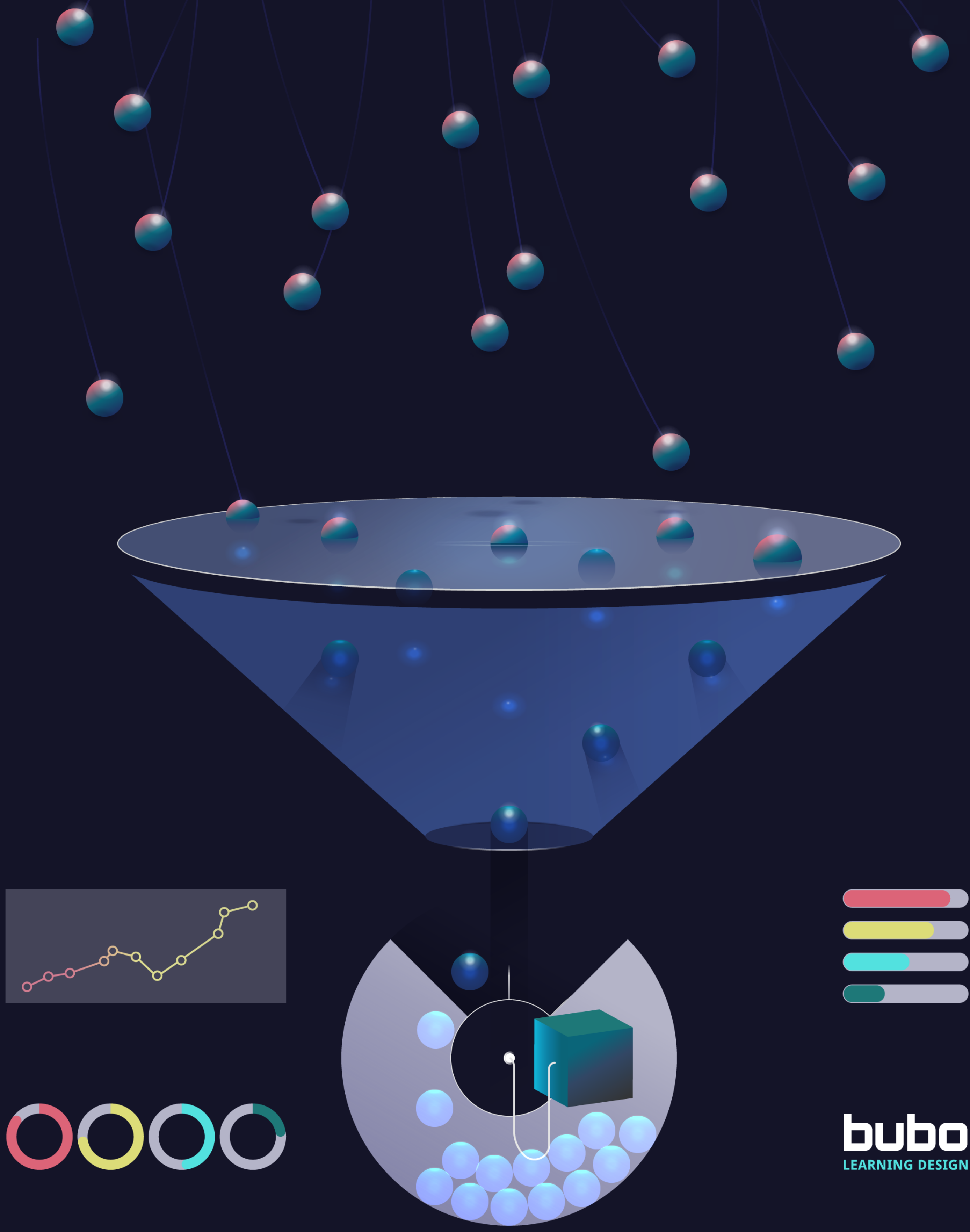
DATA UTILIZATION FOR PREDICTIVE LEARNING ANALYTICS



bubo
LEARNING DESIGN

Introduction

One of the **greatest opportunities** facing industries and tech users today is **properly utilizing** the seemingly perpetual tsunami of data generated every second. A matter that is particularly vexing for those in the field of e-learning, this flood serves to obfuscate **useful metrics** that can **improve** the **performance** of learners.

