

인터넷 커뮤니티에서의 익명성과 사용자의 단어사용 사이의 관계 분석 연구

황 병 화*, 이 은 규°

Analysis of the Impact of Anonymity on Members' Use of Words in Online Communities

Byeonghwa Hwang*, Eun-Kyu Lee°

요 약

정보화 시대가 도래하면서 현실에 존재하던 여러 요소가 인터넷 세상으로 옮겨지게 되었고, 동호회나 동아리 등의 모임도 “인터넷 커뮤니티”를 통해 나타나고 있다. 인터넷은 기본적으로 익명성을 특징으로 한다. 즉, 인터넷 커뮤니티에서 사용자는 실제의 자신을 숨길 수 있고, 사이버 세계에서만 표현되는 또 다른 나를 통해 현실과 다른 성격이 표출될 수 있다. 기존의 연구에서는, 이러한 익명성에 기반한 공격성 및 폭력성을 근거로 하여 인터넷이 사회에 미치는 부정적인 영향을 설명하였다. 과거 인터넷 사용자들이 단순히 나를 숨기기 위해 익명성이라는 가면으로 부정성을 발현시키기도 했다. 그러나, 오늘날에 와서는 이 가면이 누군가에 의해 밝혀질 수도 있다는 사실이 자리 잡았기 때문에, 익명성의 본질인 “나”를 숨기는 것에 초점을 맞추게 되었다. 즉, 익명성의 은닉성은 예전이나 지금이나 가면을 내세우기 때문에 사용자의 의지에 따라 공격성이 나타난다는 특징이 있지만, 공격성은 결국 부정적인 영향을 가져오기 때문에 더욱 주목 받을 수밖에 없음을 알 수 있다. 따라서, 본 연구에서는 인터넷의 익명성 특징에 따라 커뮤니티 사용자의 단어 사용 행태가 어떤 차이를 보이는지 알아본다. 실험 결과를 통해, 은닉성만을 나타내고자 하는 사용자들이 존재한다는 것과 이들의 단어 사용이 다른 익명성의 특징과 별다른 차이를 보이지 않는다는 점을 보인다.

키워드 : 익명성, 은닉성, 인터넷 커뮤니티, 단어 사용, 디시인사이드, 정보보호, 데이터처리

Key Words : Anonymity, concealability, online community, use of words, DCinside, privacy, data processing

ABSTRACT

With the advancement of the information technology, various “Things” that existed in reality have been transferred to the Internet world. Social meetings and clubs are also appearing through the “Internet community”. The Internet is basically characterized by anonymity. That is, users in the community can hide who they are, and a personality different from reality can be expressed through another me expressed only in the cyber world. Previous studies have tried to explain the negative impact of the Internet on society based on aggression and violence behind the anonymity. In the past, Internet users showed their negativity with a mask of anonymity simply to hide themselves. However, today, since it has been understood that this mask can be revealed by someone, the focus has been on hiding “I,” the essence of anonymity. In other words, the

※ 본 연구는 인천대학교 2024년도 자체연구비 지원으로 수행되었습니다.

• First Author : Incheon Nat'l University, Dept of Info. & Telecomm. Eng., esy3770@inu.ac.kr, 학생회원

° Corresponding Author : Incheon Nat'l University, Dept of Info. & Telecomm. Eng., eklee@inu.ac.kr, 정회원

논문번호 : 202407-140-D-RN, Received July 10, 2024; Revised August 27, 2024; Accepted September 9, 2024

concealment of anonymity is characterized by aggression appearing according to the will of the user because the mask is put forward, in the past and now, but it can be seen that aggression inevitably attracts more attention because it eventually brings a negative effect. Therefore, in this study, we examine the difference in word use behavior of community users according to the anonymity characteristics of the Internet. Experimental results have shown that there are users who only want to show concealment and that their use of words does not show much difference from other anonymity characteristics.

I. 서 론

정보화 시대가 발전함에 따라 현실에 존재하는 많은 것들이 가상의 인터넷 공간으로 옮겨지고 있다. 예를 들어, 현실의 화폐를 가상화한 암호화폐나 현실의 저작권과 같은 개념을 가상 세계로 옮겨놓은 NFT (Non-Fungible Token), 현실에서와 같은 경험을 주는 메타버스 등이 해당한다. 현실의 동호회는 가상의 세계인 인터넷 커뮤니티로 나타나고 있다.

인터넷 커뮤니티는 CMC(Computer Mediated Community)^[1]라는 이름으로 알려져 있으며, 컴퓨터를 매개로 하는 커뮤니티를 뜻한다. 한국에서는 “인벤”과 “디시인사이드”, “네이버 카페” 등이 대표적인 인터넷 커뮤니티이다. 인터넷 커뮤니티는 마치 화장실 낙서와 같아서^[2] 저속하고 저급한 말들이 오갈 수 있고, 난폭하고 무법지대와 같은 생명력을 갖고 있어서 다양한 성격의 웹사이트가 나타날 수 있다^[3]. 이들이 나타내는 교양이나 질서와 거리가 멀어 보이는 모습 때문에 현실에서는 인터넷 커뮤니티에 좋지 않은 시각을 가질 수 있다. 그럼에도 불구하고, 많은 사람들이 인터넷 커뮤니티를 이용하고 있는데, 이는 인터넷 커뮤니티 사용자들이 현실에서 잘 나타나지 않는 것은 익명성에 기인하고 있기 때문이다.

현실에서 우리는 이름을 통해서 나를 드러내며 살아가고 있다. 누군가와 소통을 시작할 때 나의 이름이 제시되고 그 누군가의 제시된 이름을 통해서 상대를 인식한다. 즉, 이름은 나를 정의하는 첫 단추인 것이다. 그러나, 인터넷에서는 이름을 감출 수가 있는데, 이를 “익명성”이라고 한다. 현실에서의 나를 숨기고 또 다른 가면을 통해 세상과 소통하는 것이 익명성의 특성인 “은닉성”이다.

과거에는 당시 뉴미디어였던 인터넷이 갖는 부정적인 영향의 원인을 익명성에서 찾았다. 나를 숨길 수 있는 특징이 있어 내면에 품고 있는 공격성을 발현할 수 있다는 것이었고, 이는 문제를 해석하는 중요한 열쇠로 풀이되었다. 하지만, 익명성의 존재 의미는 바로 은닉성이다. 정보의 바다라고 할 수 있는 인터넷 세상에서 나

의 존재를 알리고 싶어 하지 않는 특성의 사용자는 이 은닉성에 기인하고 있으며, 여기서 공격성으로 이어지는 것은 오로지 사용자의 선택에 달려있다. 현대 사회에서 인터넷이 자리 잡기 시작하고, 더 이상 새로운 미디어가 아닌 상식이 되어버린 인터넷 세상에서 사용자들은 더 이상 익명성이 완전하지 않다는 사실을 알게 된다. 인터넷에서 범죄와 같은 악의를 표출하는 일이 발생하면, 경찰과 같은 수사기관의 추적을 통해 그 범인을 잡아낼 수 있는 여러 상황을 경험하게 되며, 과거에 비해서 내 익명성이라는 가면이 벗겨질 수 있다는 사실이 알려지게 되었다. 일부 인터넷 사용자들은 본인을 추적할 수 없는 완전한 익명성을 찾게 되는데, 이것이 TOR(The Onion Ring) 웹 브라우저와 같은 형태로 등장하게 되었다.

