어지고 있다[12]. 견고성은 특정 회사나 네트워크상에서 어느 부분이 재해와 같은 이유로 동작 불능이 되는 경우에도 이로 인한 다른 부분에 대한 영향이 없어야 한다는 것을 의미한다. 신뢰성은 전체 사회와 경제 간의 상호 관계가 투명한 구조, 거버넌스 및 정책을 가져야 한다는 것이다. 즉 구조적으로 발언의 자유, 개인 기업 등을 보호하여야 하며, 글로벌 신뢰에 대한 어떠한 편향이나 시스템적인 오용을 막을 수 있도록 설계되어야 한다는 것이다. 또한 시민의 안전을 보장하고 국경을 가로질러 시장과 사회의 건강성 및 자율성을 강화해야 한다는 것이다. 지속성은 인터넷의 성공이 제한 없는 혁신을 제공한다는 점에 기반한 것으로 다양성과 분산화를 촉진하고 혁신적 발전을 위한 잠재력을 성장시키는 것이다.

이러한 비전하에 2018년부터 해당 과제에 대한 공지(Call)가 NGI Initiative라는 이름으로 시작되었다. NGI Initiative 이후에는 NGN 플래그십

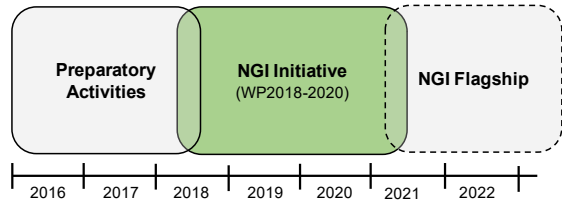


그림 2 유럽 NGI 프로젝트 로드맵

(Flagship) 과제들에 대한 공지가 예정되어 있다 [11]. 현재 NGI Initiative를 위한 공지는 ICT-24-2018-2019와 ICT-31-2018-2019가 대표적이다.

ICT-24-2018-2019 관련하여 2018년과 2019년 2번의 공지가 있었으며 2018년 4월, 2019년 3월에 과제 제안이 마감되었다. 각 공지에서 요구하는 주요한 연구 항목(Research and Innovation Actions)은 다음과 같다[13]. 먼저 2018년도에 공지된 연구 항목은 다음의 세 가지이다.

- 프라이버시 및 신뢰 향상 기술: 센서, 디바이스, 인공지능 기반 알고리즘 등이 디지털 환경에 도입됨에 따라 개인 데이터 공유 시 사용자에게 신뢰 증가와 제어성 향상을 지원하기 위한 기술
- 비집중형 데이터 거버넌스: 블록체인, 분산형 장부 기술, 개방형 데이터 및 피투피 기술에 기반한 개방형 하드웨어, 소프트웨어 에코시스템을 향상하기 위한 것으로 자율성, 데이터 주권 및 윤리적, 법적, 프라이버시 이슈를 포함
- 데이터 발견 및 식별 기술: 고도의 다양성을 가지는 데이터 소스, 서비스, 센서, 디바이스, 멀티미디어 콘텐츠 등을 검색하고 접근하기 위한 기술

2019년도에 공지된 세 가지 연구 항목은 다음과 같다.

- 전자 식별자(Electronic identifier)를 이용한 인터넷 신뢰 강화 기술: 인터넷에서 신뢰 강화를 위한 핵심적 이슈를 다루기 위한 것으로, 다양한 수준의 식별, 평판 및 신뢰를 가지는 인터넷 규모의 전자 식별자를 지원하기 위한 분산형 기술과 개인 데이터를 확인하고 평가하기 위한 새로운 비즈니스 모델을 위한 기반 제공
- 서비스와 데이터 이동성: GDPR하에서 개인 데이터 이동성 제공 이슈에 대응하기 위한 기술로, 최종 사용자에게 제공되는 서비스로부터 데이터의 분리를 포함
- 개방형 인터넷 구조 혁신: 더 나은 효율성, 규모성, 보안 및 견고성을 제공하도록 인터넷 구조의 진화를 보장하기 위한 개발자 커뮤니티 지원

ICT-31-2018-2019은 NGI를 위한 유럽과 미국 공동연구에 관련된 공지로 공동 워크숍과 같은 지원 활동, 펠로우십 프로그램을 포함한다. 연구 항목으로는 유럽과 미국의 실험 플랫폼상에서 NGI의 새로운 연구 이슈에 대한 공동 실험을 포함하고 있다.