Bubo Learning Design, LLC, in partnership with **SCAD AI**, has envisioned a way to harness this untapped resource with predictive learning analytics (PLA) enabled by artificial intelligence (AI) and natural language processing (NLP).

The pace at which data is currently being created and consumed is often left up to broad speculation, with estimates of up to 2.5 billion GB of daily user-generated data.[1]

According to the International Telecommunication Union's 2019 report on global digital development,[2] approximately 189 kbit/s of bandwidth is consumed per user in developed nations, with the United States users likely exceeding 200.

To put those numbers into a more digestible context, roughly 1.5 MB are used every minute at a kbit usage rate of 200 per second.

In an average word processor file that excludes graphics, 1.5 MB is more than sufficient to contain an entire novel.

When applying these figures to the estimated 4.1 billion internet users,[3] even at lower rates of bandwidth consumption across various developing nations, it becomes readily apparent that there is more data in existence that is being transferred and created than can be analyzed and employed for productive purposes.

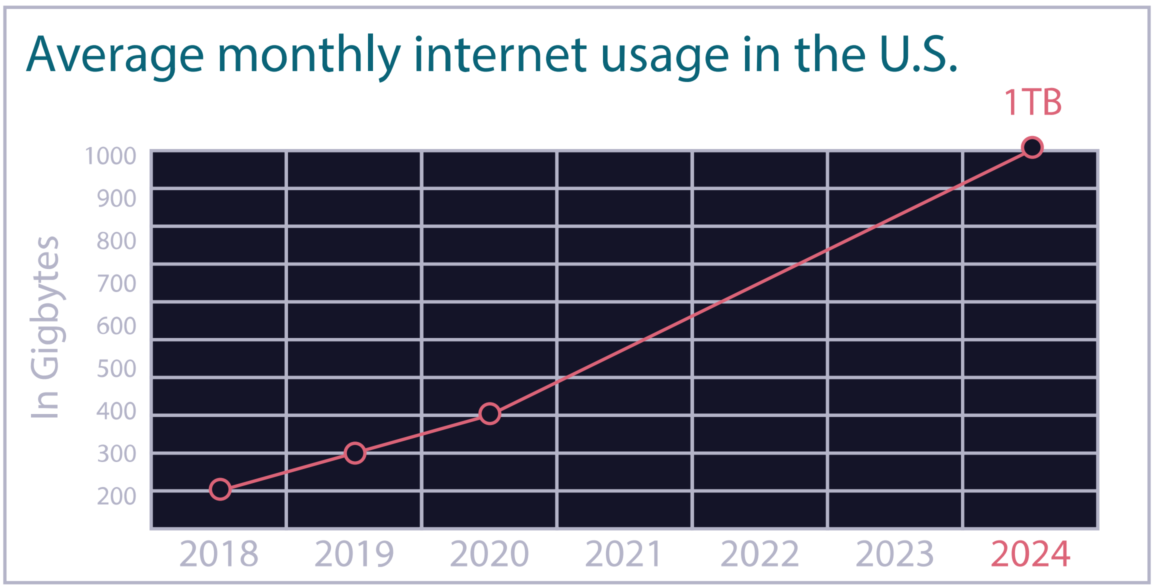
[1] Kietzmann, J., Paschen, J., & Treen, E. R. (2018). Artificial Intelligence in Advertising: How Marketers Can Leverage Artificial Intelligence Along the Consumer Journey. Journal of Advertising Research, 58(3), 263–267. https://doi.org/10.2501/jar-2018-035; Rizkallah, J. (2017, June 5). Council Post: The Big (Unstructured) Data Problem. Forbes. https://www.forbes.com/sites/forbestechcouncil/2017/06/05/the-big-unstructured-data-problem/.

