

Dataverse.vn: Thu thập dữ liệu lớn phân tán trên nền blockchain

InfoRe Technology

preprint version - updated 2021.11.21

Tóm lược: Trong nội dung này chúng tôi giới thiệu Dự án phát triển khái quát của Dataverse.vn - một “Mạng thu nhận dữ liệu truyền thông xã hội đa ngôn ngữ bởi sức mạnh cộng đồng” với sự liên kết người tham gia dựa vào những hợp đồng thông minh trên mạng blockchain. Mạng quét dữ liệu này được hình thành từ máy tính của rất nhiều người trên internet, giúp cập nhật thông tin xã hội mới - nóng một cách liên tục và trên diện rộng, phục vụ được cho nhiều quá trình khai thác dữ liệu để tạo ra lợi ích. Dự án cũng đề xuất phát hành đồng token DTVN để đại diện cho giá trị, sức sống và lợi ích của mạng cộng đồng thu thập dữ liệu này và trình bày phương pháp tạo ra một nền kinh tế thu nhỏ xoay quanh đồng DTVN để Dataverse.vn phát triển mạnh, đem lại nhiều giá trị thực tế.

Mục lục

Mục lục	2
1. Giới thiệu	4
2. Dataverse.vn: Social Media Data Miners	5
3. Đổi dữ liệu lấy token	7
4. Xử lý dữ liệu	8
5. Cổng máy Dataverse.vn	8
6. Tokenomics	9
7. Token Gift	11
8. Kế hoạch marketing	12
8.1. Affiliate Marketing	12
8.2. Influencer Marketing	13
8.3. Email Marketing	13
8.4. Các biện pháp marketing khác	13
8.5. Học bổng và từ thiện	14
8.6. Nhóm các nhà đầu tư sớm	14
9. Lộ trình phát hành token	14
10. Kế hoạch kinh doanh SMCC	15
10.1. Tiếp cận thị trường quốc tế	15
10.2. Dịch vụ làm giàu dữ liệu bằng DTVN	15
10.3. Cung cấp công cụ NFT các di sản số của cá nhân	15
11. Giá trị lõi của DTVN	16
12. Định hướng xa	17
13. Hiện thực hóa lợi nhuận - thu hồi vốn sớm	18
14. Take away	19
Giới thiệu	22
Kinh nghiệm	22
Phân tích dữ liệu, tư vấn giải pháp xử lý thông tin - Since 2012	22
Xử lý Dữ liệu lớn - Since 2011	22

Ứng dụng Trí tuệ nhân tạo - Since 2011	22
Nhân lực	23
Đội ngũ phân tích dữ liệu - 63 người	23
Đội ngũ xử lý dữ liệu - 8 người	23
Đội ngũ tư vấn cao cấp - 7 người	23
Giải thưởng	23

1. Giới thiệu

SMCC.VN¹ là hệ thống phần mềm Lắng nghe mạng xã hội được công ty InfoRe Technology² phát triển từ 2011 đến nay. Hoạt động dưới dạng một phần mềm dịch vụ (SaaS), SMCC chuyên rà quét các website tin tức, báo điện tử và các trang công khai trên mạng xã hội Facebook ở Việt Nam để tổng hợp thông tin hằng ngày; các thông tin tổng hợp được xử lý bởi Trí tuệ nhân tạo, hình thành một máy tìm kiếm thống kê tốc độ cao phục vụ nhu cầu truy xuất thông tin của người sử dụng dữ liệu để *bảo vệ thương hiệu, phát hiện khủng hoảng truyền thông, đánh giá xu hướng, tuyển dụng nhân sự, quản trị cộng đồng, chăm sóc khách hàng...* và đặc biệt là *tìm kiếm, tối ưu hóa các tệp khách hàng* cho các chiến dịch marketing của doanh nghiệp.

Năm 2016, SMCC đạt giải Nhất giải thưởng Nhân Tài Đất Việt³ và hiện đã có hơn 25 nghìn người sử dụng, truy cập các thông tin thống kê về hầu hết các lĩnh vực trong xã hội mỗi ngày trên hằng tỉ bản ghi dữ liệu.

