

---

# 공공 데이터 분석결과의 게임화를 통한 웹 방문자 성향 분석

## Analysis of Web Visitor Behavior using Public Data Analysis and Gamification

허종욱 (Jong-Uk Heo)\*, 온병원 (Byung-Won On)\*\*, 이만재 (Manjai Lee)\*\*\*

---

**요약** 빅 데이터의 등장과 함께 공공기관이 보유한 데이터를 활용하여 국민에게 개선된 서비스를 제공하고자 하는 정책이 탄력을 받고 있다. 일반 국민이 관심을 갖는 국회의원의 본회의 표결 데이터를 기반으로 국회의 의정활동을 파악하고자 하는 폴리티즈라는 웹 사이트를 개발하였다. 웹 사이트의 기능 중 하나는 실제 국회에서 처리된 법안에 대해 웹 사용자에게 찬반을 묻도록 하였으며 자신과 가까운 성향을 갖는 의원을 알 수 있도록 게이미피케이션을 적용하였다. 방문자의 투표결과를 통해 해당 법안에 대한 국민의 생각을 알 수 있음을 발견하였다. (<http://www.politiz.org>에서 서비스 중이며, 구글크롬, IE8.0이상의 웹브라우저에서 사용가능함)

### Abstract

The utilization of public data based on Big data analytics would contribute to the development of transparent public policy. As one of case studies, we developed the so-called Politiz web site in which we can understand congressional activities using roll call votes in the 18<sup>th</sup> Korean National Assembly. In the web site, a user can vote sixteen controversial bills and finally find the best match showing whom the user is close to among all congressmen. In our research, we showed a congressional social network that provides us with the holistic insight mined from the roll call data based on Big data analytics. In particular, we focus on studying how to apply the concept of gamification so as to make users more fun and engaging in the Politiz web site. In addition, our web site provides a handy tool for collecting public opinion about controversial issues occurring in our society.

**핵심어:** 게이미피케이션, 공공 데이터, 비주얼라이제이션, 성향분석

---

본 연구는 재단법인 차세대융합기술연구원의 연구과제(2011-P3-11)의 지원을 통하여 연구되었음.

\* 카이스트 웹 사이언스 석사과정; e-mail: juheo@mmc.kaist.ac.kr

\*\* 교신저자 : 서울대학교 차세대융합기술연구원 연구교수; e-mail: bwon@snu.ac.kr

\*\*\*서울대학교 차세대융합기술연구원 특임연구위원; e-mail: manjai@snu.ac.kr

## 1. 서론

공공 데이터는 정부나 지방자치 단체 또는 산하기관이 보유한 공공 목적으로 사용할 수 있는 데이터를 말한다. 주민 등록 데이터, 소득 관련 데이터, 차량 보유 데이터 등 일반 국민이 생활과정에서 발생하는 다양한 데이터는 공공 데이터로 볼 수 있다. 미국, 영국 등 선진국에서는 공공 데이터를 공개하고 이를 통해 새로운 서비스를 제공하거나 기존의 정책을 개선하고자 하는 open data 운동을 추진 중에 있으며 우리나라의 경우에도 정부나 지방자치단체를 중심으로 데이터의 공개와 이를 이용한 서비스 개발을 추진하고 있다 [1].

공공 데이터는 용량이 크며 데이터의 구성이 복잡하기에 빅 데이터 어널리틱스와 같은 분석과정을 거쳐야 의미 있는 내용을 파악할 수 있다. 공공 데이터 분석결과는 수치나 문장으로 결과를 보여주기보다는 인터랙티브 비주얼라이제이션을 통해 일반 사용자에게 전달된다. 사이트의 목적이 일반 시민에게 정부의 정책을 홍보하고자 할 경우 흥미를 돋는 내용이 있어야 원래의 목적을 달성할 수 있다. 이러한 문제를 해결하는 방법으로 게임 요소를 웹 사이트에 포함하려는 방법이 제시되고 있으며 이를 게이미피케이션(gamification)이라고 한다.