[2] Telecommunication Development Sector, Measuring digital development: Facts and figures 2019 (2019). Geneva, Switzerland; ITU Publications. https://www.itu.int/en/ITU-D/Statistics/Documents/facts/FactsFigures2019.pdf.

[3] Ibid.

There is little chance that this influx of new information will slow with the progression of technology either.

The natural growth of data usage along with the extraordinary circumstances of 2020 that has forced a substantial contingent of workers to telecommute from their homes (a change that will likely be permanent for many after the cessation of the pandemic). This has escalated the average data expenditure to **over 400 GB per household** in the US. It is set to reach over **1 TB** by 2024, even if the rate of use declines with the return to normalcy.[4]



[4] Cohen, J. (2020, June 5). Data Usage Has Increased 47 Percent During COVID-19 Quarantine. PCMag. https://www.pcmag.com/news/data-usage-has-increased-47-percent-during-covid-19-quarantine.; Rizzo, L. (2020, October 25). Americans Working From Home Face Internet Usage Limits. The Wall Street Journal. https://www.wsj.com/articles/americans-working-from-home-face-internet-usage-limits-11603638000.; Toledo, R. (2020, April 17). REPORT: The Average Household's Internet Data Usage Has Jumped 38x in 10 Years. DecisionData. https://decisiondata.org/news/report-the-average-households-internet-data-usage-has-jumped-38x-in-10-years/.; Toledo, R. (2020, September 9). ANALYSIS: The Average Household Will Surpass ISP's 1TB Data Caps Within 3 Years. DecisionData. https://decisiondata.org/internet/analysis-the-average-household-will-surpass-isps-1tb-data-caps-within-3-years/.

The vast majority of underutilized data is called "unstructured data."

Unstructured data tends to be data like emails, videos, images, and audio files that are typically not stored in a way that makes it easy to search for or employ beyond their initial production purpose.[5]

This is in contrast to **"structured data,"** which is typically quantifiable and categorical information (e.g., names, addresses, and transactions) that has been organized in a database for future application.[6]

Predictive learning analytics, which is the collection and analysis of data to make predictions of probable outcomes, in its most efficacious form, will harness both structured and unstructured data to make informed decisions about a given task.

Structured data is needed to identify essential and relevant elements for a task. Since structured data has already been sorted in some fashion and defines critical variables, it is the easiest kind of data for PLA to handle and catalyzes the program's predictions.

One example of PLA using structured data are algorithms that can make product recommendations and filter out undesirable items based on past spending history.

This implementation is among the most straightforward and commonly utilized by companies like Amazon and the digital games retailer Steam to curate and personalize their storefront for customers.

[5] Taylor, C. (2018, March 28). Structured vs. Unstructured Data. Datamation. https://www.datamation.com/big-data/structured-vs-unstructured-data.html.; Pickell, D. (2018, November 16). Structured vs Unstructured Data – What's the Difference? G2. https://learn.g2.com/structured-vs-unstructured-data.

[6] Ibid.

To use unstructured data, the process increases in complexity significantly.

Data like email messages and multimedia files often don't contain quantifiable variables used in their original format, although they might have vital information that can greatly enhance the predictive process.

The question "How are you today?" doesn't contain anything that would tell a mathematical algorithm about the sender or receiver of the message based on methods used for structured data analysis, but still may contribute to enhancing the accuracy of PLA when considered in a broader, interconnected context. This is where natural language processing comes into play.

Natural Language Processing

NLP has the lofty goal of allowing computers through AI to understand natural languages, i.e., languages spoken and written fluently by people that are often filled with vagaries and tend to be grammatically incorrect when used casually, but are nonetheless understandable due to the adaptability of the human mind.

Compared to programming languages like JavaScript and C++, these artificially constructed computer languages allow no room for error if a program is to be successfully executed.[7]

The gulf between humans and computers' abilities to intuitively derive meaning from fluid and chaotic methods of communication has been especially exasperating for e-learning designers. Owing to the semantic deficit of past and current technologies, it is commonplace for open-ended, non-binary questions in assessments to be avoided almost entirely, as they have typically required human administrators to be graded.