### 3. 분산 웹(Decentralized Web)

웹은 인터넷상에서 가장 널리 사용되는 응용플랫폼으로 일반인들에게는 흔히 인터넷과 동일한 개념으로 인식되기도 한다. 웹은 원거리의 다른 컴퓨터에 저장된 문서를 즉각적으로 이용할 수 있게 해 주는 장벽 없는 글로벌 네트워크의 하나로 간주된다. 글로벌 연결성을 제공하는 인터넷을 기반으로 웹은 글로벌 규모로 문서뿐만 아니라 다른 형태의 자원도 공유하게 해 준다.

Tim Berners Lee에 의해 개발된 웹은 분산성(Decentralization)과 보편성(Universality)의 두 가지 원칙에 기반하여 설계되었다고 할 수 있다[14]. 분산성은 웹에 자료를 포스팅하는 것에 대해 어떤 중앙 권한자의 허가를 필요로 하지 않으며, 중앙 제어 노드가 없어 단일 고장점(Single Point of Failure)이 존재하지 않는 것을 의미한다. 보편성은 누구나 어떤 내용이라도 웹상에 출판할 수 있으며, 사용자의 하드웨어나 위치, 문화, 정치적인 신념과 무관하게 서로 간 동일한 언어로 모든 컴퓨터가 통신할 수 있어야 한다는 것을 의미한다.

그러나 현재 웹에서의 이러한 원칙을 깨뜨리는 방향으로 많은 회사가 상업적 성공을 거두고 있다. 이는 중앙집중화된 서버가 고객에게 서비스를 제공하는 것뿐만 아니라 새로운 비즈니스 모델을 위한 효율적인 데이터 처리에서도 장점을 가지기 때문이다. 또한, 고유한 API나 데이터 포맷을 사용하는 것은 다른 후발 기업과의 경쟁에서 도움이 된다. 이러한 추세에 따라 데이터 자원의 폐쇄적인 사일로(Silo)가 기업의 주요한 자산이 되고 있다. 예를 들어, 전자상거래에서 상품의 구매자와 판매자는 독점적 지위를 이용하는 소수의 중앙집중형 중개자들이 지시하는 조건에 따라야 한다. 또한, 데이터 프라이버시 관점에서 이러한 독점적 중개자는 개인 정보의 과도한 보유로 데이터 주권에 대한 중대한 위협이 될 수 있다.

Tim Berners Lee가 제안한 링크드 데이터(Linked Data) 기술은 문서가 아닌 데이터에 대한 표준을 제공하여 보편성을 지원하고, 다른 원격 기기에 저장된 데이터들을 통합 처리할 수 있는 기술 제공을 통해 분산화를 지원함으로써 웹의 기본 설계 원칙을 회복하고자 하는 시도라고 할 수 있다. 이러한 링크드 데이터를 더욱 발전시켜 분산화된 소셜 웹(Social Web)을 구현하고자 하는 것이 최근 Tim

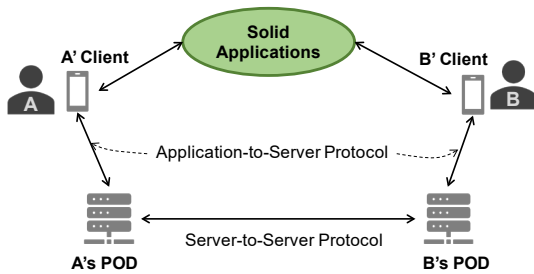


그림 3 솔리드 플랫폼 구조

Berners Lee가 MIT에서 개발하고 있는 솔리드(Solid: Social Linked Data) 플랫폼이다[15].

솔리드 플랫폼에서 각 사용자는 데이터를 중개자의 사일로가 아닌 개인 저장소인 POD(Personal Online Datastore)에 자신의 데이터를 저장한다. 이 개인 저장소는 개인 기기나 저장 서비스를 제공하는 사업자의 저장소일 수 있다. 솔리드는 이러한 분산화된 데이터를 기반으로 분산화된 인증 및 접근 제어를 제공하며, 표준화된 데이터 접근 메커니즘을 지원함으로써 분산화된 소셜 웹을 구현하고자 한다[16].