TOR는 웹사이트에 사용자가 도착할 때까지의 경로를 다양한 장소를 거치게 만들고, 그 경로를 암호화하여 추적을 어렵게 만들기 때문에, 지금까지의 익명성과는 차원이 다른 특징을 보여준다. 완전한 익명성의 등장에 따라서 불법, 범죄와 같은 악의적인 공격을 하고자 하는 사용자들은 당연하게 이 뒤에 숨게 되었고, 현재도 이러한 공격성을 나타내는 사용자들이 존재함이 여러 사건과 보도를 통해 알려져 있다. 다만, 공격성을 나타내는 사용자는 그 특성 때문에 은닉성만을 나타내는 사용자보다 더 눈에 띄는 것뿐이고, 개방된 공간에서 나를 드러내고 싶지 않아 하는 사용자들의 은닉성은 이 완전한 익명성에서 여전히 작동하고 있다.

인터넷 커뮤니티에는 중앙화된 인터넷 커뮤니티 사업자에게 개인정보를 제공하고 활동을 보장받는 회원 제도가 존재한다. 따라서, 인터넷 커뮤니티 사용자를 회원과 비회원 두 부류로 나눌 수 있는데, 개방된 공간에서 나를 드러내고 싶지 않아 하는 사용자는 대체로 비회원에 속하게 된다. 개인정보를 제공하고 활동을 보장받는 회원과 달리, 비회원은 드러나는 정보가 거의 없어서 사업자는 이들을 관리하기 위해 IP주소의 일부를 남기는 방법을 사용하고 있다. 비회원 대부분은 국내 IP주소를 남기고 있지만, 종종 완전한 익명성의 사용자는 탈출 노드를 거쳐온 해외 IP주소가 남게 된다. 이렇게 회원과

비회원으로, 그리고 비회원일 경우에는 IP주소의 차이로 완전한 익명성이 나타나는 사용자와 그렇지 않은 사용자를 구별할 수 있다. 그렇다면, 이 완전한 익명성이 나타나는 사용자는 회원과 비회원 중 국내 IP주소 사용자와 어떻게 다를까? 이 질문에 답하기 위해서는 먼저 익명성을 기준에 따라 분류해야 하고 그 결과 나타나는 사용자 사이의 차이를 단어에 속성을 부여하는 방식을 통해 그 비율을 비교해 보아야 한다.

이러한 연구를 진행하기 위해서는 인터넷 커뮤니티가 다음과 같은 조건을 만족해야 한다. (i) 사용자가 익명을 사용해 활동하고 있는 커뮤니티, (ii) 사용자들의 익명이 분류되어 있는 커뮤니티, (iii) 사용자의 이용률이 높아서 대표성을 펼 수 있는 커뮤니티, (iv) 사용자의 게시글에 공개적으로 접근이 가능한 커뮤니티. 본 연구에서 국내 인터넷 커뮤니티를 조사한 결과, 디시인사이드가 가장 적합한 연구 대상이라고 판단한다. 디시인사이드는 위의 조건을 모두 만족시키는 동시에 해당 커뮤니티가 우리 사회에서 독창적인 인식을 가지고 있으므로, 개별 사용자가 익명성을 사용하는 방식이 우리가 가지는 커뮤니티에 대한 일반적인 인식과 차이점을 보이는지 여부를 확인하는 것도 흥미로운 관찰이 될 것으로 생각된다.

II. 선행 연구

해외에서 진행된 인터넷 커뮤니티와 관련된 개념적 연구는 다양한 방면으로 이루어져 있었는데, 그 중에서 특정 커뮤니티를 대상으로 연구를 진행한 일본의 2ch^[2]와 미국의 4chan^[3], reddit^[13]과 관련된 사례가 있었다. 해당 연구들의 목적은, 인터넷 커뮤니티가 일반적으로 긍정적인 시선으로 보여지지 않지만 그럼에도 불구하고 왜 이들이 항상 여러 가지 사회적 이슈의 중심에서 있는지를 알아보는 데 있었다. 국내에서의 특정 커뮤니티에 관한 연구는 커뮤니티를 이용하여 사회적인 현상을 설명하는데 목적을 두는 경우가 많았다^[14,15]. 특히 과거 악명을 떨쳤던 “일간 베스트 저장소”는 당시에 인터넷 우익으로서 혐안과 반일의 최전선에 있었고^[16,17], 최근 떠오르고 있는 “인터넷 커뮤니티가 여성혐오를 만들어 간다.”라는 담론에도 등장하고 있다^[18].

게시글을 작성하는 사용자는 인터넷 커뮤니티를 이루고 있는 최소 단위이며, 이들이 모여서 갤러리라는 담론을 만들어 간다. 따라서, 본 연구에서는 갤러리를 통해서 하나의 현상을 설명하고자 하는 것보다는 사용자 개인에 집중하는 것을 목표로 둔다. 공동체가 유리되어 가는 현대 사회에서는 더더욱 공동체라는 틀보다 그

공동체를 이루고 있는 개인에 더 관심을 가져야 하기 때문이다.

본 연구의 목표가 선행 연구들과 다르며, 그 연구 방법에도 차이를 가진다. 예를 들어, 기존의 연구들에서는, 사용자들이 사용하는 ID를 분류함으로써 대중들의 익명성에 대한 선호도를 분석하거나^[3], 커뮤니티내의 신고게시판에 게시된 글을 분석함으로써 커뮤니티의 악용사례를 분석하는^[14] 방법을 적용하였다. 그러나, 본 연구에서는 커뮤니티에서의 익명성 및 은닉성이 사용자 개인의 언어 사용에 어떻게 미치는지를 확인하는 것을 목표로 하며, 이를 위해서 커뮤니티에 게시된 문장을 평가하는 방법을 적용한다. 문장을 평가하는 방법은 크게 3가지 방법으로 분류될 수 있다: (i) 유전적 정보를 생성하여 정량적으로 평가하는 방법^[25], (ii) 품사 정보와 템플릿을 이용하여 문장을 축소하는 방법^[26], (iii) 형태소 분석을 통해 추출된 단어에 회귀분석을 수행하는 방법^[27]. 본 연구에서는 그 중에서 젊은 세대가 많이 사용하는 신조어를 적절히 추출하기 위해 세 번째 방법을 적용한다.

즉, 본 연구에서는 디시인사이드의 사용자들에 초점을 맞춰, 사용자의 익명성에서 나타나는 특징에 따라 익명성을 분류하고, 그들이 사용하는 언어에는 속성을 부여하여, 이들의 관계성이 어떻게 나타나는지 알아보고, 그 중에서도 완전한 익명성으로 나타나는 사용자들이 디시인사이드에서는 어떠한 모습으로 나타나는지 알아본다.