[7] Philbin, C. A. (2017, November 22). Natural Language Processing: Crash Course Computer Science #36. YouTube. https://www.youtube.com/watch?v=fOvTtapxa9c.

This avoidance of questions that require critical thinking creates two additional problems:

Firstly, binary questions with answers that can be memorized through rote learning do not allow for a proper demonstration of mastery over a given subject.

The difference between rote learning and true comprehension of the concepts at hand is vast.

While useful for obtaining a basis of knowledge to build from, **rote learning** does **not** explicitly enable a deeper understanding of how to apply that knowledge to different contexts; it may only allow the learner to know the answer to that one specific question.

As an example of the limits of rote learning, in a study of non-Arabic speaking Muslims who memorized the Quran written in Classical Arabic, it was found that even though the subjects could with good accuracy identify and interpret passages from the text, it did not give them a sufficient understanding of Classical Arabic as a language to be fluently applied elsewhere.[8]

Conversely, a meta-analysis of vocabulary instructional methods showed “that the most effective vocabulary teaching methods included **both definitional and contextual information** in their programs.”[9]

Second, in many learning management systems (LMS), it can be fairly simple to “game” the systems that have simplistic question making.[10]

[8] Saleem, A. (2015). Does memorization without comprehension result in language learning?.

[9] Stahl, S. A., & Fairbanks, M. M. (1986). The Effects of Vocabulary Instruction: A Model-Based Meta-Analysis. Review of Educational Research, 56(1), 72–110. https://doi.org/10.3102/00346543056001072.

[10] Baker, R. S. et al. (2004). Off-task behavior in the cognitive tutor classroom. CHI ‘04: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, 383–390. https://doi.org/10.1145/985692.985741.

By abusing hint features or repeatedly taking a test until getting a passing grade after learning the right and wrong answers, a learner can appear to be sufficiently knowledgeable with little effort.[11] In corporate training assessments where prospective or new employees are often given several chances to pass a certification exam, these behaviors are particularly worrisome. They must be accounted for to prevent potential workplace mishaps caused by a lack of training.

The frustration that undertraining creates may lead to the employee's exit from the company, voluntarily or otherwise. Replacing employees who have left their position through termination or resignation is quite a significant cost; an estimated average of **20% of that employee's annual salary** for workers earning less than $75,000.[12]

Proper implementation of NLP in a fine-tuned LMS will alleviate both of these issues. Allowing the creation of questions that require fluency of the subject material to answer that can then be analyzed by the system with minimal need for human input, thereby reducing the reliance on abusable, binary questions for evaluation.

Data Mining

To achieve the level of natural language comprehension required for an LMS and other similarly complex applications, many of the most sophisticated NLP efforts use **"data mining"**; the collection and arrangement of data into meaningful patterns.[13]

By forming a large set of interlinking components from data mined from the internet or internal systems, an AI can learn language through the frequency of word use, morphology, syntax, use in conjunction with other words, and many other components.

[11] Aleven, V., et al. (2004). Toward Tutoring Help Seeking. Lecture Notes in Computer Science, 3220, 227–239. https://doi.org/10.1007/978-3-540-30139-4_22.

[12] Boushey, H., & Glynn, S. J. (2012, November 16). There Are Significant Business Costs to Replacing Employees. Center for American Progress. https://www.americanprogress.org/issues/economy/reports/2012/11/16/44464/there-are-significant-business-costs-to-replacing-employees/.

[13] Castro, F. et al. (2007). Applying Data Mining Techniques to e-Learning Problems. Evolution of Teaching and Learning Paradigms in Intelligent Environment Studies in Computational Intelligence, 183–221. https://doi.org/10.1007/978-3-540-71974-8_8.