본고에서는 공공 데이터의 분석결과를 인터랙티브 비주얼라이제이션으로 구현한 웹 사이트 개발과 관련하여 게이미피케이션을 도입하는 과정과 이를 통하여 방문자의 성향을 파악하는 방법을 제시한다.

## 2. 관련 사례

공공 데이터를 비주얼라이제이션으로 보여주는 여러 웹 사이트들이 있다. 유럽의 에너지 문제를 CO2 배출량, 에너지 재사용률, 에너지 의존도 등 다양한 방식으로 제공한다 [2]. 영국의 국민으로부터 받은 세금이 어떻게 사용되는가를 인터랙티브 방식으로 구현한 페이지도 좋은 예이다[3]. OECD 국가들의 행복도를 주택, 치안, 교육 등 여러 요소로 구분하여 각 요소에 가중치를 부여할 경우 어느 나라가 가장 살기 편한 나라인가를 보여 주기도 한다[4]. 이러한 웹 사이트는 우수한 비주얼로 방문자에게 좋은 반응을 얻을 수 있으나 댓글과 같은 직접적인 방식을 사용하기 전에는 방문자의 의견을 받아들이기 부족한 점이 있다.

게임을 교육에 사용하려는 노력은 게이미피케이션이라는 용어가 등장하기 이전부터 에듀테인먼트라는 용어와 함께 사용되었다. 게이미피케이션이라는 용어는 교육 외에도 마케팅, 환경, 건강 등 여러 가지 목적에 따라 사용되는 웹 사이트에 게임 요소를 추가하면서 강조되어 왔다. 게이미피케이션의 대표적인 사례로 손꼽히는 어플리케이션으로는 Foldit을 들 수 있다[5]. 2008년 워싱턴 대학에서 개발된 이 어플리

케이션은 단백질의 구조를 찾기 위해 개발된 게임으로 단백질 구조를 어떻게 잘 접는가를 점수와 연동하도록 하여 참가자의 경쟁심을 북돋우고 있으며 이를 이용하여 단백질 구조를 발견하는 데 실제로 사용되고 있다.

## 3. 웹 사이트 구현과 게임화

우리나라의 경우 정부나 공공기관이 공개한 데이터가 빈약하여 영향력이 있는 공공 데이터 어플리케이션을 개발하는 것이 쉽지 않다. 이미 공개된 데이터 중 국민의 관심을 얻을 수 있는 데이터로는 국회의 표결데이터가 있다. 표결 데이터는 국회 회의록에 포함되어 있으며 이를 기반으로 데이터 분석이 가능하다. 폴리티즈라고 명명된 어플리케이션은 18대 국회의 본회의 투표 데이터로부터 의원 간의 유사도를 계산하고 이를 바탕으로 개별의원의 노드를 배치하는 방식으로 의원 간의 유사도를 비주얼로 보여주고 있다[6].

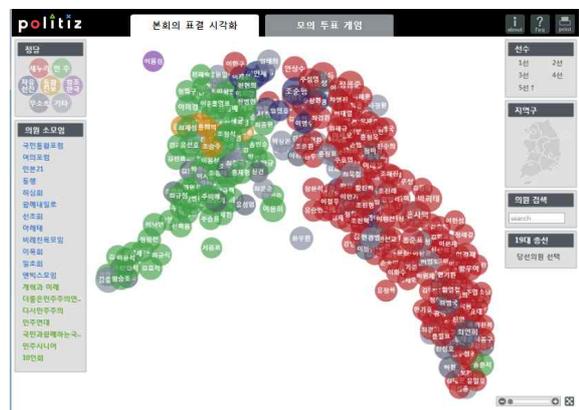


그림 1 폴리티즈 웹 사이트

그림 1에서 보는 바와 같이, 폴리티즈 웹사이트의 핵심은 가까운 위치에 배치되어 있으면 의원의 투표성향이 유사하다는 것을 알려주는 것이다. 여기에 의원이 소속된 정당, 비례대표, 여러 번 선출된 의원의 선수, 출신 지역구 등의 정보를 함께 제공한다. 웹 사이트를 방문자가 직접 참여할 수 있는 게임 요소를 추가하도록 하였다.