III. 디시인사이드

디시인사이드는 1999년 7월에 개설된 디지털 카메라 동호인 웹사이트에서 출발한 인터넷 커뮤니티로서 2010년대를 거쳐 크게 발전해 2024년 현재는 대한민국 1위 인터넷 커뮤니티의 자리를 차지하고 있다^[4]. 이들은 2010년대 논란이 점화되었던 극우 성향의 커뮤니티 사이트 “일간 베스트”의 모태이기도 하며, 매년 여러 가지 사건 사고에 빠지지 않았다. 최근에는 디시인사이드 우울증 갤러리에서 촉발된 여고생 투신 사건^[5]이 주목받았다. 신림역에서 발생한 칼부림 사건^[6] 이후 홍기난동 예고 게시글^[7]이 다발적으로 올라와 경찰에서 수사에 착수하는 등, 지금까지도 디시인사이드 뿐만 아니라 인터넷 커뮤니티 이용자에 대한 부정적인 인식이 줄어들지 않고 있다. 그러나 이러한 사회적으로 문제가 있는 행동은 극히 일부에 불과하며 이들이 디시인사이드 전체를 대표할 수는 없다. 왜냐하면 디시인사이드는 2010년대 일본의 인터넷 문화였던 “니코동(ニコニコ動

畫)”에서의 영향과, 한때 우리나라에서 유행했던 UCC(User Created Contents)에서 영감을 받아 “합성 필수 갤러리” 등지에서 해외 인터넷 밈(Meme)을 소스로 활용해 여러 가지 국내 인터넷 밈을 생산해 냈고, 여기에 “병맛 동영상”이라는 이름을 붙여 널리 퍼질 수 있게 만들었던 인터넷 문화의 중심지이기 때문이다. 이러한 인터넷 밈의 생산과 파생되는 재생산을 통해 인터넷 문화를 주도하였고, 특히 COVID-19의 영향으로 인터넷 커뮤니티와 관련 없던 일반인들도 접근할 수 있는 여건이 만들어지면서 더욱 활동이 활발해지게 되었다. 따라서 이제는 인터넷 커뮤니티인 디시인사이드가 갖는 부정적인 영향에만 초점을 두고 담론을 만들어가는 것이 아닌 인터넷 문화의 중심지로서 각각의 소그룹들이 갖는 영향력에 대한 연구가 필요한 시점이다.

3.1 디시인사이드의 갤러리 분류

디시인사이드는 주제에 관련된 소그룹을 개설하여 소통하는 방식을 채택하고 있으며, 이를 갤러리라고 부른다. 갤러리는 사이트에서 자체적으로 관리하는 메인 갤러리와 사용자들이 자체적으로 운영할 수 있는 마이너 갤러리와 미니 갤러리로 구분된다. 메인 갤러리는 사이트 관리자가 개설하고 사용자는 자유롭게 글을 작성할 수 있지만, 이용 약관에 위배되는 글을 작성하거나 신고가 누적되면 사이트 관리자가 해당 게시글을 삭제 처리한다. 마이너 갤러리는 사이트 관리자가 아닌 사용자가 사전 개설 신청을 하여 허가를 받아 개설된 갤러리로서, 메인 갤러리와 비슷한 방향성을 띄고 있으나 사실이기 때문에 사용자가 관리자의 역할을 맡을 수 있다. 마이너 갤러리는 그 이용자가 많아질 경우 메인 갤러리로 승격될 수 있다는 특징을 갖고 있다. 미니 갤러리는 2020년 12월에 도입된 시스템으로 마이너 갤러리와는 달리 사용자가 많아도 정규 갤러리로 승격되지 않는다. 이는 회원만 이용할 수 있는 비공개 개설과 가입을 통해 활동할 수 있는 멤버제도 등 폐쇄성을 제공하여 기존 디시인사이드의 메인 및 마이너 갤러리와는 다른 방향성을 갖고 있다.

3.2 디시인사이드의 익명성 분류

디시인사이드에서는 자체적으로 익명성의 분류를 회원(Membership)과 비회원(Non membership)으로 나누고 있다. 그리고 사용자들은 이 방식을 좀 더 세분화하여 고정닉, 비고정닉, 유동닉으로 구분하고 있다. “고정닉 (Permanent nickname user)”은 회원으로 활동하면서 닉네임을 고정해두는 사용자를 의미한다. “고닉”이라고 줄여서 부르며 이렇게 고정해 둔 닉네임은

Representation of users	Description
	Permanent nickname user: A user is the member of the Internet community, and uses a fixed nickname inside it. The gray icon next to the nick name represents her membership status, and the yellow square represents her using a permanent nickname.
	Temporary nickname user: A user is the member of the Internet community, and uses a nickname that may change frequently. The temporary nickname often consists of meaningless characters.
	Non-member: A user is NOT a member of the Internet community so that the nickname is not with the gray icon. Instead, it is followed by the first two parts of the IP address that the user is using.

그림 1. 디시인사이드에서의 사용자 분류

Fig. 1. Three types of users and their representations in DCINSIDE

다른 사용자가 사용할 수 없기 때문에 자신을 특정할 수 있는 단서를 제공하기도 한다. 그림 1에서 보이는 바와 같이, 이들의 닉네임 뒤에는 회색 아이콘에 노란색 표시가 존재한다. “비고정닉 (Temporary nickname user)”은 회원으로 활동하지만 닉네임이 고정되어 있지 않은 사용자를 의미한다. 내가 원하는 닉네임이 있지만 누군가가 고정닉으로 사용하는 경우, 혹은 비회원처럼 활동하고 싶지만 회원에게만 제공되는 서비스를 이용하고 싶은 사용자들이 비고정닉으로 활동한다. 이들의 닉네임 뒤에는 회색 아이콘이 붙지만 노란색 표시가 붙지 않는다. “유동닉 (Non-member)”은 비회원으로 활동하는 사용자들로 닉네임을 유동적으로 바뀌가며 사용하는 특징이 있다. 대체로 간편하게 칠 수 있는 “○○”이라는 닉을 채택하고 있어서 “○○”이 유동닉의 대명사로 불리고 있다. 이들은 회원이 아니기 때문에 닉네임 뒤에 회색 아이콘이 붙지 않고, 대신에 사용자가 접속하는 IP주소의 앞 두 단위가 표시된다.

3.3 키즈나 아이 마이너 갤러리

본 연구에서는 마이너 갤러리에 초점을 둔다. 메인 갤러리는 사이트 자체에서 관리하고 있기 때문에, 주제와 맞지 않는 게시글이 갤러리에 올라와도 약관에 위반되거나 신고를 통해 삭제되지 않는다면 그대로 남아있기 때문에, 표본 수집 시 주제와 상관없는 게시글이 포함될 가능성이 높아서 제외되었다. 미니 갤러리는 역사가 짧아 분석 대상의 글이 적고, 회원제와 같이 특정한 조건을 맞춰야 하는 어려움이 존재하기에 제외되었다. 따라서 주제와 관련된 게시글 여부를 어느 정도 관리할 수 있고, 역사가 오래되어 충분한 표본의 개수를 제공할 수 있는 마이너 갤러리를 선택한다. 마이너 갤러리에는 다양한 소그룹이 존재하며 그 성격이 각각 다르다. 특히, 관리자가 아닌 사용자가 관리하는 특성상 메인 갤러리보다도 그 수가 압도적으로 많이 존재한다. 본 연구에

서는 취미 카테고리를 대표할 수 있을 정도의 규모를 지닌 마이너 갤러리를 탐색하기 위해, 전체 마이너 갤러리 순위 중 20위 이내로 “대홍갤”이라 불리는 조건을 넣었고 이를 만족하는 동시에 최근 주목받기 시작한 장르 조건을 추가하였다.