Apple's Siri and Amazon's Alexa use data mining to great effect to reinforce their ability to understand and replicate human speech whenever they are used.[14] Siri and Alexa send conversations spoken to the digital assistants back to databases to be analyzed. Recently, both companies now allow users to opt-out of having their data collected.[15]

Despite the ability to opt-out from data collection and repeated explanations from cybersecurity experts that there's **no evidence** to suggest that digital assistants are always recording our conversations,[16] there is a pervasive public belief that they are spying on us,[17] partly due to their capacity to so presciently ascertain our interests and spending habits.

The fact is, however, that they don't need to listen to us to make predictions.

Browser history, cookies, and location data all provide enough information to create a surprisingly detailed profile for predictive algorithms to use for various purposes.[18]

The larger the base of data and the more points connecting said data, the greater the ability for an AI to make accurate inferences.[19] With databases as large as the ones possessed by Amazon and Apple, that only serves to further enrich their PLA and NLP endeavors.

[14] Day, M., Turner, G., & Drozdiak, N. (2019, April 10). Amazon Workers Are Listening to What You Tell Alexa. Bloomberg. https://www.bloomberg.com/news/articles/2019-04-10/is-anyone-listening-to-you-on-alexa-a-global-team-reviews-audio.; McMillan, R. (2013, April 19). Apple Finally Reveals How Long Siri Keeps Your Data. Wired. https://www.wired.com/2013/04/siri-two-years/.

[15] Bohn, D. (2019, August 3). Amazon will let you opt out of human review of Alexa recordings. The Verge. https://www.theverge.com/2019/8/2/20752418/amazon-alexa-human-review-recordings-opt-out-eu.; Leswing, K. (2019, October 28). Apple lets users delete Siri recordings in new iPhone update after apologizing for handling of user data. CNBC. https://www.cnbc.com/2019/10/28/ios-13point2-has-new-siri-privacy-settings-including-deletion-and-opt-out.html.

[16] Smith, E. A. (2019, February 1). No, Your Phone Is Not Eavesdropping on You. Scientific American. https://blogs.scientificamerican.com/observations/no-your-phone-is-not-eavesdropping-on-you/.; Tidy, J. (2019, September 5). Why phones that secretly listen to us are a myth. BBC News. https://www.bbc.com/news/technology-49585682.; BBC. (2019, April 11). Smart speaker recordings reviewed by humans. BBC News. https://www.bbc.com/news/technology-47893082.

[17] Auxier, B., et al. (2019, November 15). Americans and Privacy: Concerned, Confused and Feeling Lack of Control Over Their Personal Information. Pew Research Center: Internet, Science & Tech. https://www.pewresearch.org/internet/2019/11/15/americans-and-privacy-concerned-confused-and-feeling-lack-of-control-over-their-personal-information/.

[18] Auxier, B., et al. (2019, November 15). Americans and Privacy: Concerned, Confused and Feeling Lack of Control Over Their Personal Information. Pew Research Center: Internet, Science & Tech. https://www.pewresearch.org/internet/2019/11/15/americans-and-privacy-concerned-confused-and-feeling-lack-of-control-over-their-personal-information/.

[19] Philbin, C. A. (2017, November 22). Natural Language Processing: Crash Course Computer Science #36. YouTube. https://www.youtube.com/watch?v=fOvTtapxa9c.; Smith, M. A. (2015). Output from Statistical Predictive Models as Input to eLearning Dashboards. Future Internet, 7(2), 170–183. https://doi.org/10.3390/fi7020170.

Indeed, PLA and NLP have made remarkable advances in the kinds of behavior that they can predict, far beyond something as mundane as music preferences.

During a 2016 study where researchers sought to determine the outcomes of court cases tried by the European Court of Human Rights using NLP for textual analysis of past judgments, they found that they could correctly deduce the decisions with an impressive **79% average accuracy**.[20]

For educators, such early insight could benefit struggling, or likely-to-struggle learners when trying to intervene on their behalf to improve their performance.

In their data analysis of at-risk college students taking e-learning courses, Baker et al. concluded that the **first two weeks** of a semester was a sufficient amount of time in which they could forecast the likelihood of course success based on student activity.[21]

Having this degree of foreknowledge would lessen the mental burden for all involved and allow for better time management. There is no imminent threat of failure to drive frantic cram sessions or other last-minute attempts to boost grades.