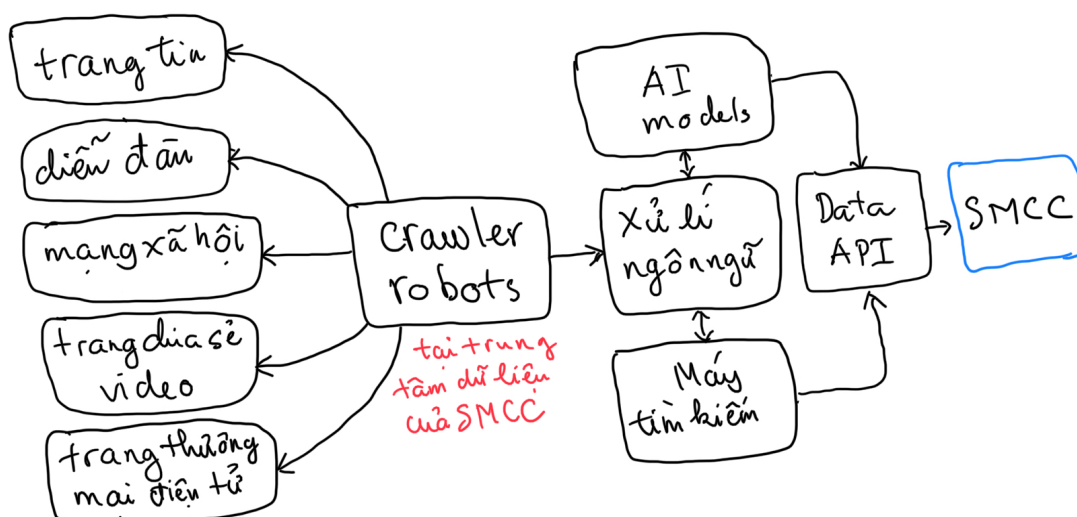
Kỉ niệm 10 tuổi, SMCC đang đặt mục tiêu tiến ra thị trường nước ngoài để cung cấp cho hàng triệu nhà phân tích dữ liệu các công cụ và cách thức làm việc theo hướng data driven hiện đại. Dự án cũng muốn xây dựng một cơ chế “bản vị dữ liệu”, lấy dữ liệu làm đối tượng cơ sở để đánh giá các giá trị khác.

SMCC sở hữu một hệ thống robot tự động có khả năng rà quét thông tin trên diện rộng, công suất thu thập tin tức mỗi ngày có thể lên tới hàng chục triệu bản ghi, hoạt động liên tục 24/24. Tuy nhiên sự bùng nổ internet di động khiến lượng thông tin trên internet ngày càng lớn, hệ thống robot tự động của SMCC chỉ bao phủ được thông tin thị trường Việt Nam. Muốn phục vụ cho thị trường các nước trong khu vực và nhiều quốc gia khác trên thế giới, SMCC cần có hệ thống robot thu thập thông tin mạnh mẽ hơn.

Trong paper này, dự án mô tả một hệ thống thu thập dữ liệu lớn phân tán, dùng phương thức hợp tác với cộng đồng để cùng tạo lập giá trị trên các khối dữ liệu lớn, gọi là Dataverse.vn (Social Media Data Miners). Dataverse.vn cho phép những người sử dụng SMCC cài đặt robot rà quét thông tin trên máy tính của mình để thu thập dữ liệu thô theo sự điều hướng của SMCC, gửi các dữ liệu hữu ích tới máy chủ SMCC để đổi lấy DTVN.

Các giao dịch DTVN được bảo chứng nhờ Ethereum Blockchains Smart Contract⁴, giúp cho DTVN được mã hóa dưới dạng token, có thể tăng giá trị theo thời gian và có thể dễ dàng sử dụng cho việc trao đổi, mua bán, tạo thành một tokenomy⁵ trên không gian số.

Mô hình Dataverse.vn giúp thu hút nhiều người tham gia cộng tác với SMCC để rà quét dữ liệu từ nhiều thị trường khác nhau và hưởng lợi, nhờ vậy SMCC có thể phát triển tới nhiều thị trường lớn trên thế giới.



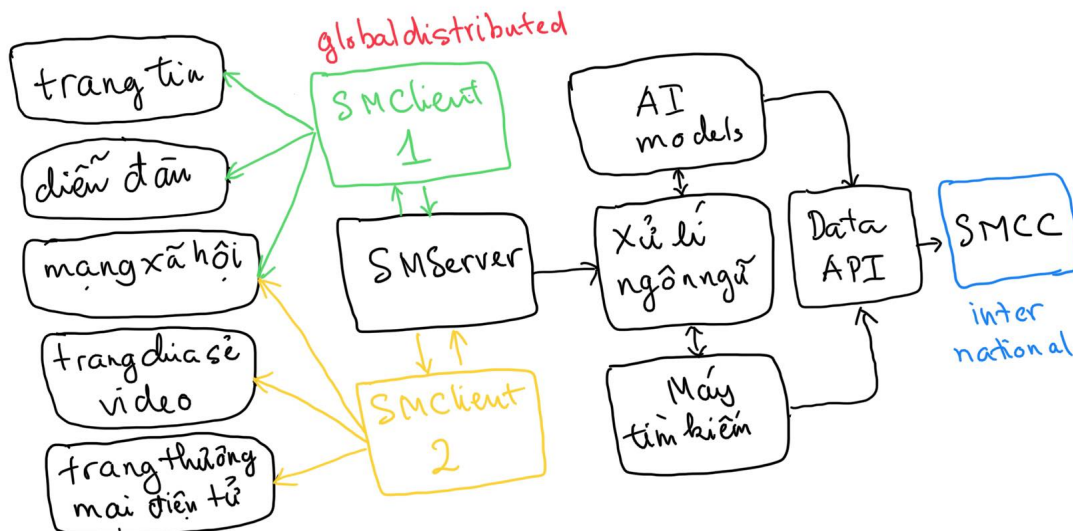
Mô hình thu thập và xử lý dữ liệu hiện tại của SMCC

