

아이는 마을이 키우고 시는 조직이 키운다

전시형 (서울과학기술대학교)

목 차

- 1. 서 론
- 2. 인공지능 도입의 비기술적 난제
- 3. 인공지능이 조직에 안착하기 위한 대안
- 4. 결 론

1. 서 론

2016년 3월 구글 딥마인드 사의 알파고 (AlphaGo)가 프로 바둑기사 이세돌 9단과의 대국을 4승 1패로 승리하면서 학계와 재계와 정계에 충격을 안겼다. 인간의 창의성은 절대 우월하다는 믿음이 과도했다는 반성과 함께 인공지능을 혁신의 도구로 도입하겠다는 시도가 이어졌다[1]. 의료, 보안, 방법, 로봇 등 다양한 업종에서 화려한 성공사례를 홍보했으며 실제로 성과가 있었으나 상당수 사람들은 일자리 축소, 사생활 침해, 인간의 자율성 무시와 같은 위협에 두려움을 느끼는 형국이다. 미국의 소프트웨어 회사 PEGA는 세계 6개국 총 6,000명의 소비자들을 대상으로 설문조사를 실시했다. 응답자의 70% 이상이 인공지능에 대한 두려움을 느낀다고 답했고, 심지어 그 중 1/4은 기계가 세계를 장악하는 상황을 우려하기까지 했다[2]. 인공지능 기술에 힘입어 다수가 중요한 정보에 쉽고 빠르게 접근한다는 점에는 긍정적인 수 있어도 인공지능이 가져올 수도 있는 해악에 대한 두려움이 크다는 현실을 나타낸다.

이렇게 기대와 걱정이 공존하는 가운데, 인공지능 기술을 통해 조직에 혁신을 가져와서 눈에 띄는 성과를 내려는 시도는 충분히 성숙하지 못한 기술 수준과 고질적인 전문인력 부족만이 아니라 기술 외적인 장벽을 극복하지 못해 실패하곤 했다. 조직의 규모를 막론하고, 구성원이 도저히 초연하게 받아들이지 못할 ‘일자리 감소’처럼 인공지능 기술이 가져올 변화 자체가 부정적으로 여겨지는 상황이 태반이었기 때문이다[2]. 전문가의 영역으로 여겨졌던 지식과 경험이 인공지능 도입으로 인해 유익하지 못한 방향으로 재평가를 받을 수 있다는 가능성은 앞서 언급한 일자리 감소로 이어지기에 위협으로 느낄 수 밖에 없다. 때문에 인공지능 기술 도입에 직접적으로 반대하지는 못하더라도 여러 사안마다 개인정보 및 영업기밀 보호와 같이 일견 타당한 명분을 들어 인공지능 기술 적용을 늦추거나 이미 적용한 기술에 대한 무력화를 시도하는 사례가 나오기도 한다.

보신주의, 무사안일주의, 부서 간 알력으로 나타나는 인공지능 배척 기조는 조직 내 생존, 단기적으로는 성과평가에 기인한다. 단일 조직이 도입

하려는 작은 업무영역의 일부가 되는 인공지능 기술은 무난히 안착하는 편이지만, 범조직적으로 영향을 미치는 인공지능은 구설수에 오르기 마련이다. 인공지능 도입에 따른 성과를 어떻게 평가하고 배분할 것인지에 영향받는 조직들은 민감하게 반응하며 관련한 조직이 협업해야만 방해가 없을 것이다. 결국 인사가 만사라는 금언은 인공지능에도 적용이 된다.