[20] Aletras, N. et al. 2016. Predicting judicial decisions of the European Court of Human Rights: a Natural Language Processing perspective. PeerJ Computer Science 2:e93 https://doi.org/10.7717/peerj-cs.93.

[21] Baker, R. S. et al. (2015). Analyzing Early At-Risk Factors in Higher Education e-Learning Courses. EDM.

Feedback Loops

A key component of formulating accurate predictions is the currency of the information available. To keep a database growing and relevant, feedback loops, an integral function of AI, are established.

These loops consist of four basic steps: observation of actions and events, analysis against previously stored data, prediction, and a recommendation of action.[22]

The cycle repeats, ideally with minimal oversight necessary from an administrator. This independence is perhaps the greatest asset to companies seeking to implement AI systems.

While the establishment of a database large enough to draw valid conclusions from, and the expense of hardware with enough processing power to handle it may seem daunting at first (though those requirements are becoming **easier** and **cheaper** to implement by the day),[23] the automated nature of **AI processing ultimately requires far fewer resources** than if a company were to hire a new batch of employees to perform the same functions.

Even so, SCAD AI is still able to go one step further towards refining this process's efficiency.

[22] Dhar, A. (2019, July 18). Get More Out Of Feedback Loops With AI. Forbes. https://www.forbes.com/sites/forbestechcouncil/2019/07/18/get-more-out-of-feedback-loops-with-ai/.

[23] Rubens, N., Kaplan, D., & Okamoto, T. (2014). E-Learning 3.0: Anyone, Anywhere, Anytime, and AI. Lecture Notes in Computer Science, 7697, 171–180. https://doi.org/10.1007/978-3-662-43454-3_18.

Our Solution

With SCAD AI's methods, they can take existing systems to cut down on the cost of acquiring a new database and create an AI system that doesn't require the maintenance and labor of a dedicated machine learning operations division.

The AI is then able to generate additional data through feedback loops to quickly bring the system up to speed and allow for a precise and tailored experience for **your specific needs**.

This process has already been successfully implemented with SCAD AI's work on the continuing medical education project Adaptrack, the talent acquisition resource HumanEven, the in-development conversational agent, and application programming interface CatalystBot, along with many other smaller individualized programs.

Bubo Learning Design, together with SCAD AI, is seeking government funding to research applications of AI and NLP to create dynamic learning paths to be paired with content while identifying gaps where new material will be required.

In addition to content development, Bubo Learning Design's efforts will go towards finding learners' critical engagement points to allow for self-reported information such as preferences, recent accomplishments, technical issues, and other data that will help optimize their education in a process called **"people analytics"**; an emerging field of HR assessment where learning is just one part of managing the overall employment experience of personnel onboarding, reassignment, and growth into new roles.[24]

[24] Bersin, J. (2016, July 1). People Analytics Market Growth: Ten Things You Need to Know. JoshBersin.com. https://joshbersin.com/2016/07/people-analytics-market-growth-ten-things-you-need-to-know/.

By establishing key performance indicators from structured and unstructured data mined by intuitive, customized AI with NLP capabilities, we will be not only able to **assist struggling learners** but also identify **ideal career trajectories**. We will suggest promotions to leadership positions for exceptionally suited individuals, and deliver a bespoke curriculum for maximum engagement, all while improving communication between the learners through greater connectivity, enabled by a comprehensive support platform.

With our expertise in the domain of e-learning, Bubo Learning Design tirelessly works to advance the performance and well-being of learners everywhere, whether they're in the corporate, federal, or public spheres.

Now more than ever, we need to be able to bring our communities together safely and encourage their growth, as each member of that flourishing community allows for our learning program to adapt and grow from them, in turn to deliver a truly transformative experience.

About Bubo

Bubo Learning Design, LLC has been providing video, web-based training, virtual simulations, and other cutting-edge modalities for learning to government and corporate clients since 2015. Our reputation for success is due to providing superior, innovative learning and training services throughout the United States and Canada, with clients based in Texas, Oregon, Washington, California, Minnesota, Michigan, and D.C.