2. Dataverse.vn: Social Media Data Miners

Có thể coi Dataverse.vn như một sự mở rộng và tiến hóa của hệ thống thu thập dữ liệu đang vận hành SMCC. Các robot của SMCC thay vì chỉ hoạt động dựa trên các máy chủ tập trung có thể được phân tán hóa và hoạt động trên nhiều máy tính ở khắp nơi trên thế giới với những phần cứng không thuộc quyền sở hữu của SMCC.

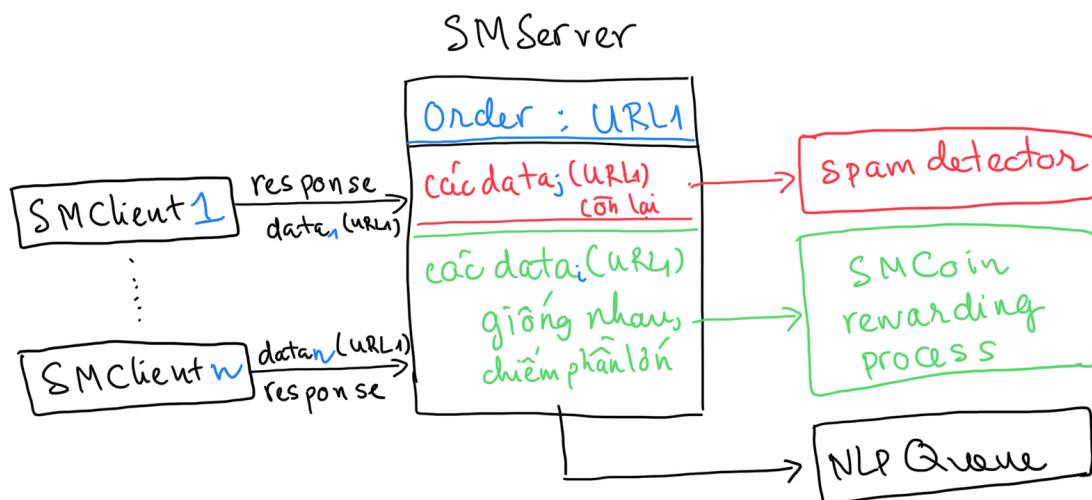
Các máy tính phân tán có thể là máy tính cá nhân, máy chủ hoặc một chiếc điện thoại thông minh được kết nối internet và được cài đặt robot của SMCC (gọi là SMClient). Mỗi SMClient thuộc về một tài khoản trả phí dịch vụ⁶ trên SMCC, mỗi tài khoản có thể nuôi nhiều SMClient. Hệ thống Trí tuệ nhân tạo trên máy chủ tập trung của SMCC (gọi là SMServer) trả về cho mỗi SMClient các nhiệm vụ rà quét dữ liệu khác nhau khi được hỏi.

Để bắt đầu một nhiệm vụ quét dữ liệu, SMClient hỏi SMServer để nhận nhiệm vụ. Hoàn thành nhiệm vụ rồi quét, SMClient gửi dữ liệu thô mình quét được cho SMServer. SMServer xử lý dữ liệu thô, tính toán trùng lặp, lọc spam, chuyển cho các thành phần xử lý tiếp theo trong hệ thống SMCC... và quy đổi các dữ liệu đạt tiêu chuẩn ra DTVN của chủ SMClient (gọi là Data Miner).