방문자의 호응을 얻기 위해서는 의원이 실제 참여한 법안에 대해 찬반을 입력하고 결과적으로 가장 가깝게 투표한 의원을 보여주는 방식을 사용하자는 모의 투표 아이디어가 도출되었다. 18대 국회의 경우 2000건이 넘는 표결안이 상정되었으나 대부분의 표결에서 반대 의견은 극소수이며 반대 의견은 불참으로 표시하는 경우가 대부분이다. 따라서 찬반 의견의 비율이 높은 쟁점법안을 먼저 선정하고 이 중에서 안보, 복지, 재정, 환경 등 다양한 분야의 법안을 포함한 16개 법안을 선정하였다(표 1 참조). 쟁점법안 중 천안함, 세종시와 관련된 2건의 법안은 부결 처리된 내용으로 가결과 부결의 데이터를 포함하여 균형을 맞추도록 하였다.

표 1 쟁점법안

번호	일시	법안	결과
1	2010년 5월	G20 정상회의 성공 개최를 위한 경호 관련 특별법안	수정가결
2	2011년 6월	변호사시험법 개정	원안가결
3	2009년 3월	독점규제 및 공정거래에 관한 법률 개정	원안가결
4	2011년 9월	도시개발법 일부 개정	원안가결
5	2010년 10월	지방세특례제한법 일부 개정	원안가결
6	2009년 4월	국민건강보험법 일부 개정	원안가결
7	2011년 11월	한미 FTA 이행을 위한 관세법 특례에 관한 법률 일부 개정	수정가결
8	2010년 2월	국군부대의 아프가니스탄 파견 동의안	원안가결
9	2011년 11월	약사법 일부 개정	원안가결
10	2010년 6월	북한의 천안함에 대한 규탄 및 대응조치 결의안에 대한 수정안	부결
11	2010년 12월	4대강 사업 관련 친수구역 활용에 관한 특별법안	원안가결
12	2010년 6월	신행정수도 건설을 위한 특별법 전부 개정	부결
13	2009년 4월	저작권법 일부 개정	수정가결
14	2011년 4월	청소년보호법 일부 개정	원안가결
15	2009년 1월	언론중재 및 피해구제 등에 관한 법률 일부 개정	원안가결
16	2009년 7월	방송법 일부개정법률안 2차 투표	수정가결

선정된 법안에 대해서는 왜 법안에 찬성하는가와 반대하는가에 대한 설명을 제시하여 참가자로 하여금 선택에 도움을 줄 수 있도록 하였다. 예를 들어 “G20 정상회의 성공 개최를 위한 경호 관련 특별법안”의 경우 행사를 원활하게 추진하기 위해 특정장소에서 집회와 시위의 자유를 일시적으로 제한하여야 한다는 찬성의견과 집회의 시위의 자유는 어떠한 경우라도 제한되어서는 안 된다는 반대의유를 함께 제공하였다.

법안의 선정과 함께 문제가 된 것은 모든 의원을 대상으로 할 것인가의 문제였다. 모든 의원을 대상으로 할 경우 화면에 표시하기가 어렵고 모의실험을 할 경우 특정 의원만 자주 등장할 확률이 높기 때문에 모든 의원을 대상으로 하기보다는 특정 조건에 맞는 의원의 세트를 먼저 정하고 이를 대상으로 게임을 진행하도록 하였다.

의원의 세트를 구성하는 방법으로는 정당, 선수, 지역구, 출신 고등학교, 출신 대학교, 출생지역, 성별, 연령, 재산규모, 경력을 사용하였다. 사용자가 제시된 조건 중 일부를 입력하면 조건에 맞는 의원의 수를 볼 수 있으며 의원의 수가 8명에서 16명 사이의 의원이 선정될 경우에만 게임을 시작할 수 있도록 하였다. 이는 게임의 복잡도를 적정 수준으로 맞추기 위해서였다. 의원의 선택과정이 너무 복잡하다고 생각하는 사용자를 위해 임의의 의원을 대신 선정하는 기능을 포함하였다(그림 2 참조).