분산 웹에 대한 또 다른 주목할만한 활동은 분산 웹 서밋(Decentralized Web Summit)이다[17]. 이 서밋은 웹 설계자, 엔지니어, 학자, 언론인 등과 같은 다양한 분야의 웹 이해관계자들이 분산 웹 구축을 위해 필요한 기술들과 영향을 논의하기 위한 장으로 만들어졌다. 2016년, 2018년 두 번의 회의를 개최하였으며, 하이퍼링크 창시자인 제나두 프로젝트의 Ted Nelson, 웹 창시자인 Tim Berners Lee, 인터넷 창시자인 Vinton Gray Cerf 등 관련 주요 인물들이 모두 참석하여 분산 웹을 위한 주요 이슈를 논의하였다. 2016년 회의에서는 현재의 웹이 프라이버시 침해와 검열의 위협성을 가지고 있으며, 인류의 문화 자산을 영구적으로 저장하기에는 용량이 부족하다는 문제를 공유한다. 따라서 개방적이

며 데이터를 분산 저장하고 처리함으로써 검열에 자유로우며 중앙집중적 제어 없이 전 세계의 수백만 대의 컴퓨터를 연결할 수 있는 분산 웹에 대한 비전을 제시하였다. 이러한 비전하에 현재 구축되고 있는 IPFS, Blockstack, ZeroNet, COALA 등과 같은 다양한 분산형 응용에 대한 소개와 현재까지의 진행 상황들에 대한 발표가 이루어졌으며 실제로 분산 웹 기술을 활성화시키기 위한 방안에 대한 논의가 이루어졌다. 2018년 회의를 통해서서는 실제로 분산 웹을 확장하기 위한 전략에 대한 논의가 주로 이루어졌다. 즉 어떻게 수백만 명의 사람들이 실제로 분산 웹을 사용하게 할 것인가, 어떤 코드가 작동하고 어떤 코드가 여전히 누락되어 있는가, 분산 웹의 활성화에 장애가 되는 문제를 해결하기 위해 어떤 협력이 필요한가에 대한 논의가 주로 이루어졌다.

### III. 결론

본 고에서는 4차 산업혁명을 위한 디지털 변환의 기반 인프라인 인터넷에서의 데이터 집중화 문제를 살펴보고 집중화 문제에 대응하는 인터넷 소사이어티와 유럽 등에서의 관련 연구 동향을 정리하였다.

빅데이터와 인공지능 기반의 4차 산업혁명 시대에 데이터는 현시대의 석유만큼이나 소중한 자원이라고 할 수 있다. 다가올 데이터 중심 사회에서 데이터 주권 확보는 데이터의 상업적 가치뿐만 아니라 개인 정보의 보호라는 측면에서도 가장 중요하게 고려되어야 할 사항이다. 따라서 4차 산업혁명의 주도국으로 발돋움하고자 하는 한국의 입장에서든 현재의 소수 기업에 의한 데이터 집중화 문제를 해결하기 위한 기술에 대한 선행 연구가 필수적이라고 할 수 있다. 현재 인터넷 기술은 실질적



으로 미국이 주도하고 있으며 유럽도 NGI 프로젝트에서 볼 수 있는 바와 같이 미래 사회 인프라 기술 확보를 위해 큰 노력을 기울이고 있음을 알 수 있다. 이러한 상황에 대한 효율적인 대응을 위해서는 한국이 경쟁력을 가질 수 있는 적절한 연구개발 모델이 수립되어야 할 것이다.