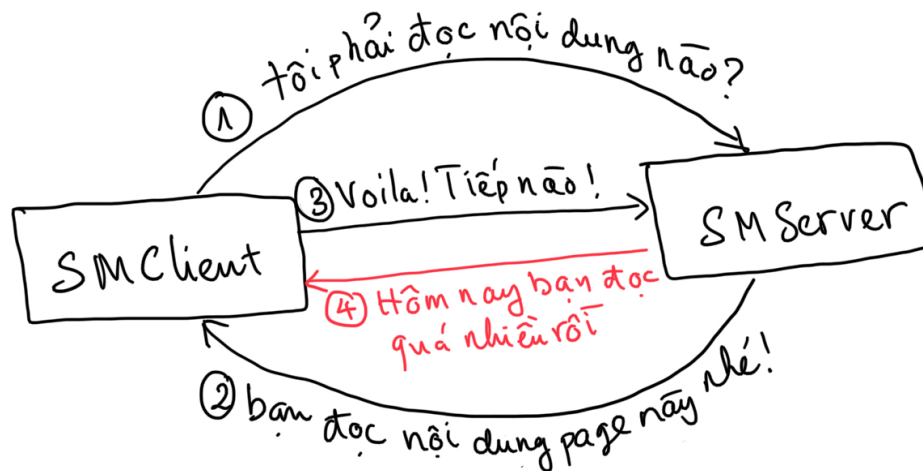


Mô hình thu thập dữ liệu có sự tham gia của các robot phân tán

Khi Data Miner có nhu cầu hiện thực hóa các DTVN mình đang sở hữu, một smart contract dựa trên nền tảng Ethereum theo chuẩn ERC20 kí giữa SMCC và Data Miner sẽ được thực hiện để bảo chứng cho lượng DTVN mà Data Miner được nhận..



3. Đổi dữ liệu lấy token



Mô hình giao tiếp giữa SMServer và SMClient

Cách thức đơn giản nhất để sở hữu DTVN là trở thành Data Miner và điều khiển các SMClient để gửi dữ liệu về SMServer, theo các bước sau:

- Đồng ý với các quy định, đăng kí tài khoản Data Miner miễn phí trên SMCC.VN
- Tải về phần mềm SMClient.
- Đăng kí robot theo tài khoản của Data Miner để nhận robot id.
- Cài đặt SMClient trên các máy tính, điện thoại.
- Cấu hình SMClient theo hướng dẫn, kèm theo robot id đã đăng kí.
- Bật SMClient ở chế độ làm việc khi muốn tham gia đào dữ liệu.

Khi SMClient ở chế độ hoạt động, phần mềm tự giao tiếp với SMServer và tiến hành các công việc thu thập dữ liệu cần thiết theo chỉ dẫn của SMServer. Các dữ liệu tốt, đã được thẩm tra chéo bởi nhiều SMClient khác sẽ được trả công bằng DTVN, ghi nhận trên hệ thống SMCC bằng cách tính công cho Data Miner.

Data Miner có trên một lượng DTVN nhất định (tùy quy định của SMCC tại từng chu kỳ trao đổi) có thể yêu cầu SMCC hiện thực hóa lượng token mình được nhận bằng một giao dịch dựa vào Ethereum Smart Contract kí giữa SMCC với Data Miner, chi phí giao dịch do Data Miner chi trả.

4. Xử lý dữ liệu

Dữ liệu được định hướng thu thập từ SMServer, nhờ vậy dữ liệu tổng hợp từ các SMClient có thể lấy mẫu diện rộng các thị trường cần khảo sát tùy ý.

Ví dụ muốn tập trung khảo sát thị trường Myanmar, SMServer khi được hỏi sẽ gửi cho các SMClient những địa chỉ website lớn, địa chỉ Facebook giàu ảnh hưởng ở Myanmar để làm hạt nhân rà quét. Sử dụng trí tuệ nhân tạo có khả năng xử lý 294 loại ngôn ngữ khác nhau⁷, SMServer có thể chế biến các dữ liệu thô nhận được thành dữ liệu tinh và sắp xếp vào các khu vực phù hợp trong kho dữ liệu của mình. Người dùng SMCC nhờ thế có thể được trải nghiệm thống kê thông tin xã hội, thị trường mà mình quan tâm ở thời gian thực với máy tìm kiếm thống kê liên tục nhận được dữ liệu mới.

Quá trình xử lý dữ liệu trên đã được SMCC ứng dụng ổn định nhiều năm trên kho dữ liệu tiếng Việt với hàng chục tỉ bản ghi. Kiến trúc hạ tầng phần cứng chuyên dụng của công ty InfoRe Technology cho phép mở rộng xử lý tới hàng nghìn tỉ bản ghi với hạ tầng phần cứng lắp thêm. Vì thế, SMCC sẽ trở nên mạnh mẽ hơn nhiều so với hiện nay khi Dataverse.vn hoạt động.