인공지능은 도구적인 개념이면서도 당분간은 조직이 경쟁력을 강화하기 위한 핵심역량에 속하기 때문에, 아이는 마을이 키운다는 아프리카 속담처럼 인공지능도 조직이 키워 나가야 조직의 역량으로 자리매김할 수 있다. 특히 정보기술 업계에 속하지 않은 조직이 인공지능 기술을 적용한 초기에는 가시적인 성과를 내기가 힘들다. 이 단계에서는 안일하게 효과성을 폄하하지 않고 관계 부서가 협업하여 보완사항을 도출함으로써 조직 차원으로 인공지능 기술을 활용하는 습관을 들여야 한다. 기술 부서만이 아니라 평가를 담당하는 인사 부서와 기획 부서 등 지원 부서까지 동원해야 가능하다. 때문에 인공지능 기술을 성공적으로 적용하는 조직은 단순히 기술에 친화적이거나 전문인력을 충분히 구인했기보다는 경영을 잘하는 조직이다.

2. 인공지능 도입의 비기술적 난제

알파고 열풍 이래로 인공지능 기술을 도입하기 위해 애써온 기업이 비기술적 사안으로 고심했던 사례를 들어보고자 한다.

콜센터 사업을 하는 A사는 상담품질 관리 시스템을 운영한다. 관계법령에 따라 상담대화를 녹취하며 품질 담당자가 무작위로 상담대화를 청취하여 상담 품질 지표를 기입해 왔다. 음성인식(Speech-to-Text, STT) 기술을 이용하여 녹취한

음성을 텍스트로 변환한 후, 이를 자연어처리(Natural Language Processing, NLP) 기술을 이용하여 비속어 사용 여부, 스크립트 준수 여부, 폭언 사용 여부 등을 모니터링함으로써 상담품질을 관리하며 상담사를 보호하는 인공지능 프로젝트를 성공적으로 수행했다. 구축한 기능에 있어서는 이전과는 달리 상담 대화 전수 검사가 이루어져 고객사의 호평을 받았다. 다만 상담대화 카테고리 분류에 있어서는 학습 데이터의 품질이 좋지 못해 기대했던 수준으로 인공지능 모델을 개발하지 못했다. 상담대화 카테고리는 전적으로 상담사가 업무 프로그램의 메뉴 버튼을 클릭하는 방식으로 생성할 수밖에 없다. 때문에 상담사의 숙련도와 성실성 뿐만 아니라 개성에 따라 카테고리를 기입하게 되므로 기준이 일관적이지 않았다. 이러한 데이터로 학습한 인공지능 모델은 상담사마다 기준이 갈리는 몇몇 카테고리를 더욱 분류하지 못했다.

이 문제를 해결하려면 상담대화 카테고리 별로 모범적인 사례만 추려서 해당하는 상담을 수행했던 상담사를 투입하여 인공지능 모델이 학습하기 충분한 양으로 데이터를 생성하게 해야 한다. 카테고리 간 데이터 불균형은 필연적으로 발생하기 마련이라, 상담 건수가 불충분한 카테고리는 해당 상담사 혹은 기준을 인계 받은 숙련자가 카테고리 지정 작업을 해야 한다. 인건비 절감을 이유로 비숙련자를 투입하면, 품질 불량 사태를 재현할 뿐이다. 더불어 상담대화 카테고리 자체를 재설계하는 방안을 고려할 필요가 있다. 상담수 콜센터의 상담대화 카테고리는 점진적으로 개선해 온 산물이라 MECE (Mutually Exclusive Collectively Exhaustive, 상호배제와 전체포괄)적이지 못한 편이다. 별 수 없이 상담사가 '기타' 카테고리에 지정해야 하거나, 특정 카테고리가 실은 동일한 값의 다른 카테고리에 포함이 되는 관계일 때도 있다. 비숙련자가 카테고리 지정 작업을 하기 힘든 이유이기도 하며, 이 같은 상황은 인공지능의

카테고리 분류 학습에도 악영향을 끼친다. 상담대화 카테고리 재설계는 인공지능 분류 성능을 획기적으로 개선할 방안이지만, 고객사를 설득하며 대부분 협력업체 소속인 상담사가 받아들여지게 하는 변화관리가 불가피하다. 이해관계자가 모여 서비스 품질 향상을 기치로 공감대를 형성하면서 당장의 매출과는 무관한 의사결정을 해야 할 텐데, 갑을 관계를 가리지 않고 실무자는 일상업무를 수행하기에도 빠듯한 실정이므로 시도할 엄두조차 내지 못했다. 자체 콜센터를 운영하는 기업은 비교적 추진이 용이하지만 콜센터 위탁사업을 하는 기업은 의사결정을 이끌어야 할 고객사가 다수이므로 아무래도 엄두를 내지 못할 방안이다.