버츄얼 유튜버(Virtual Youtuber, Vtuber)는 2016년 12월 “키즈나 아이”에 의해 정의된 용어^[9]로 유튜버를 통해 창작활동을 하는 유튜버(Youtuber)와 가상공간을 뜻하는 버츄얼(Virtual)을 합쳐 가상 캐릭터가 유튜버로 활동하는 것을 의미한다^[9]. 구글 트렌드 분석에 따라 버츄얼 유튜버에 대한 관심도를 확인해 보면, 세계적으로나^[10] 국내적으로도^[11] COVID-19 유행 발생 이후로 관심도가 급증하였으며 이러한 관심은 지금까지도 계속해서 이어져 오고 있다. 일본 버츄얼 유튜버 업계의 쌍두마차인 “홀로라이브 프로덕션”과 “니지산지”와 같이 기업에서 운영하는 형태를 참조하여, 우리나라에서도 2021년 “이세계아이돌”과 “V-LUP” 그리고 2023년 “스텔라이브”와 같은 기업형 버츄얼 유튜버들이 대거 등장하였고 현재까지도 인기를 이어가고 있다. 따라서 세계 최초의 AI 버츄얼 유튜버 “키즈나 아이”를 포함하여 버츄얼 유튜버 전반에 대하여 다루는 갤러리인 “키즈나 아이 마이너 갤러리”를 대상으로 연구를 진행한다^[12].

IV. 익명성 및 단어의 속성

4.1 익명성

선행연구에서는^[1] 익명성을 정체성 정보 유형에 따라 개념화하였고, 이를 기반으로 인식 유형에 따라 사회적 익명성, 시각적 익명성, 명명적 익명성, 소재적 익명성으로 분류하였다. 본 연구에서는 해당 분류 방법과 디자인사이드에서 시행하고 있는 익명성의 분류를 통합하여, 그림 2와 같이 특징에 따라 익명성을 정의한다.

첫 번째는, 회원으로서 특정할 수 있고 자신을 드러내고 있는 익명성이다. 이들은 회원으로 등록되어 있기

에 외부 기관에 의해 개인정보가 드러날 가능성이 존재하고, 닉네임을 통해서 자신을 드러내고 있기 때문에 누구인지 특정할 수 있다. 여기에는 기존 고정 닉네임 사용자와 고정 닉네임을 선점하지 못해 비고정 닉네임을 사용하는 사용자가 해당한다. 이러한 익명성을 “고닉 (Whitebox member)”으로 표현한다. 기본적으로 고닉은 고정 닉네임을 사용하는 사용자들의 줄임말이지만 본 논문에서는 비고정닉 중에서도 고닉과 같이 자신의 정체성을 드러내고자 하는 사용자들까지 포함한다. Whitebox member의 경우 실명제와 달리 추적이 가능하지만 회원으로서 실명인증이 되어있지 않다는 차이점이 있다. 또한 실명제 회원은 글 게시와 동시에 본인의 정보가 노출되지만 Whitebox member는 글 작성자의 신원을 스스로 드러내지 않는다는 것에서 차이가 있다.

두 번째는, 회원으로서 특정하기 어렵고 자신의 정보를 드러내지 않은 익명성이다. 이들은 회원으로 등록되어 있기에 마찬가지로 외부 기관에 의해 개인정보가 드러날 가능성이 존재하지만, “ㅇㅇ”과 같은 비고정 닉네임을 사용하고 있기 때문에 자신을 드러내고 있지는 않아 누구라고 특정하기 어려운 사용자가 해당한다. 이러한 익명성은 커뮤니티 내에서 회원이지만 속은 텅 비어 있다는 의미로 사용하는 깡통과 같은 계정을 뜻하는 “깡통계 (Blackbox member)”로 표현한다.

세 번째는, 개인 정보를 특정하기 어렵지만 자신의 정보를 드러내고 있는 익명성이다. 이들은 비회원이기 때문에 게시물 목록과 같이 회원이 사용할 수 있는 서비스를 이용하지 못해 누구인지 특정할 수는 없지만, 국내에서 제공되는 인터넷 서비스를 이용하고 있기 때문에 외부 기관에 의해서 충분히 개인정보가 드러날 수 있고 닉네임 뒤에 IP주소의 앞 두 단위가 나타나고 있다는 특징이 있다. 이러한 익명성은 커뮤니티 내에서 통용되고 있는 “유동 (Traceable non-member)”으로 표현한다.

마지막으로, 비회원으로서 특정하기 어렵고 자신의 정보도 드러내지 않은 익명성으로 완전한 익명성으로 나타나는 사용자다. 이들은 대체로 VPN이나 TOR 브라우저를 사용하기 때문에 외부 기관에 의해 개인정보가 드러날 염려가 없다. 이들을 유추할 수 있는 정보는 오직 닉네임 뒤 IP주소의 앞 두 단위뿐이다. 여기에는 해외에서 서비스되는 IP주소를 사용하는 유동 사용자가 해당한다. 이러한 익명성은 커뮤니티 내에서 통용되고 있는 “토르 (Tor user)”로 표현한다.

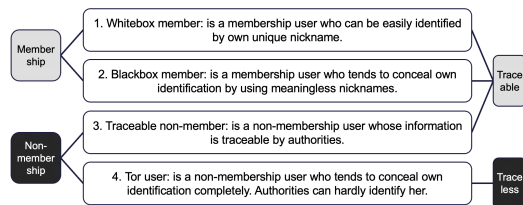


그림 2. 익명성 및 은닉성 특성 기반으로 분류된 사용자
Fig. 2. Categorization of 4 types of users with respect to anonymity and concealability in this study.

4.2 단어의 속성

인터넷 커뮤니티에서 사용되는 단어의 속성을 그림 3과 같이 분류한다. 우선, “안녕”, “오늘”과 같이 일반적으로 사용하는 단어는 “일반 단어 (Everyday word)” 속성으로 분류하고, “텐핑”, “대황숙”, “그홀아”와 같이 커뮤니티 내부에서만 사용하는 단어는 “커뮤니티 단어 (Community word)” 속성으로 분류한다. 커뮤니티에서의 언어 사용으로 고소 또는 고발당하는 사례를 참조하여 추가적으로 분류한다. 법적 처벌 관련 세부 분류에 모욕이나 명예훼손으로 상달하는 사례^[20]와 최근 이슈에 올랐던 통신 매체 이용 음란죄로 상달하는 사례^[21]에 참고하여 욕설과 비속어 그리고 각종 비하 단어에 해당하는 “욕설 및 비속어 (Insulting word)” 속성과 대체로 성(性)과 관련된 단어인 “천박한 단어 (Vulgarism)” 속성으로 분류한다.

Categorization of words	Description and examples
Everyday word	"hello", "today", "안녕", "오늘", "어뻐어"
Community word	"텐핑", "대황숙", "그홀아"
Vulgarism	Words related to sex in an indecent or offensive way. "성(性)"과 관련된 단어
Insulting word	Slang, offensive words, racism, etc.