5. Cơ máy Dataverse.vn

Dataverse.vn tạo thành một cơ máy lớn gồm cả phần mềm và phần cứng, một hệ thống mạng lưới các robot được kết nối tập trung với máy chủ của SMCC, chuyên rà quét và thu thập dữ liệu. Công suất hoạt động của cơ máy này có thể lớn hơn khả năng tiêu thụ của hệ thống SMCC, vì thế SMCC sẽ không phải là hệ thống duy nhất sử dụng Dataverse.vn.

Các hệ thống Datalake, Data Platform... của các doanh nghiệp lớn, doanh nghiệp có nhu cầu làm giàu dữ liệu, cập nhật dữ liệu một cách thường xuyên có thể thông qua cơ chế

của Dataverse.vn để dùng DTVN trả cho các Data Miner nhằm đổi lấy dữ liệu cần cập nhật, giống như SMCC.

Nhu cầu thu thập và sử dụng dữ liệu của các hệ thống dữ liệu lớn thông qua Dataverse.vn giúp thu hút nhiều Data Miner tham gia vào hệ thống. Khả năng cung cấp dữ liệu thường xuyên của các Data Miner giúp họ tích lũy được nhiều DTVN và hình thành tài sản. Các DTVN đóng vai trò của vật ngang giá cho quá trình cung - cầu dữ liệu này, giúp hình thành một nền kinh tế thu nhỏ quanh đồng DTVN - một Tokenomy.

6. Tokenomics

Để khởi tạo cho hệ sinh thái của Dataverse.vn với DTVN, SMCC lưu trữ 10 tỉ bản ghi dữ liệu đã thu thập được từ 2016 tới nay trên hệ thống và phát hành 1 tỉ DTVN cho chính mình. Các DTVN này được chia làm 5 phần:

- 32% tương ứng 320 triệu DTVN sử dụng để phát hành cho các nhà đầu tư sớm (Data Investor) qua hình thức Token Gift (xem mục số 6).
- 32% tương ứng 320 triệu DTVN dành cho SMCC với cam kết không chuyển nhượng trong vòng 3 chu kì (trừ việc sử dụng cho token burning).
- 32% tương ứng 320 triệu DTVN được sử dụng để trả thưởng cho các Data Miners, khuyến khích việc thu thập và đóng góp dữ liệu.
- 2% tương ứng 20 triệu DTVN dành cho việc thúc đẩy chiến dịch affiliate marketing (tiếp thị giới thiệu) giữa các nhà đầu tư sớm để quá trình token sale diễn ra thuận lợi, định vị được nhiều nhà đầu tư có tầm nhìn tương tự nhau. Các DTVN không sử dụng hết được dùng để tặng thưởng xuyên cho các quỹ từ thiện và tặng học bổng cho sinh viên.
- 2% tương ứng 20 triệu DTVN dành cho phát triển chương trình influencer marketing (tiếp thị với người có tầm ảnh hưởng), thông qua những người nổi tiếng trong cộng đồng đầu tư blockchain, dữ liệu lớn, trí tuệ nhân tạo... được nhận DTVN để quảng bá cho dự án Dataverse.vn.

SMServer gửi cho nhiều SMClient ngẫu nhiên khác nhau các yêu cầu rà quét dữ liệu giống nhau (các SMOrder) và đối chiếu chéo các bản ghi nhận được (các SMResponse) để tránh dữ liệu bị làm giả.

Với số lượng n đủ lớn các SMClient gửi về cùng một SMResponse giống nhau trong cùng một cửa sổ thời gian và m là mức giá trả thưởng cho mỗi bản ghi được kết nạp vào kho dữ liệu của SMCC, mỗi bản ghi chính xác SMClient nhận được m/n DTVN cho Data Miner của mình.

Mỗi khi số lượng dữ liệu của SMCC tăng lên thêm 10 tỉ bản ghi so với thời điểm bắt đầu của chu kì trước đó, m sẽ giảm đi một nửa do tổng lượng DTVN trả thưởng thay đổi ở chu kì mới. Mỗi chu kì được kì vọng giúp SMCC thu thập được 10 tỉ bản ghi dữ liệu.

Phần DTVN được sử dụng để chia sẻ cho các Data Miner tại một chu kì luôn được chia làm hai nửa bằng nhau, một nửa dùng để chia cho tất cả các Data Miner góp được số dữ liệu kì vọng trong chu kì, một nửa dùng để trả cho Data Miner ở các chu kì tiếp sau.