유통사업을 하는 B사는 매장에서 품질이 발생하면 매출에 타격을 입으므로 Vision AI를 도입하여 매대에서 진열한 상품이 사라지면 해당 매대 담당자에게 알람이 가는 시스템을 구축하고자 했다. 상품마다 진열 방식과 품질을 판단하는 기준이 달라 인공지능 모델 개발이 쉽지는 않았지만, 파일럿 프로젝트였으므로 일단은 rule과 휴리스틱도 동원하여 일부 상품의 품질을 인지하는 인공지능 모델을 배포했다. 매대에 카메라를 설치하고 1주일 정도 모니터링을 하던 프로젝트팀은 특이한 상황에 맞닥뜨렸다. 이 파일럿 프로젝트를 추진하게 된 근거 중에는 어떤 상품이 일주일 혹은 특정한 단위기간에 몇 회 이상 품질이 발생했다는 현황조사가 있었는데, 모니터링 기간 중에는 품질이 거의 발생하지 않았다. 프로젝트팀은 인공지능 모델의 결함을 의심하여 서버에 저장한 매대 사진을 전수 조사했고, 실제로 품질이 발생하지 않았음을 확인했다. 해당 매대의 상품이 팔리지 않은 상황은 아니었다. 단지 파일럿 프로젝트 기간에는 이전과는 달리 매대 담당자가 설 틈 없이 매대에 상품을 채웠기에 품질이 발생하지 않았던 것이다. 매대 담당자 입장에서는 파일럿 프로젝트 때문에 업무가 더욱 과중해진 셈이었다.

이렇게 된 데에는 부서 간의 입장 차이가 있었다고 본다. 파일럿 프로젝트를 추진한 부서는 본사 조직이고, 매대를 관리하는 부서는 지점 조직이다. 처음에는 매대에서 발생한 품질을 조기에 알리면 매출 손실을 막고자 하던 담당 부서가 환영할 거라 예상했다. 분명히 그러한 효익은 담당 부서가 바라는 바이기는 하지만, 파일럿 프로젝트 전후의 실무자 인터뷰를 바탕으로, 품질 상황을 본사 조직이 안다는 점이 부담으로 작용했다고 추측한다. 그렇다고 품질 상황을 본사 조직이 모도록 구축하기는 불가능했다. 디지털 전환(Digital Transformation, DX)을 추진하는 부서가 프로젝트 진척을 모르는 건 어불성설이었다. 더구나 품질 추이를 데이터로 확인할 수 있으면 매장 운영에 바쁜 지점 조직 대신 본사 조직에서 해당 상품의 진열 확대나 (이미 본사 조직이 매대 진열 관리인 Plan-O-Gram, POG를 관리한다.) 여타의 개선방안을 제시하는 식의 DX 성공사례를 이어나가는 것도 기대할 만했기 때문이었다. 파일럿 프로젝트 자체는 이슈 없이 끝났지만, 실무 부서의 호평도 두드러지지 않은 편이었다. 아쉽게도 본사업 추진에 탄력을 받지 못했다.

