Press contact:
Michele Moore Duhen
Tel.: +44 370 905 3408
Email: michele.mooreduhen@capgemini.com

Capgemini uses AI to help identify and track sperm whale population

Winning solution for internal competition at Capgemini, sponsored by AWS, uses computer vision capabilities to identify images of individual sperm whales with 97.5% accuracy

Paris – December 3, 2020 – Capgemini has developed an intelligent data solution which is designed to accurately identify sperm whales using computer vision for image processing. The solution, the brainchild of a winning team during the Group's annual Global Data Science Challenge (GDSC), uses Amazon Web Services (AWS) machine learning (ML) services to analyze thousands of photos of sperm whales and identify matches of the individual whales, helping scientists to track migration routes, look at the social structure of the sperm whale groups, and protect the whales' natural habitats. Capgemini's GDSC is a group-wide internal competition where hundreds of employees from all over the world compete in small teams to address real world challenges using artificial intelligence (AI).

Sperm whales can be identified by their tail fins, known as flukes, which show unique markings and structures – similar to human fingerprints – depending on the whale's unique lived experience. Scientists use images of an individual whale's flukes to document key information about its behaviors, migratory paths, and movement patterns.

However, matching new whale pictures with historical ones can be tedious and very labor intensive. Lisa Steiner, a marine biologist and renowned expert for sperm whale research in the Azores, explains the process: *"I have to manually assist the program to pick out the contour for each half of the fluke. If the photos are good, this process doesn't take very long; however, if there isn't a lot of contrast between the fluke and background or there is a lot of glare on the edge of the fluke, I have to follow the contour manually."*

A Capgemini employee experienced the cumbersome approach when she volunteered for an expedition led by Steiner. Seeing a way to help improve Steiner's day-to-day activities, Capgemini proposed replacing Steiner's dated legacy software by harnessing the power of computer vision image processing capabilities to help eliminate the need for manual matching. Thus, the challenge for the 2020 GDSC was born: develop a tool, which uses a machine learning algorithm, to automate the task of sorting and identifying pictures of specific whales to help Steiner in her quest.

Nearly 700 teams from across Capgemini participated in the challenge, and after several months, the winning team – called 'AI Sensing' - presented their solution consisting of a pre-trained deep neural network (ResNet 101) that was fine-tuned with roughly 4,500 pictures containing flukes from more than 2,200 individual whales. The model runs on a GPU-based Amazon Elastic Compute Cloud (Amazon EC2) compute cluster using Amazon SageMaker (AWS's service that helps everyday developers and data scientists build, train, and deploy machine learning models quickly in the cloud and at the edge), which can process the photos in three hours. The ML algorithm is capable of automatically cropping a new picture using AWS Lambda by removing unnecessary parts of the photo and leaving only the fluke in the center, then comparing it to all other pictures in the database and finding matches for a given sperm whale with 97.5% accuracy.

The solution has freed up Steiner to focus her time on tasks like migration tracking and habitat protection rather than spending hundreds of hours manually tagging photos. Steiner is pleased with the results, and recently matched a sperm whale that had not been identified since 2002 due to changes in the contour of its fluke.

"I found quite a cool match that had not been detected previously, due to some changes in the contour," noted Steiner. *"This goes back to 2002, so there were three separate IDs for this animal, instead of one. If I had just clicked match on the first photo, I wouldn't have found it. I'm loving this program!"*

Capgemini continues to work with Steiner to improve the usability of the application as well as to encourage other researchers to try it out. The long-term aim is that people from all walks of life who have spotted and taken a picture of a sperm whale in the Atlantic Ocean can one day upload their pictures into the tool, which would scale up the research capabilities with more data and could lead to the discovery of new whale matches and insights into these beautiful creatures.

Zhiwei Jiang, CEO of the Insights and Data Global Business Line at Capgemini, said, *"The Global Data Science Challenge highlights both Capgemini's commitment to driving positive change through its internal initiative and how using machine learning technology delivers real-world solutions. It's been an honor to see so many Capgemini colleagues from around the globe come together to create an impactful solution that has been instrumental in helping Lisa's work with sperm whales. Congratulations to the 'AI Sensing' team for their efforts this year, and I'm already looking forward to seeing the inspiring outputs from next year's challenge."*

"We are delighted to work with AWS Premier Consulting Partners like Capgemini who use AWS machine learning technology to build innovative solutions that address sustainability," said Mike Miller, Director, AI Devices, Amazon Web Services, Inc. *"This unique solution for Lisa Steiner is a testament to Capgemini's capabilities, and the potential of machine learning to positively impact whale conservation and solve some of the world's toughest challenges. We are optimistic about the potential of machine learning and its ability to materially benefit society."*

Capgemini believes that digital transformation should benefit all of humanity, and its projects like the GDSC will continue to focus on addressing issues relating to sustainability by unleashing human energy through technology for an inclusive and sustainable future.

To see the full story on the Global Data Science Challenge, including video, click here.

About Capgemini
Capgemini is a global leader in consulting, digital transformation, technology, and engineering services. The Group is at the forefront of innovation to address the entire breadth of clients' opportunities in the evolving world of cloud, digital and platforms. Building on its strong 50-year heritage and deep industry-specific expertise, Capgemini enables organizations to realize their business ambitions through an array of services from strategy to operations. A responsible and multicultural company of 265,000 people in nearly 50 countries, Capgemini's purpose is to unleash human energy through technology for an inclusive and sustainable future. With Altran, the Group reported 2019 combined global revenues of €17 billion.

Visit us at www.capgemini.com.