Chu kì số	Tổng lượng dữ liệu mục tiêu	Lượng DTVN trả thưởng	m
1	10 tỉ bản ghi	160 triệu DTVN	0.016
2	10 tỉ bản ghi	80 triệu DTVN	0.008
3	10 tỉ bản ghi	40 triệu DTVN	0.004
4	10 tỉ bản ghi	20 triệu DTVN	0.002
Các chu kì thu nhận dữ liệu lặp lại vô tận với m nhỏ đi một nửa ở mỗi chu kì mới			

Data Miner cần là người dùng miễn phí của SMCC nhưng cũng có thể dùng DTVN hoặc VNĐ để nâng cấp các gói dịch vụ của mình để quản lí được nhiều SMClient hơn.

SMCC cũng đẩy mạnh kinh doanh tài khoản dưới dạng SaaS ở nhiều thị trường khác nhau trên thế giới nhằm nâng cao lợi nhuận của công ty và đảm bảo giá trị cho DTVN. Lợi nhuận thu được từ việc kinh doanh của SMCC được sử dụng để tái đầu tư cho hệ thống phần cứng và công nghệ, giúp SMCC duy trì được hệ thống dữ liệu lớn.

Data Miner có thể trao đổi, giao dịch peer to peer với nhau qua Smart Contract. Khi cộng đồng Data Miner đủ lớn, có phân bố rộng rãi, cùng xây dựng được khối dữ liệu có giá trị, DTVN sẽ được đưa lên các sàn giao dịch token để hòa vào mạng lưới tiền mã hóa toàn cầu, giúp các Data Miner và Data Investor nâng cao tính thanh khoản của DTVN.

Lực cung dữ liệu

Khi giá của DTVN giảm dưới mức trung bình mà một Data Miner cần bỏ ra để nhận được 1 DTVN, các Data Miner và SMCC sẽ có xu hướng mua vào nhiều DTVN hơn, giúp giá DTVN gia tăng. Khi giá DTVN gia tăng hơn mức chi phí trung bình để có được 1 DTVN, các Data Miner tham gia vào mạng lưới nhiều hơn, kích thích đẩy nhanh các chu kỳ thu thập dữ liệu. Sau một số chu kỳ, nếu lượng DTVN trả thưởng dành cho Data Miner trở nên thiếu hấp dẫn, SMCC có thể nâng cao tỉ lệ trả thưởng bằng các DTVN mình đang nắm giữ.

Lực cầu dữ liệu

SMCC là hệ thống đóng vai trò thúc đẩy nhu cầu dữ liệu ở giai đoạn ban đầu. Tuy nhiên, để đẩy mạnh nhu cầu dữ liệu để tạo thành động cơ giao dịch cho tokenomy của Dataverse, dự án sẽ hợp tác với các doanh nghiệp lớn đang sở hữu Datalake, CDP - Customer Data Platform, Data Warehouse... như các ngân hàng, các tổ chức tín dụng, các doanh nghiệp bán lẻ có tệp khách hàng lớn, các doanh nghiệp cung cấp platform... để cùng thụ hưởng các dòng chảy dữ liệu từ Dataverse. Các doanh nghiệp có nhu cầu cập nhật, rà quét dữ liệu internet thông qua Dataverse để tạo ra đầu vào dữ liệu cho hệ thống của mình có thể thu mua DTVN và thông qua Dataverse trả cho các Data Miner theo cách tương tự như SMCC.

Nguồn cung dữ liệu từ các Data Miner và nguồn cầu dữ liệu từ SMCC và các doanh nghiệp lớn giúp duy trì cân bằng và tạo ra giao dịch liên tục cho DTVN. Nhu cầu làm giàu dữ liệu của các doanh nghiệp đang ngày một lớn và Dataverse là một giải pháp dễ dàng, đơn giản, chi phí không cao, nhờ vậy sẽ tạo ra một tokenomy ổn định và có tiềm năng phát triển mạnh mẽ.

7. Token Gift

Dự án sử dụng 320 triệu DTVN được chia làm nhiều suất quà tặng để tặng kèm khách hàng của SMCC mua các tài khoản sử dụng SMCC Premium trong vòng ít nhất 1 năm với mức giá 8 triệu đồng/tháng. 28 khách hàng sớm nhất của chương trình được tặng 800 nghìn DTVN, 100 khách hàng tiếp theo được tặng 400 nghìn DTVN, 256 khách hàng tiếp theo được tặng 200 nghìn DTVN và cứ như vậy cho tới khi hết 320 triệu DTVN trong

chương trình token gift. Số lượng DTVN trong mỗi gói quà tặng có thể được điều chỉnh trong từng giai đoạn để phù hợp với tình hình thị trường.

Với SMCC Premium, người sử dụng có thể tìm ra nhiều dữ liệu phục vụ tốt cho các quá trình kinh doanh của mình, có thể bán lại quyền sử dụng cho người dùng khác hoặc có thể thường xuyên đúc phiên bản NFT của các nội dung mà mình muốn lưu giữ, thấy có giá trị để xây dựng kho tài nguyên số cho bản thân.