인공지능이 조력자로 (AI assistant) 인식이 되어야 부담 없이 적극적으로 도입하며 또 다른 분야로 적용 범위를 넓혀갈 수 있는데, 감시자로만 여겨지는 바람에 실패는 아니지만 성공도 아니게 된 사례였다. 품질이라는 병목이 부정적으로만 인식이 되고 발견이 되었을 때에 압박을 받을 수 밖에 없는 상황이었다. 품질 발생이 가시적으로 드러난다면 평가를 받는 입장에서 어떻게 받아들일지에 대해 이해관계자들이 공론화하는 게 바람직했다. 이를 통해 걸림돌을 사전에 제거했다면 인공지능은 조력자로서 환영 받았을지도 모른다. 다행히 품질 감지 효과 자체는 인정 받았으므로 언젠가 B사에서 의사결정만 하면 매대의 품질만이 아니라 물류 창고의 안전재고 수량 파악이나 재고

실사 등 현장에서 감당할 만한 점진적인 속도로 적용 범위를 넓힐 수 있다.

업계에서 선두를 달리는 제조업 C사는 유통매장의 매대 전면에 자사 제품을 진열한 비율을 정기적으로 조사한다. 1차로는 매장에 방문하는 영업대표가 육안으로 헤아려서 보고하고, 2차로는 리서치업체에 위탁하여 무작위로 선별한 매대의 진열 비율을 보고 받는다. 자체적으로 단기직원을 고용하여 조사하기도 한다. 인공지능을 활용할 만한 분야를 전방위로 조사하던 C사는 이 업무를 Vision AI를 통해 자동화하고자 했다.

영업대표는 모바일 앱을 통해 자사 제품의 진열 비중을 보고하는데, 이 모바일 앱을 통해 매대를 촬영한 이미지에서 자사 제품과 경쟁사 제품의 전면 진열 비율을 산출하겠다는 아이디어였다. 이렇게 하려면 Vision AI는 매대에서 상품을 검출하고 자사와 경쟁사 제품을 분류해야 했다. 해당 업종의 제품을 검출하는 인공지능 모델을 개발하기도 쉽지 않았지만 포장이 수시로 달라지기까지 하는 상품 class 수백 개를 분류하는 인공지능 모델 개발은 더욱 어려웠다. 더군다나 매대를 한 번만 촬영하여 1개 이미지를 추론할 수 있었다면 그나마 용이했겠지만, 매장이 대체로 거리를 크게 둘 만큼 넓지는 못하므로 일종의 파노라마 기능을 개발하는 등 난관이 많았다. 그렇게 공을 들였음에도 중소 유통매장은 다른 카테고리의 상품을 한 군데에서 진열한다든가 조명 색상이 유별나다든가 하는, 개발자 입장에서는 일일이 대응하기 힘든 현상이어서 인공지능 모델의 분류 성능이 크게 떨어지기도 했다. 인공지능 기술의 당시 한계는 이를 추진한 조직도 잘 알고 있었으나 현장의 영업대표는 크게 반발했다.

매대 전면 진열 비율은 영업대표의 KPI였기 때문에 민감한 사안이었다. 테스트 매장에서 분류 성능이 좋아 시범 사용을 시작했던 것인데도 전국 각지의 다양한 현장에서는 앞서 언급한 대로 유독

결과가 ‘나쁜’ 사례가 속속 발생했다. 추진 부서와 영업부서 간에 어떤 논쟁이 있었는지 정확히는 모르지만 추진 부서로서는 상당히 부담을 느낄 만한 항의가 있었던 것으로 안다.

복기를 하자면, 단계적 추진이라는 로드맵을 이해관계자가 숙지했어야 했다. 개발 기간이 촉박하여 모바일 앱 내에서 구현하는 촬영 UX (User eXperience, 사용자 경험) 개발에도 미진한 부분이 있었고, 서비스 초기에는 발생할 수 밖에 없는 오검출, 오분류를 터치 한두 번에 신고하는 기능 역시 기획은 했으나 개발을 포기할 수 밖에 없었던 시점이었다. 입장이 달라서이겠지만, 간편한 신고 기능이 없더라도 ‘이의 제기’ 같은 버튼을 만들어 관리 부서가 확인하는 업무 절차를 만들면 어떨을까 하는 아쉬움이 남는다. 물론 IT 업계 종사자가 아닌 사용자 입장에서는 시스템이 주는 중압감이 상당할 수 밖에 없다. 본업도 스트레스가 많은데, 모바일 앱이 얼토당토 않은 결과를 보여주는 스트레스까지 참기는 어려웠을 것이다.