그림 3. 게시글에 사용된 단어의 속성 분류
Fig. 3. Categorization of 4 attributes for words defined in this study.

V. 제안하는 방법론

5.1 가설 설정

익명성의 특징과 단어 사용 비율 사이의 관계에 대해 알아보기 위해서 다음과 같은 가설을 설정한다.

(1) 고닉 ID, 강통계 ID, 유동 ID는 모두 외부에 의해 추적 가능한 익명성으로 대체로 비슷한 단어 사용 비율을 나타낼 것이다.

(2) 토르 ID는 추적이 어려운 완전한 익명성을 나타내고 있으며, 만약 앞선 3가지 특징의 익명성과 다른 양상을 보인다면, 익명성의 특징이 다른 것은 완전한 익명성이 온전히 악의적이고 공격적인 행위를 위해 작동되고 있으며, 단지 개방된 공간에서 나를 숨기기 위한 특성은 옳다고 볼 수 있다. 만약, 같은 양상을 보인다면 익명성의 특징이 달라도 사용자의 단어 사용 행태는 비슷하다고 볼 수 있기 때문에, 악의를 갖고서만 완전한 익명성을 사용하는 것이 아닌 개방된 공간에서 나를 숨기기 위한 특성이 여전히 존재한다고 볼 수 있다.

(3) 게시글 전체에 대한 형태소 분석과 게시글 한 줄에 대한 분석을 통해, 익명성의 특징에 따라 전체적으로 어떤 속성의 단어를 사용하고 있는지와 그 단어 속성에 따라 정해지는 게시글의 속성은 어떠한지가 나타날 것이다. “천박한 단어”나 “비속어”의 경우, 한 번이라도 사용된다면 문장 전체를 오염시킬 수 있는 특성을 갖고 있다. 따라서, 전체적인 단어 속성에서는 크게 눈에 띄는 비율을 갖지 않는다고 하더라도, 게시글의 속성에는 영향을 미쳐 한 줄에 대한 분석에서는 그 비율이 어느 정도 높게 나타날 것이다.

5.2 평가 방법

디시인사이드의 단어 사용을 평가하기 위한 절차는 다음과 같다. (i) 웹 크롤링을 통해 한 페이지당 게시글 수집하고 이를 병합하여 하나의 텍스트 파일로 생성한다. (ii) 중복된 게시글을 제거하여 완성된 게시글 텍스트 파일을 생성한다. (iii) 게시글 텍스트 내용을 익명성에 따라 분류한다. (iv) 전체적인 형태소와 한 줄에 대한 형태소를 분석하여 단어 속성을 부여한 후 결과값을 산출한다.

웹 크롤링은 방문한 웹사이트의 요소에 접근하여 원하는 정보를 가져온다^[22]. 게시글의 제목은 작성자의 의도를 포함하기 때문에, 본 연구에서는 디시인사이드 게시글의 제목에 초점을 둔다. 이를 위해, 셀레니움 라이브러리를 사용하고 있으며 CSS selector를 통해 웹 페이지의 문서 객체 모델 (DOM)의 요소에 접근하여 원하는 부분을 파이썬 리스트에 추가하고 텍스트 파일로 내보내는 방법을 사용한다. 파이썬 리스트는 익명성의 분류에 따라서 각각의 텍스트 파일로 존재하는데, 이를 호출하여 KoNLPy라는 자연어 처리 시스템을 통해 어떤 형태소를 갖고 있는지 파악한다. KoNLPy는 한국어 정보처리를 위한 파이썬 패키지로 파이썬 프로그래밍 언어로 한국어 정보처리를 할 수 있게 한다^[23]. 본 연구에서는 가장 직관적으로 형태소를 분류를 하는 Okt Class를 사용한다. 사용자 사진은 분석 패키지가 사용할 수 있도록 전처리한 단어의 모음집으로, 신조어와

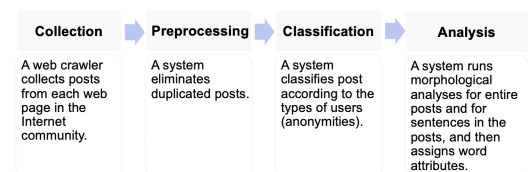


그림 4. 사용자의 단어 사용을 평가하는 방법에 관한 단계
Fig. 4. Procedure of the proposed method evaluating the use of words.

같이 쉽게 변하는 단어나 고유명사를 분석기가 찾을 수 있게 하는 역할을 한다. Okt Class를 위한 사용자 사진을 생성하여 형태소 분석에 이용한다.

형태소 분석에서는 문장의 의미 전달에 있어서 가장 중요한 체언과 용언에 초점을 둔다. 특히, 명사, 형용사, 동사를 중심으로 분류하고, 이외 형태소는 고려하지 않는다. 다만, 현재의 형태소 분석기는 자연어 처리 후 나오는 결괏값에 있어 완전하지 않다는 단점이 있기 때문에, 명사, 형용사, 동사와 관련된 사용자 사진을 만들어 문장 처리와 결괏값에 대한 보정을 할 필요가 있다. 본 연구에서는 보정된 분석기를 통해 두 가지 방법으로 “익명성 특징-단어 사용” 비율을 산출한다. (1) “사용된 형태소 전체에 대한 속성” 비율은, 게시글에서 사용된 모든 형태소를 분류하고, 분류된 형태소가 단어 속성을 명시한 리스트에 존재하는 여부를 확인하여 있다면 해당 속성을 부여한다. (2) “게시글 한 줄이 갖는 속성” 비율은, 게시글 리스트에서 한 줄을 불러와서 형태소를 분류하고, 분류된 형태소가 단어 속성을 명시한 리스트에서 천박한 언어나 욕설 및 비속어를 포함할 경우, 그와 관련된 특성비를 통해 맞는 속성을 부여한다. 포함하지 않는 경우에는, 커뮤니티 단어 속성을 확인하여 커뮤니티 단어 속성과 일반적인 단어 속성으로 분류한다. 두 가지 방법을 수행하는 이유는 전체적인 단어 사용이 주는 인상과 비속어와 같이 부정적인 단어가 들어갔을 때 주는 인상이 다르기 때문이다.

본 연구에서는 2017년 2월 6일 처음 게시된 3번 게시글부터 작업 당시 마지막인 2023년 3월 1일에 게시된 7,506,774번 게시글까지 남아있는 총 6,060,728개의 게시글 중에서 고닉 (Whitebox member) 3,166,088개 (52.239%), 강통계 (Blackbox member) 1,128,499개 (18.620%), 유동 (Traceable non-member) 1,757,890개 (29.005%), 토르 (Tor user) 8,251개 (0.136%)로 나눌 수 있었고 이를 분석한 결과는 다음과 같다.

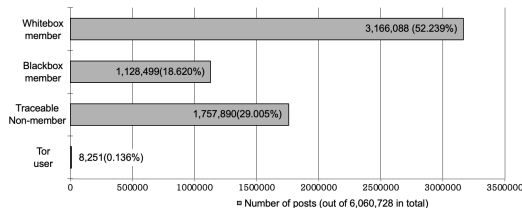


그림 5. 분류된 사용자에 의해 작성된 게시글 개수 및 비율
Fig. 5. Numbers of posts written by 4 types of users (anonymities) and their percentages.