Dự án coi các khách hàng này là những Nhà đầu tư sớm cho DTVN và gọi họ là các *Early Data Adopter (EDA)*. Các EDA có thể dùng chính SMCC để kiểm tra chất lượng dữ liệu của toàn mạng Dataverse.vn và định hướng thường xuyên cho kế hoạch đầu tư vào DTVN của mình.

Các DTVN ban đầu chỉ được chuyển giao dưới dạng quà tặng nhưng khi nhu cầu mở tài khoản của Data Miner trên SMCC gia tăng hoặc khi tỉ lệ trả thưởng DTVN cho mỗi bản ghi dữ liệu thay đổi (vì tổng lượng dữ liệu tăng lên gấp đôi), các EDA có thể thu được nhiều lợi ích từ lượng DTVN mình đang nắm giữ. Các EDA đồng thời là Data Miner thường có thể có thêm các lợi ích kép do tham gia đào dữ liệu sớm.

Kể từ đầu năm 2022, mỗi quý, tùy vào tình hình thị trường, tối đa 1% số DTVN của SMCC có thể được burn để kích thích việc tăng giá trị quy đổi của DTVN sang các coin khác. Lượng DTVN được burn trích từ 32% DTVN ban đầu mà SMCC giữ cho mình. Nếu không burn, SMCC có thể airdrop cho những người sử dụng SMCC và các Data Miner.

8. Kế hoạch marketing

Để thúc đẩy việc thực hiện Token Gift và phát triển cộng đồng Data Miner được thuận lợi, dự án dự kiến triển khai hai hình thức marketing chính là Tiếp thị liên kết - giới thiệu (Affiliate Marketing)¹⁰ và Tiếp thị với người có tầm ảnh hưởng (Influencer Marketing)¹¹. Dự án có các marketing campaign chính sau:

8.1. *Affiliate Marketing*

Để khuyến khích các EDA giới thiệu dự án Dataverse.vn cho nhau, dự án có chính sách tặng thưởng thời hạn sử dụng SMCC Premium và DTVN cho nhà đầu tư giới thiệu được

các nhà đầu tư khác, theo các luật sau:

- Hình dung, Tuấn là một người dùng bất kì của SMCC.
- Tuấn thấy tiềm năng của dự án Dataverse.vn, Tuấn giới thiệu cho Hưng trở thành EDA. Khi điền form đăng kí trở thành EDA, Hưng cung cấp địa chỉ email của Tuấn.
- Nhờ giới thiệu Hưng thành EDA, Tuấn nhận được thêm 2 tháng sử dụng tài khoản SMCC Premium và được tặng thêm 20000 DTVN.
- Nhờ Tuấn là người dùng SMCC giới thiệu, Hưng được giảm 16 triệu đồng, chỉ phải trả 80 triệu đồng cho mỗi 1 năm thuê bao SMCC Premium có tặng kèm Token Gift.

Khi điền form đăng kí trở thành EDA. Mỗi người dùng SMCC có thể giới thiệu nhiều EDA nhưng tổng số EDA tối đa của Dataverse.vn không vượt quá 400 (do chỉ có tối đa 400 suất Token Gift).

Tổng lượng DTVN dùng để tặng thưởng trong marketing module này là 20 triệu DTVN. Khi Dataverse.vn không còn suất Token Gift nào, lượng DTVN còn lại của campaign 8.1 sẽ được dành cho campaign 8.5.

8.2. Influencer Marketing

Dựa vào các phân tích về mức độ ảnh hưởng và thể loại nội dung của những người có tầm ảnh hưởng trên mạng xã hội do đội ngũ của SMCC cung cấp, dự án tiến hành tiếp cận các influencer này và tặng các gói DTVN để nhận được sự hỗ trợ truyền thông của họ.

Tổng số DTVN cho việc này là 20 triệu DTVN, chiếm 2% tổng lượng DTVN được lưu hành. Influencer Marketing được kì vọng sẽ giúp Dataverse.vn xây dựng nhanh chóng cộng đồng các Data Miner ban đầu là những người Việt Nam nhờ sự ảnh hưởng trên mạng xã hội của các Influencer.

8.3. Email Marketing

Hệ thống SMCC hiện đang có 25 nghìn người đã đăng kí sử dụng. Dự án gửi email mời người dùng của SMCC tham gia chương trình Token Gift và mời người dùng SMCC tham gia vào dự án Affiliate Marketing ở phần 8.1. Quà tặng mà những người giới thiệu được nhận với mỗi lượt thành công là 2 tháng sử dụng SMCC miễn phí và 20 nghìn

DTVN [theo 8.1]. Marketing campaign này được kì vọng sẽ nhận được sự hưởng ứng của nhiều người yêu quý và mong muốn SMCC trở nên tốt hơn nữa.