인터넷 데이터 집중화 문제에 대한 대응 기술 개발의 첫 단계는 미국과 유럽에서 이루어지고 있는 최신 연구에 대한 분석과 이를 바탕으로 한 관련 기술 프레임워크 개발이라고 할 수 있다. 개발 프레임워크를 기반으로 한국이 확보 가능한 기술 목록 도출 및 산학연 협력 연구전략 개발이 이루어져야 할 것으로 보인다. 연구개발 전략의 관점에서 글로벌 인터넷 표준 기술을 독자적으로 주도하기 어려운 한국의 입장에서 볼 때 현재 미국, 유럽에서 이루어지고 있는 관련 연구 활동에 공동연구 형태로 적극적으로 참여하는 것이 하나의 방안이 될 수 있다. 국가적 차원에서도 본 고에서 기술한 현재 미국과 유럽 간의 NGI 공동연구와 같이 미국, 유럽과의 국제 공동연구 프로그램 개발이 이를 위한 효율적인 방안이 될 수 있을 것이다.

**용어해설**

**인터넷(The Internet)** 다양한 네트워크가 존재하는 환경에서 글로벌 범위의 종단 간 연결을 지원하는 하는 TCP/IP 기반의 인터넷 워킹 인프라. 분산 네트워크 및 종단 간 통신을 원칙으로 하며 다양한 네트워크 간을 연결하는 네트워크라는 측면에서 흔히 네트워크들의 네트워크(Network of Networks)로 불림

**월드와이드웹(World Wide Web)** 다양한 컴퓨터 간 문서 공유를 지원하는 인터넷상의 대표적인 응용. 보편성과 분산성을 원칙으로 하며, 정보를 요청/응답 기반으로 교환하기 위한 http 프로토콜과 문서의 링크를 표현하는 하이퍼텍스트로 작성되는 html 언어로 구성됨

**약어 정리**

|       |                                   |
|-------|-----------------------------------|
| AR/VR | Augmented Reality/Virtula Reality |
| CDN   | Content Delivery Network          |

|       |  |
|-------|--|
| DINRG | Decentralized Internet Infrastructure Research Group |
| GDPR  | General Data Protection Regulation                   |
| IAB   | Internet Architecture Board                          |
| ISOC  | Internet SOCIety                                     |
| NGI   | Next Generation Internet                             |
| POD   | Personal Online Datastore                            |
| SOLID | SOcial LIInked Data                                  |

**참고문헌**

- [1] A. K. Molavi et al., "Taking a Long Look at QUIC: An Approach for Rigorous Evaluation of Rapidly Evolving Transport Protocols," in Proc. Internet Measurement Conf., London, UK, Nov. 2017, pp. 290-303.
- [2] <https://www.internetworldstats.com/stats.htm>.
- [3] Internet Health Report 2018, <https://internethealthreport.org/2018/>.
- [4] J. Arkko et al., "Considerations on Internet Consolidation and the Internet Architecture," IRTF Internet draft draft-arkko-iab-internet-consolidation-02, July 09, 2019.
- [5] The Global Internet Phenomena Report, <https://www.sandvine.com/hubfs/downloads/phenomena/2018-phenomena-report.pdf>, October 2018.
- [6] The data economy: The world's most valuable resource, The Economist, pp. 14-17, May 2017.
- [7] World Economic Forum, "Will the Future Be Human? Yuval Noah Harari," Jan. 2018, <https://youtu.be/hL9uk4hKyq4>
- [8] EU GDPR homepage, <https://eugdpr.org>.
- [9] Internet Society Global Internet Report, "Consolidation in the Internet Economy," 2019.
- [10] IRTF DINRG homepage, <https://datatracker.ietf.org/rg/dinrg/about/>.
- [11] O. Bringer, "The Next Generation Internet - An Internet of Humans," EU NGI forum, 2018.
- [12] NLnet Foundation and Gartner, "Next Generation Internet 2025: A study prepared for European Commission DG Communication Networks, Content & Technology," 2018.
- [13] EU Next Generation Internet Initiative, <https://ec.europa.eu/digital-single-market/en/next-generation-internet-initiative>.
- [14] L.-D. Ibanez et al., "Redecentralizing the Web with Distributed Ledger," IEEE Intel. Syst., vol. 32, no. 1, 2017, pp. 92-95.
- [15] Solid homepage, <https://solid.mit.edu>.
- [16] E. Mansour et al., "A Demonstration of the Solid Platform for Social Web Applications, WWW'16 Companion," in Proc. Int.

conf. Companion World Wide Web. Montreal, Canada, Apr. 2016, pp. 223-226.

[17] Decentralized Web Summit, <https://www.decentralizedweb.net>.