To date, our largest clients include Amazon, LinkedIn, Dave & Busters, the US Air Force, Thomson Reuters, and Ally Bank.

Our learning modules are designed to make the learner feel truly engaged. We replicate real-life scenarios that make the learning relevant and practical. These scenarios reflect what their day-to-day job experience is or will be like.

It is an exciting time to be in learning technologies, as conventional methods and tools are giving way to more intriguing, dynamic, and immersive opportunities. It is now possible to provide employees an interactive learning experience that adequately prepares them for their on-the-job roles in more ways than ever before.

For more information on Bubo Learning Design, LLC, visit **https://www.bubold.com**. See the "Our Clients" section to view our broad experience with both commercial and government clients.

About SCAD AI

Niraj Swami, President and CEO of **SCAD.AI** will lead the effort to create data lakes and structuring data to drive performance improvement over time that will be paired with AI and Machine-learning algorithms.

SCAD.AI's ventures primarily focus on the fields of skills development, career trajectory projections, analytics, and human capital.

Our partner, **SCAD.AI**, has also made inroads utilizing AI/Machine Learning to build performance supporting applications.

Example applications that are similar but not entirely equivalent to what we might employ include:

KnowBot: https://vimeo.com/163047019

Pulse: https://youtu.be/Woa8yoA_bg4

HumanEven: https://youtu.be/_vFh76QJmH8

Aletras, N., Tsarapatsanis, D., Preoţiuc-Pietro, D., & Lampos, V. 2016. Predicting judicial decisions of the European Court of Human Rights: a Natural Language Processing perspective. PeerJ Computer Science 2:e93 https://doi.org/10.7717/peerj-cs.93

Aleven, V., McLaren, B. M., Roll, I., & Koedinger, K. (2004). Toward Tutoring Help Seeking. Lecture Notes in Computer Science, 3220, 227–239. https://doi.org/10.1007/978-3-540-30139-4_22

Auxier, B., Rainie, L., Anderson, M., Perrin, A., Kumar, M., & Turner, E. (2019, November 15). Americans and Privacy: Concerned, Confused and Feeling Lack of Control Over Their Personal Information. Pew Research Center: Internet, Science & Tech. https://www.pewresearch.org/internet/2019/11/15/americans-and-privacy-concerned-confused-and-feeling-lack-of-control-over-their-personal-information/.

Baker, R. S., Corbett, A. T., Koedinger, K. R., & Wagner, A. Z. (2004). Off-task behavior in the cognitive tutor classroom. CHI '04: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, 383–390. https://doi.org/10.1145/985692.985741

Baker, R. S., Lindrum, D., Lindrum, M.J., & Perkowski, D. (2015). Analyzing Early At-Risk Factors in Higher Education e-Learning Courses. EDM.

BBC. (2019, April 11). Smart speaker recordings reviewed by humans. BBC News. https://www.bbc.com/news/technology-47893082.

Bersin, J. (2016, July 1). People Analytics Market Growth: Ten Things You Need to Know. JoshBersin.com. https://joshbersin.com/2016/07/people-analytics-market-growth-ten-things-you-need-to-know/.

Bohn, D. (2019, August 3). Amazon will let you opt out of human review of Alexa recordings. The Verge. https://www.theverge.com/2019/8/2/20752418/amazon-alexa-human-review-recordings-opt-out-eu.

Boushey, H., & Glynn, S. J. (2012, November 16). There Are Significant Business Costs to Replacing Employees. Center for American Progress. https://www.americanprogress.org/issues/economy/reports/2012/11/16/44464/there-are-significant-business-costs-to-replacing-employees/.

Castro, F., Vellido, A., Nebot, À., & Mugica, F. (2007). Applying Data Mining Techniques to e-Learning Problems. Evolution of Teaching and Learning Paradigms in Intelligent Environment Studies in Computational Intelligence, 183–221. https://doi.org/10.1007/978-3-540-71974-8_8.