상품 분류 성능이 99%인 인공지능 모델을 만들었다면 모두 해결이 되는 문제이긴 했다. 하지만 다른 인공지능 솔루션 업체가 나서서 파일럿 프로젝트를 했으나 결과는 대동소이했다. 복합적인 한계가 분명하게 존재하는 현장들이므로, 이 기능은 장기간 지속적으로 인공지능 모델을 고도화하는 식으로 완성도를 높여야만 영업부서가 불만을 표출하지 않을 수준으로 개발이 가능하다고 본다. 애초에 추진 부서와 영업부서가 대립해서는 꾸준히 진행할 수가 없는 사안이었다.

어떤 메리트를 영업부서에 제공했으면 보기 곱스러운 모바일 앱의 보고서 화면을 차아줄 수 있었을까? 예를 들어, 인센티브 금액을 올린다면 어떨까? 이전에는 매대 전면 진열 비율을 육안으로 세어 모바일 앱에서 수작업으로 보고했기에 신뢰도에 한계가 있었다고 볼 수도 있다. 당장은 아니지만 궁극적으로는 인공지능을 통해 이 수치를

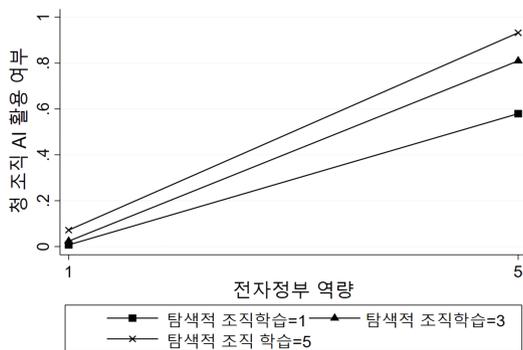
투명하게 처리하게 될 것이므로 인센티브 금액을 올려도 회사는 수용이 가능하지 않을까 한다. 물론 탁상공론의 아이디어이며, 이런 방향으로도 논의가 이루어졌으면 활로가 나오지 않았을까 하는 아쉬움이 남았을 뿐이다.

3. 인공지능이 조직에 정착하기 위한 대안

중앙정부조직을 중심으로 한 정부의 인공지능 도입을 분석한 연구에 따르면, 탐색적 조직 학습에 능숙한 ‘청’ 조직이 전자정부 역량의 증가에 따라 인공지능 활용 확률도 증가하는 양상을 보였다[3].

탐색적 조직 학습 수준은 완전히 새로운 일의 방식을 시도하는 데에 능숙하고, 혁신적인 외부 조직과 전문가에게 개방적이며, 기존과 다른 창조적인 정책대안을 추구할 여유를 갖고 있다고 자부하는 조직에게서 높게 나타난다. ‘청’ 조직의 탐색적 조직 학습 수준이 높은 이유는 집행적 업무를 주로 수행하고 업무는 세분화되어 있으며, 이에 따라 혁신이 발생할 여지가 상대적으로 높다는 논의에 기반한다[3].

다시 말해, ‘청’ 조직은 목표가 보다 가시적이기에 인공지능을 적용해서 얻을 수 있는 성과 역시



(그림 1) 탐색적 조직학습과 전자정부 역량의 상호작용 그래프[3]

가시적으로 도출할 수 있다. 분명히 달성해야 할 목표 아래에서는 적절하게 업무를 분담한 조직 간 협업 역시 용이할 것이므로 앞서 언급한 비기술적 난제를 피할 수 있었던 것으로 본다.