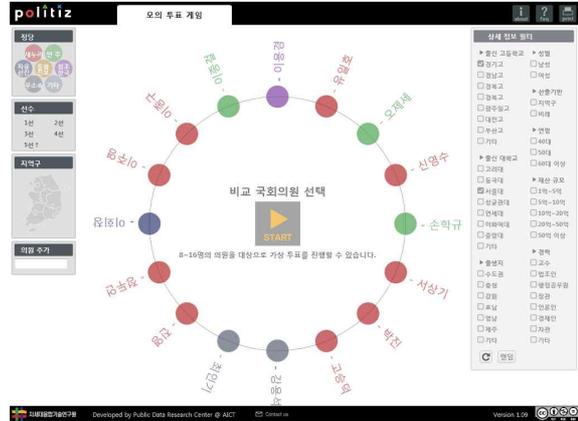


그림 2 관심의원 선택과정

후보 의원이 선정된 다음 사용자는 제시되는 법안에 대해 찬성, 반대, 모름의 셋 중 하나를 선택하도록 하였다. 그림 3에서 보는 바와 같이, 매 선택의 결과는 화면에 보이나 의원은 당적만을 표시할 뿐 실명을 보여주지 않도록 하여 게임의 금급증을 높이도록 하였다. 이 때 사용자(a)와 후보 의원(b)간의 유사도는

$$sim(a,b) = \frac{x}{y}$$

로 정의하였으며 분모 y는 사용자의 모의 투표 회수이며 x는 a와 b가 찬성, 반대의견이 같은 회수를 말한다. 유사도는 16건의 투표가 진행되는 중간 지속적으로 업데이트되며 결과를 그림4와 같이 보여준다.



그림 3 사용자의 모의 투표 과정

16건의 투표자를 끝내면 사용자와 가장 가깝게 투표한 세 의원을 시상식 형태로 중앙에 배치하고 나머지 의원은 사진만으로 의원을 표시하였다.

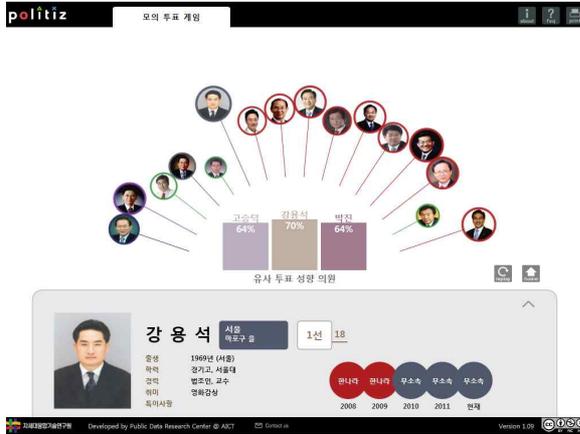


그림 4 모의 투표 결과 표시 예

자신과 가장 가까운 의원이나 다른 의원의 정보를 추가적으로 보고자 할 경우 해당의원을 클릭할 경우 당직 변화, 재산, 투표 비용과 함께 위키백과에 기록되어 있는 개인 문서를 보여준다. 일반 사용자는 의원에 대한 객관적인 데이터를 보기를 원할 것으로 추정되어 자신을 홍보하는 의원의 웹사이트 대신 위키백과를 연결하여 중립성을 지키도록 하였다.

#### 4. 결과 분석

폴리티즈 웹 사이트는 2012년 3월 20일 공개되었다. 웹 사이트 구현을 위해 의원의 데이터베이스를 구축하였으며 MySQL, PHP와 같은 도구를 사용하였다. 웹 사이트의 비주얼은 일반적으로 널리 사용되는 플래시 대신 HTML5의 캔버스 기능을 사용하였다. 미래의 웹 사이트는 모든 웹 브라우저가 지원하는 HTML5로 통합될 것이기에 개발의 난이도는 높았으나 모든 브라우저를 사용할 수 있다는 장점이 있다.