VI. 실험 결과

6.1 전체 게시글에 대한 형태소 분석 결과

전체 게시글에 대한 형태소 분석 결과는 다음과 같으며, 그림 6과 그림 7에서 보여준다. “고닉”의 전체 3,166,088개의 글 중에서는 19,269,739개의 형태소를 찾아냈으며, 이 중 일반 단어 속성은 9,344,411개 (48.492%), 커뮤니티 단어 속성은 2,653,654개 (13.771%), 천박한 단어 속성은 224,476개 (1.165%), 욕설 및 비속어 속성은 322,694개 (1.675%)이다. 4개의 속성에 해당하지 않는 기타 형태소는 6,724,504개 (34.896%)로 나타났다. 이 중에서 의미 전달에 중요하지 않은 기타 형태소를 전체 표본에서 제외하게 되면, 총 12,545,235개의 형태소가 되고, 다시 계산된 비율은 일반 단어 속성이 74.486%, 커뮤니티 단어 속성이 21.153%, 천박한 단어 속성이 1.789%, 욕설 및 비속어 속성이 2.572%로 나타났다.

“강통계”의 전체 1,128,499개의 글 중에서 7,163,763개의 형태소를 찾아냈으며, 이 중 일반 단어 속성은 3,372,686개 (47.080%), 커뮤니티 단어 속성은 1,057,009개 (14.755%), 천박한 단어 속성은 99,577개

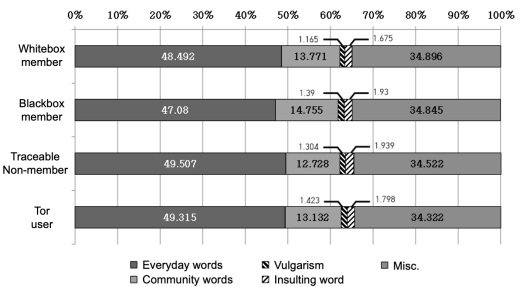


그림 6. 전체 게시글에 대한 형태소 분석 결과 (기타 형태소 분류 포함)
Fig. 6. Result of morpheme analysis for entire posts (including misc. morpheme).

Fig. 6. Result of morpheme analysis for entire posts (including misc. morpheme).

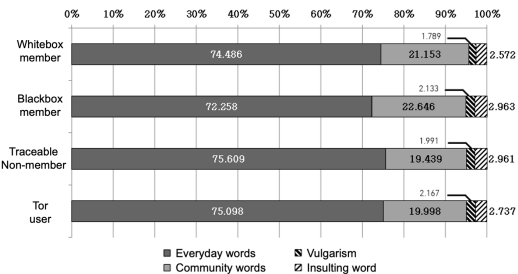


그림 7. 전체 게시글에 대한 형태소 분석 결과 (기타 형태소 분류 제외)
Fig. 7. Result of morpheme analysis for entire posts (excluding misc. morpheme).

Fig. 7. Result of morpheme analysis for entire posts (excluding misc. morpheme).

(1.390%), 욕설 및 비속어 속성은 138,292개 (1.930%)이다. 여기에 해당하지 않는 기타 형태소는 2,496,199개 (34.845%)로 나타났다. 이 중에서 의미 전달에 중요하지 않은 기타 형태소를 전체 표본에서 제외하게 되면 총 4,667,564개의 형태소가 되고, 다시 계산된 비율은 일반 단어 속성이 72.258%, 커뮤니티 단어 속성이 22.646%, 천박한 단어 속성이 2.133%, 욕설 및 비속어 속성이 2.963%로 나타났다.

“유동”의 전체 1,757,890개의 글 중에서 12,081,472개의 형태소를 찾아냈으며 이 중 일반 단어 속성은 5,981,229개 (49.507%), 커뮤니티 단어 속성은 1,537,745개 (12.728%), 천박한 단어 속성은 157,494개 (1.304%), 욕설 및 비속어 속성은 234,266개 (1.939%)이다. 여기에 해당하지 않는 기타 형태소는 4,170,738개 (34.522%)로 나타났다. 이 중에서 의미 전달에 중요하지 않은 기타 형태소를 전체 표본에서 제외하게 되면 총 7,910,734개의 형태소가 되어 다시 계산된 비율은 일반 단어 속성이 75.609%, 커뮤니티 단어 속성이 19.439%, 천박한 단어 속성이 1.991%, 욕설 및 비속어 속성이 2.961%로 변화되어 나타났다.

“토트”의 전체 8,251개의 글 중에서 53,625개의 형태소를 찾아냈으며 이 중 일반 단어 속성은 26,445개 (49.315%), 커뮤니티 단어 속성은 7,042개 (13.132%), 천박한 단어 속성은 763개 (1.423%), 욕설 및 비속어 속성은 964개 (1.798%)이다. 여기에 해당하지 않는 기타 형태소는 18,411개 (34.332%)로 나타났다. 이 중에서 의미 전달에 중요하지 않은 기타 형태소를 전체 표본에서 제외하게 되면 총 35,214개의 형태소가 되어 다시 계산된 비율은 일반 단어 속성이 75.098%, 커뮤니티 단어 속성이 19.998%, 천박한 단어 속성이 2.167%, 욕설 및 비속어 속성은 2.737%로 변화되어 나타났다.

6.2 게시물 한 줄에 대한 형태소 분석 결과

게시글 한 줄에 대한 형태소 분석 결과는 그림 8에서 보여준다. “고닉”의 전체 3,166,088개의 게시물 중에서 일반 단어 속성의 게시글은 850,707개 (26.869%), 커뮤니티 단어 속성의 게시글은 2,082,793개 (65.785%), 천박한 단어 속성의 게시글은 82,976개 (2.121%), 욕설 및 비속어 속성의 게시글은 149,612개 (4.725%)로 나타났다.

“강통계”의 전체 1,128,499개의 게시물 중에서 일반 단어 속성의 게시글은 241,765개 (21.424%), 커뮤니티 단어 속성의 게시글은 801,396개 (71.014%), 천박한 단어 속성의 게시글은 31,543개 (2.795%), 욕설 및 비속어 속성의 게시글은 53,795개 (4.767%)로 나타났다.

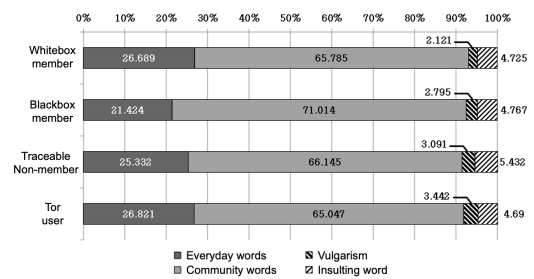


그림 8. 게시물 한 줄에 대한 형태소 분석 결과
Fig. 8. Morphological analysis for each line of post.

“유동”의 전체 1,757,890개의 게시물 중에서 일반 단어 속성의 게시글은 445,327개 (25.332%), 커뮤니티 단어 속성의 게시글은 1,162,749개 (66.145%), 천박한 단어 속성의 게시글은 54,330개 (3.091%), 욕설 및 비속어 속성의 게시글은 95,484개 (5.432%)로 나타났다.