‘청’ 조직과는 달리 ‘부’ 조직은 대부처주의(大部處主義) 기조에 따라 한 부처가 다양하고 유사한 기능들을 복합적으로 편제하므로 복잡성이 증가하고 비정형적 업무를 많이 실시한다. 때문에 조직 민첩성이 높아야 인공지능 기술을 높이는 양상이 보인다. 이 같은 조직은 새로운 이슈에 빠르게 대응하는 구조적 역량을 확립하도록 해야 하며 [3], 기업 역시 기술보다는 경영 관점에서 먼저 조직을 정비할 필요가 크다.

4. 결론

적절한 개인정보 비식별화 기준, 적절한 군집화 기준은 실무 부서가 제일 잘 안다. 사전 인터뷰만 잘 해도 인공지능 모델 학습에 도움이 되는 정보를 많이 얻을 수 있다. 실무자가 이야기한 대로 특성 중요도가 나올 때도 많다. 하지만, 부서 간 상대평가인 상황이라면 담당 임원이 다른 부문 혹은 사업부가 벌이는 일에는 그다지 협조하지 않는다. 기껏 비용과 시간을 들여 도입한 인공지능 기술이 성공적으로 조직에 정착하게 하려면 구성원이 위협으로 느끼게 하지 말아야 한다.

인공지능 기술로 큰 혁신을 이끌어 내려면 단일 부서 내에서만 쓰는 기능 뿐만 아니라 사업부 차원, 전사 차원으로 영향을 미치는 사안에 적용해야 한다. 콜센터 상담 대화 녹취 음성을 텍스트로 변환하기 위해서는 음성인식 (STT) 엔진을 도입하고, 속기사를 고용하여 생성한 데이터로 학습하면 충분한 성과를 얻는다. 그러나 그렇게 확보한 텍스트 데이터를 통해 상담 품질 지표를 산출하려면 협력사를 포함한 이해관계자의 관심과 협력을

이끌어 내야 한다. 협업할 체계를 구축해야 지속이 가능하다.

또한 정량적으로 협업을 이끌 장치가 필요하다. 세부적인 업무 성과가 가시적이려면 측정이 가능해야 한다. 인공지능을 적용한 초기에 퀀텀점프를 하지는 못한다. 작은 성공 경험이 쌓이면서 조직적으로 역량을 갖추어 임계점을 넘어서는 순간 퀀텀점프라 할 만한 대성공을 거두는 게 수순이다. 작은 성공경험을 인지하려면 때로는 퇴보하기까지 하는 소소한 성과를 측정해야 하며, 이러한 개선 효과를 경영진이 격려하고 인정해야 한다. 결국은 부서 평가 문제로 귀결하지만 인사 정책만은 아니다. 경영진과 실무자 간 신뢰가 바탕으로 전형적인 경영의 일환으로서 접근해야 지속가능한 기업으로 성장하는 가운데 인공지능을 조직 곳곳에서 유용하게 쓰게 될 것이다.

저 자 약 령



전 시 형

이메일 : sihyoung.jum@gmail.com

- 2003년 동국대학교 컴퓨터멀티미디어공학 (학사)
- 2022년 서울과학기술대학교 산업정보시스템공학 (석사)
- 2003년-2022년 롯데정보통신 / 시기술팀장
- 2022년-현재 메조미디어 / 데이터테크팀장
- 관심분야: 인공지능, 빅데이터, 예측, 최적화, 추천, 광고, 스마트 팩토리

참 고 문 헌

- [1] 윤영주, “‘이세돌 vs 알파고’ 세기의 대결 이후 5년...AI 어디까지 왔나”, 시타임스, 2021
- [2] 이창섭, 이현정, “인공지능 혁신에 대한 기대와 불안 요인 및 영향 연구”, 한국콘텐츠학회 논문지, 19(9), 37-46, 2019.
- [3] 한명성, “정부의 인공지능 도입에 관한 분석: 중앙정부조직을 중심으로”, 한국콘텐츠학회 논문지, 22(2), 281-293, 2022. 성, 국민대학교 석사학위 논문, 2020.