웹 사이트 방문기록은 게이미피케이션의 새로운 가능성을 보여 주었다. 방문자의 IP 주소와 함께 모의 투표에 대한 후보의원 데이터와 찬반 데이터를 확보할 수 있기 때문이다. 모의 투표 기록의 일부는 웹 사이트를 테스트하기 위한 데이터로 포함되어 있기에 확보된 데이터 중 개발자와 베타 테스터의 데이터는 제외하였다. 또한 모든 투표에 무조건 찬성, 또는 반대를 표시한 데이터는 삭제하고 나머지 데이터를 활용하여 투표데이터를 분석한 결과는 표 2와 같다.

천안함 관련 법안의 경우 야당의 안전이 부결된 것을 본다면 국민건강보험과 관련된 법안을 제외하고는 사이트 방문자의 야권 성향이 높은 것을 볼 수 있다. 본 연구를 시작할 당시 게이미피케이션이라는 개념은 부수적인 것이었으나 이러한 기능을 활용한다면 사이트 방문자의 정치현안에 대한 데이터 수집에 활용할 수 있음을 확인하였다.

표 2 폴리티즈 웹사이트 방문자들의 투표 결과

법안	법안	찬성	반대	기권
1	G20 정상회의 성공 개최를 위한 경호 관련 특별법안	43%	55%	1%
2	변호사시험법 개정	33%	53%	15%
3	독점규제 및 공정거래에 관한 법률 개정	26%	66%	8%
4	도시개발법 일부 개정	40%	44%	16%
5	지방세특례제한법 일부 개정	29%	61%	9%
6	국민건강보험법 일부 개정	61%	26%	14%
7	한미 FTA 이행을 위한 관세법 특례에 관한 법률 일부 개정	36%	54%	10%
8	국군부대의 아프가니스탄 파견 동의안	47%	43%	10%
9	약사법 일부 개정	31%	58%	11%
10	북한의 천안함에 대한 규탄 및 대응조치 결의안에 대한 수정안	57%	35%	9%
11	4대강 사업 관련 천수구역 활용에 관한 특별법안	28%	66%	6%
12	신형정수도 건설을 위한 특별법 전부 개정	39%	50%	11%
13	저작권법 일부 개정	36%	57%	7%
14	청소년보호법 일부 개정	30%	64%	6%
15	언론중재 및 피해구제 등에 관한 법률 일부 개정	44%	46%	10%
16	방송법 일부개정법률안 2차 투표	27%	67%	6%

#### 5. 결론

공공 데이터를 활용하기 위한 하나의 예로 국회의원의 본회의 투표결과를 분석하고 결과를 비주얼라이제이션 한 폴리티즈라는 웹 사이트를 소개하였다. 의원 간의 투표성향 그래프를 보여주는 것 외에 모의 투표라는 개념을 도입한 과정을 제시하였다. 게이미피케이션은 사용자의 관심을 끌 수 있는 유효한 수단이며, 모의투표는 자신의 속내를 쉽게 드러내지 않는 방문자의 의견을 읽을 수 있는 수단으로 사용될 수 있음을 발견하였다.

- [1] 이만재, "빅 데이터와 공공 데이터 활용," Internet and Information Security, vol.2 no.2 pp.1-21, 2011.
- [2] Open Knowledge Foundation, "Europe's energy," <http://energy.publicdata.eu/ee/vis.html>, 2008.
- [3] Open Knowledge Foundation, "Where does my money go?," <http://wheredoesmymoneygo.org/dashboard/>, 2010.
- [4] OECD, "Better life initiative, your life index," <http://www.oecdbetterlifeindex.org/>, 2011.
- [5] UW Center for Game Science, Foldit, <http://foldit.it/portal/>.
- [6] 이만재, 온병원, "빅 데이터 접근방식의 공공 데이터 비주얼라이제이션 사례," 정보와 통신, vol.29 no.11 pp.11-17, 2012.