“토트”의 전체 8,251개의 게시물 중에서 일반 단어 속성의 게시글은 2,213개 (26.821%), 커뮤니티 단어 속성의 게시글은 5,367개 (65.047%), 천박한 단어 속성의 게시글은 284개 (3.442%), 욕설 및 비속어 속성의 게시글은 387개 (4.690%)로 나타났다.

6.3 결과 분석

전체 게시글을 대상의 형태소 모음에 대한 단어 속성 비율을 자세히 살펴보면, 전체적으로 일반 단어 속성이 차지하는 비율이 가장 많았고 다음으로 기타 형태소가 많았다. 커뮤니티 단어 속성은 전체적으로 13~14%의 비율을 차지하고 있었고, 천박한 단어 속성과 욕설 및 비속어 속성은 5% 미만으로서 많이 나타나지 않았다. 의미 전달에 있어 필수적이지 않은 기타 형태소 비율을 빼고 계산된 단어 속성 비율에서도 비슷한 결과가 도출되었다. 이 경우, 일반 단어 속성이 70% 이상을 차지하고 있고 커뮤니티 단어 속성은 20%대로 상승하였으나 큰 상승 폭을 보이지는 않았다. 이러한 결과는, 사회의 인식과는 달리, 인터넷 커뮤니티 사용자들이 전체적으로 커뮤니티와 관련된 언어나 성희롱 및 수치스러운 발언, 욕설 및 비속어를 자주 사용하지 않는다는 것을 보여준다. 오히려, 일상생활에서 널리 사용되는 단어를 많은 비율로 사용하고 있음을 보여준다.

그러나, 일반 단어 속성을 오염시킬 수 있는 다른 속성들의 특성을 고려하여 게시물 한 줄에 대한 분석 결과를 살펴보면, 전체적으로 커뮤니티 단어 속성이 차지하는 비율이 가장 많은 것을 확인할 수 있다. 그 다음으로 일반 단어 속성이 많이 나타난다. 이러한 결과를 통해, 일반인들과는 달리 커뮤니티 사용자임을 알 수

있게 한다. 천박한 단어 속성과 욕설 및 비속어 속성은 최대 5%대까지 차지하는 비율이 상승하였고, 이를 통해서 부정적인 단어는 단일 상태에 있을 때보다 문장에 섞여 들어가 있을 때 더 눈에 띄게 되는 사실을 확인할 수 있다. 특히, 커뮤니티 단어 속성과 일반 단어 속성의 관계는 단일 상태와 문장에 섞여 있을 때가 서로 역전되고 있음을 보여주고 있으며, 게시글의 속성으로 봤을 때 인터넷 커뮤니티 사용자는 커뮤니티 단어 속성을 보여주는 게시글을 일반적인 언어 속성만 드러나는 게시글보다 더 많이 작성하고 있음을 보여준다.

마지막으로, 전체적인 실험 결과로 봤을 때, 4가지 익명성 사이에서 큰 차이가 없이 대체로 비슷한 비율을 보여주고 있다. 이는 4가지 익명성 사용자들이 같은 틀 안에서 단어를 사용하고 있음을 보여준다.

6.4 가설 증명

본 논문에서는 실험 결과를 통해 앞서 제시했던 가설에 대한 증명을 할 수 있다. 첫 번째로, “고닉”, “강통계”, “유동”은 공통적으로 외부에 의해 추적이 가능한 익명성이며, 이러한 특성으로 인해 대체로 비슷한 단어를 사용하고 있음을 그 비율의 유사함을 통해 확인할 수 있다.

두 번째로, “토르”는 추적이 어려운 완전한 익명성이지만 앞선 3가지 특징의 익명성과 같은 양상을 보이고 있기 때문에, 익명성의 특징이 달라도 사용자의 단어 사용 행태는 비슷하다고 볼 수 있다. 이를 통해 악의를 갖고서만 완전한 익명성을 사용하는 것이 아니며²⁴⁾, 개방된 공간에서 나를 숨기기 위한 특성인 은닉성이 여전히 존재하고, 결국 완전한 익명성 상태에서 은닉성만을 나타내는 존재를 증명할 수 있다.

마지막으로, 게시글 전체에 대한 형태소 분석과 게시글 한 줄에 대한 분석을 통해, 익명성의 정도에 따라 전체적으로 어떤 속성의 단어를 사용하고 있는지와 그 단어 속성에 따라 정해지는 게시글의 속성을 확인할 수 있었다. 천박한 단어 속성과 욕설 및 비속어 속성의 경우, 전체적인 부분에서는 큰 비율을 차지하고 있지는 않았으나, 이들이 문장에 들어가서 게시글의 속성을 오염시키기 때문에 게시글 한 줄에 대해서는 어느 정도 비율을 차지하기 시작하는 것을 확인할 수 있었다.

VII. 결 론

본 연구에서는 인터넷의 익명성 특징에 따라 커뮤니티 사용자의 단어 사용 행태가 어떤 차이를 보이는지 알아보았다. 실험 결과를 통해, 은닉성만을 나타내고자

하는 사용자들이 존재한다는 것과 이들의 단어 사용이 다른 익명성의 특징과 별다른 차이를 보이지 않는다는 점을 보였다. 본 연구는 다음과 같은 의의를 가진다. 우선, 기존의 디시인사이드와 같은 인터넷 커뮤니티에 대한 연구는 이들의 부정적, 폐쇄적인 특성에 따른 영향이나 정치색에 대한 연구가 주된 주제였다. 그러나, 본 연구는 인터넷 커뮤니티 분석에 대한 관점을 바꿔서, 게시글 사용례가 익명성의 특징과 관련되어 있는지를 알아보았다. 또한 디시인사이드 갤러리의 익명성을 분류하고, 이를 형태소 분석을 통해 단어에 속성을 부여하는 방식을 갖는 방법론을 제안하고 적용하였다. 보다 직관적으로 결과를 보임으로써 인터넷 커뮤니티에 대한 내용적인 판단이 보다 쉽게 내려질 수 있도록 하였다. 본 연구에서는 취미에 관련된 갤러리를 통해 분석하였지만, 이후 더 확장하게 된다면 타 갤러리에도 적용할 수 있으며, 궁극적으로는 이들 갤러리에 대한 분석을 확장하여 디시인사이드라는 거대한 담론에 다가갈 수 있게 될 것으로 기대한다.