Cohen, J. (2020, June 5). Data Usage Has Increased 47 Percent During COVID-19 Quarantine. PCMag. https://www.pcmag.com/news/data-usage-has-increased-47-percent-during-covid-19-quarantine.

Dhar, A. (2019, July 18). Get More Out Of Feedback Loops With AI. Forbes. https://www.forbes.com/sites/forbestechcouncil/2019/07/18/get-more-out-of-feedback-loops-with-ai/.

Kietzmann, J., Paschen, J., & Treen, E. R. (2018). Artificial Intelligence in Advertising: How Marketers Can Leverage Artificial Intelligence Along the Consumer Journey. Journal of Advertising Research, 58(3), 263–267. https://doi.org/10.2501/jar-2018-035

Leswing, K. (2019, October 28). Apple lets users delete Siri recordings in new iPhone update after apologizing for handling of user data. CNBC. https://www.cnbc.com/2019/10/28/ios-13point2-has-new-siri-privacy-settings-including-deletion-and-opt-out.html.

McMillan, R. (2013, April 19). Apple Finally Reveals How Long Siri Keeps Your Data. Wired. https://www.wired.com/2013/04/siri-two-years/.

Philbin, C. A. (2017, November 22). Natural Language Processing: Crash Course Computer Science #36. YouTube. https://www.youtube.com/watch?v=fOvTtapxa9c.

Pickell, D. (2018, November 16). Structured vs Unstructured Data – What's the Difference? G2. https://learn.g2.com/structured-vs-unstructured-data.

Rizkallah, J. (2017, June 5). Council Post: The Big (Unstructured) Data Problem. Forbes. https://www.forbes.com/sites/forbestechcouncil/2017/06/05/the-big-unstructured-data-problem/.

Rizzo, L. (2020, October 25). Americans Working From Home Face Internet Usage Limits. The Wall Street Journal. https://www.wsj.com/articles/americans-working-from-home-face-internet-usage-limits-11603638000.

Rubens, N., Kaplan, D., & Okamoto, T. (2014). E-Learning 3.0: Anyone, Anywhere, Anytime, and AI. Lecture Notes in Computer Science, 7697, 171–180. https://doi.org/10.1007/978-3-662-43454-3_18

Saleem, A. (2015). Does memorization without comprehension result in language learning?.

Smith, E. A. (2019, February 1). No, Your Phone Is Not Eavesdropping on You. Scientific American. https://blogs.scientificamerican.com/observations/no-your-phone-is-not-eavesdropping-on-you/.

Smith, M. A. (2015). Output from Statistical Predictive Models as Input to eLearning Dashboards. Future Internet, 7(2), 170–183. https://doi.org/10.3390/fi7020170.

Stahl, S. A., & Fairbanks, M. M. (1986). The Effects of Vocabulary Instruction: A Model-Based Meta-Analysis. Review of Educational Research, 56(1), 72–110. https://doi.org/10.3102/00346543056001072

Taylor, C. (2018, March 28). Structured vs. Unstructured Data. Datamation. https://www.datamation.com/big-data/structured-vs-unstructured-data.html.

Telecommunication Development Sector, Measuring digital development: Facts and figures 2019 (2019). Geneva, Switzerland; ITU Publications. https://www.itu.int/en/ITU-D/Statistics/Documents/facts/FactsFigures2019.pdf.

Tidy, J. (2019, September 5). Why phones that secretly listen to us are a myth. BBC News. https://www.bbc.com/news/technology-49585682.

Toledo, R. (2020, April 17). REPORT: The Average Household's Internet Data Usage Has Jumped 38x in 10 Years. DecisionData. https://decisiondata.org/news/report-the-average-households-internet-data-usage-has-jumped-38x-in-10-years/.

Toledo, R. (2020, September 9). ANALYSIS: The Average Household Will Surpass ISP's 1TB Data Caps Within 3 Years. DecisionData. https://decisiondata.org/internet/analysis-the-average-household-will-surpass-isps-1tb-data-caps-within-3-years/.