References

- [1] B. H. Spitzberg, “Preliminary development of a model and measure of computer-mediated communication (CMC) competence,” *J. Computer-Mediated Commun.*, vol. 11, no. 2, pp. 629-666, Aug. 2006.
(<https://doi.org/10.1111/j.1083-6101.2006.00030.x>)
- [2] M. Matsumura, A. Miura, Y. Shibauchi, Y. Osawa, and M. Ishizuka, “The dynamism of 2 channel,” *J. Inf. Process. Soc. Japan*, vol. 45, no. 3, pp. 1053-1061, Mar. 2004.
(<https://cir.nii.ac.jp/crid/1050282812860484736>)
- [3] M. Bernstein, A. Monroy-Hernández, D. Harry, P. André, K. Panovich, and G. Vargas, “4chan and/b: An analysis of anonymity and ephemerality in a large online community,” in *Proc. Int. AAAI Conf. Web and Social Media*, vol. 5, no. 1, pp. 50-57, Barcelona, Spain, Aug. 2011.
(<https://doi.org/10.1609/icwsm.v5i1.14134>)
- [4] Similarweb, *Website Traffic - dcinside.com Category Rank #1*(in Korea, Republic of), Retrieved Jul. 8, 2024, from <https://www.simil>

- arweb.com/website/dcinside.com.
- [5] Channel A, *Sexual crimes committed in 'Depression Gallery', a live broadcast of extreme choice for teenagers*(2023), Retrieved Jul. 8, 2024, from <https://www.youtube.com/watch?v=e1fHId0E6-c>.
- [6] JTBC News, *Horrific stabbing in broad daylight near Sillim Station - 1 dead, 3 injured*(2023), Retrieved Jul. 8, 2024, from <https://www.youtube.com/watch?v=5VNgNC-C DpU>
- [7] S. Kim, *DC Inside's "Report of a stabbing incident" is attracting attention - 'Murder notices' mostly DC*(2023), Retrieved Jul. 8, 2024, from <https://www.segye.com/view/20230809503075>.
- [8] A.I.Channel, *[Self-introduction] Nice to meet you! I am Kizuna AI*(<https://www.youtube.com/watch?v=NasyGUeNMTs>)(2016), Retrieved Jul. 8, 2024, from <https://www.youtube.com/watch?v=NasyGUeNMTs>
- [9] C.-Y. Kim and H.-W. Han, "A study on subjectivity of virtual youTuber," *Humanities Contents*, vol. 58, pp. 187-209, Sep. 2020. (<https://doi.org/10.18658/humancon.2020.09.187>)
- [10] Google Trends, *'Virtual Youtuber' / Worldwide standard in last 5 years*, Retrieved Jul. 8, 2024, from https://trends.google.co.kr/trends/explore?date=today%205-y&q=%2Fg%2F11f5d65ct_&hl=ko
- [11] Google Trends, *'Virtual Youtuber' / Korean standard in last 5 years*, Retrieved Jul. 8, 2024, from https://trends.google.co.kr/trends/explore?date=today%205-y&geo=KR&q=%2Fg%2F11f5d65ct_&hl=ko
- [12] *Introduction Kizuna AI minor Gallery*, Retrieved Jul. 8, 2024, from <https://gall.dcinside.com/mgallery/board/lists?id=kizunaai>
- [13] K. E. Anderson, "Ask me anything: What is Reddit?," *Library Hi Tech News*, vol. 32, no. 5, pp. 8-11, Jul. 2015. (<https://doi.org/10.1108/LHTN-03-2015-0018>)
- [14] E. G. Park, "A study on the analysis of the characteristics of harmful information in the cyber community," M.S. Thesis, Hanyang University, 2012. (<https://repository.hanyang.ac.kr/handle/20.500.11754/137302>)
- [15] M.-W. Noh, S.-N. Lee, H.-S. Yi, and J.-Y. Choi, "Meaning construction of boasting in the cyberspace interaction: A case study on a Korean internet message board for boasting," *J. Cybercommun. Academic Soc.*, vol. 29, no. 3, pp. 5-48, Sep. 2012. (<https://www.dbpia.co.kr/journal/articleDetail?nodeId=NODE01975200>)
- [16] M. Song, "Korea-Japan net nationalism site and the Anti-Korean sentiment, anti-Japanese consciousness," *J. Image and Cultural Contents*, vol. 10, pp. 127-148, Jun. 2016. (<https://doi.org/10.24174/jicc.2016.06.10.127>)
- [17] S. Kim, "Understanding the hate speech phenomenon in Japan and South Korea from '2 Channel' and 'Ilbe'," Ph.D. Dissertation, University of Tsukuba, Jul. 2018. (<https://doi.org/10.15068/00153824>)
- [18] J. Eom, "Strategic misogyny and its contradictions: Focusing on the analysis of posts from the internet community ilgan best jeojangso(Daily Best Storage)," *Media, Gender & Culture*, vol. 31, no. 2, pp. 193-236, Jun. 2016. (<https://www.dbpia.co.kr/journal/articleDetail?nodeId=NODE06702526>)
- [19] J. Choi, "Study on the different anonymity perceptions impacting on posting malicious message according to types of online communities," Ph.D. dissertation, Hanyang University, 2016. (<https://www.dbpia.co.kr/journal/detail?nodeId=T13960579>)
- [20] Lawtalk Search, *Consultation Result : "Cyber defamation/insult"*, Retrieved Jul. 8, 2024, from <https://www.lawtalk.co.kr/?keyword=디시인사이드>
- [21] Lawtalk Search, *Consultation Result : "Korean law for vulgarism"*, Retrieved Jul. 8, 2024, from <https://www.lawtalk.co.kr/?keyword=디시인사이드>
- [22] C. Castillo, "Effective web crawling," *ACM*

SIGIR Forum, vol. 39, no. 1, pp. 55-56, Jun. 2005.

(<https://doi.org/10.1145/1067268.1067287>)

- [23] E. Park and S. Cho, "KoNLPy: Korean natural language processing in Python," in *Proc. Annual Conf. Human and Cognitive Language Technol.*, pp. 133-136, Chuncheon, Korea, Oct. 2014.

(<https://scienceon.kisti.re.kr/srch/selectPORSrchArticle.do?cn=NPAP12013363>)

- [24] C. Tran, K. Champion, A. Forte, B. M. Hill, and R. Greenstadt, "Are anonymity-seekers just like everybody else? An analysis of contributions to Wikipedia from Tor," in *Proc. IEEE Symp. Sec. and Privacy*, pp. 186-202, San Francisco, USA, May 2020.

(<https://doi.org/10.1109/SP40000.2020.00053>)

- [25] S. H. Yang and Y.-S. Kim, "A quantitative evaluation method of the quality of natural language sentences based on genetic algorithm," *J. KISS(B): Software and Appl.*, vol. 26, no. 11, pp. 1372-1381, 1999.

(<https://www.dbpia.co.kr/journal/articleDetail?nodeId=NODE00602960>)

- [26] S.-S. Lee, K.-W. Yeom, J.-H. Park, and S.-B. Cho, "A sentence reduction method using part-of-speech information and templates," *J. KISS*, vol. 35, no. 5, pp. 313-324, 2008.

(<https://www.dbpia.co.kr/Journal/articleDetail?nodeId=NODE00992696>)

- [27] H. Kim, "Extraction method of new word from online community: The application of new method to morphological analysis," 2019. (<https://www.dbpia.co.kr/journal/detail?nodeId=T15011210>)

황 병 화 (Byeonghwa Hwang)



2017년 3월~현재: 인천대학교
정보통신공학과 학사과정
<관심분야> 무선통신, 인터넷
커뮤니티, 데이터공학
[ORCID:0009-0004-0658-5916]

이 은 규 (Eun-Kyu Lee)



2000년 2월: 동국대학교 전자
공학과 공학사
2002년 2월: KAIST 전자 공
학과 석사
2014년 6월: UCLA 컴퓨터학
과 박사
2015년 2월~현재: 인천대학교
정보통신공학과 교수

<관심분야> 컴퓨터보안, 사물인터넷, 인공지능, 커
넥티드 자율주행차

[ORCID:0000-0003-1459-5458]