8.4. Các biện pháp marketing khác

Dựa trên dữ liệu phân tích của SMCC, dự án có thể thực hiện nhiều chiến dịch marketing đa kênh như sử dụng robot gọi điện bằng trí tuệ nhân tạo¹², sử dụng quảng cáo trực tuyến, tổ chức hội thảo trực tuyến vv... để tiếp cận các EDA và các Data Miner.

8.5. Học bổng và từ thiện

Lượng DTVN của campaign 8.1 sau khi Dataverse.vn đã tặng hết 400 suất Token Gift được chuyển thành DTVN để tặng cho các quỹ từ thiện và tặng học bổng cho sinh viên. Khi cộng đồng cùng tạo dựng được giá trị cao cho DTVN, đây sẽ là các khoản tài trợ có nhiều ý nghĩa. Việc này cũng giúp tạo dựng cộng đồng Data Miner mạnh hơn từ sự tham gia của giới sinh viên, góp phần phát triển giá trị và các hướng ứng dụng cho DTVN.

8.6. Nhóm các nhà đầu tư sớm

Dự án chia nhà đầu tư sớm ra làm 4 nhóm, căn cứ theo thời gian tham gia:

28 nhà đầu tư sớm đầu tiên của Dataverse.vn được gọi là *Innovative Investor*. Innovative Investor được tặng thêm 3 năm sử dụng SMCC Premium miễn phí (ngoài thời lượng đã trả phí), trị giá 288 triệu đồng. Các Innovative Investor được công nhận là sáng lập viên của dự án Dataverse.vn và cùng tham góp ý kiến định hướng xây dựng dự án cho đội ngũ phát triển.

Các nhà đầu tư từ thứ 29 đến thứ 128 được tặng thêm 2 năm sử dụng SMCC Premium trị giá 192 triệu đồng (Nhóm *Early Adopter*).

Các nhà đầu tư từ thứ 129 đến 256 được tặng thêm 1 năm sử dụng SMCC Premium trị giá 96 triệu đồng (Nhóm *Early Majority*).

Các nhà đầu tư có thể mua gói sử dụng SMCC Premium nhiều năm để được tặng nhiều gói DTVN nhưng chỉ nhận thêm tối đa không quá 3 năm quyền sử dụng SMCC Premium. Tất cả các nhà đầu tư sớm đều có thể chuyển nhượng quyền sử dụng SMCC Premium cho các cá nhân, tổ chức khác.

9. Lộ trình phát hành token

Sau khi có đủ 28 Innovative Investor, SMCC sẽ tham vấn chính các nhà đầu tư sớm nhất này và công bố lộ trình phát hành token chi tiết phù hợp, ưu tiên quyền lợi của các Innovative Investor. Dự kiến lộ trình phát hành token của Dataverse.vn sẽ được công bố vào tháng 11/2021 và bổ sung vào mục này.

Theo dự kiến hiện tại, các nhà đầu tư sớm được nhận token khi nhóm của mình đạt số lượng thành viên tối đa. Ví dụ, khi nhóm *Innovative Investor* có đủ 28 thành viên, dự án bắt đầu chuyển các gói token cho các thành viên của nhóm này. Tương tự đối với các nhóm Early Adopter khi dự án có nhà đầu tư thứ 128, nhóm Early Majority với nhà đầu tư thứ 256 và các nhà đầu tư còn lại khi số Token Gift được phát hết.

10. Kế hoạch kinh doanh SMCC

10.1. Tiếp cận thị trường quốc tế

Trước mắt, khi dữ liệu từ Dataverse.vn ổn định, chúng tôi tiếp cận phân tích các fanpage sử dụng tiếng Anh và cung cấp công cụ SMCC cho các online marketer đang bán hàng ở thị trường Mỹ, Anh, Úc... để phục vụ nắm bắt xu thế thị trường.

Chúng tôi cũng đang đàm phán hợp tác với một số doanh nghiệp công nghệ Việt Nam đã cung cấp dịch vụ, sản phẩm ở các thị trường nước ngoài để cùng phân phối sản phẩm SMCC (có thể dưới dạng white label hoặc rebranding).

10.2. Dịch vụ làm giàu dữ liệu bằng DTVN

Công suất thu thập dữ liệu của Dataverse.vn có thể lớn hơn nhu cầu khai thác của SMCC trong lúc nhiều doanh nghiệp có nhu cầu bổ sung, làm giàu dữ liệu cho các Data Platform của mình. Vì thế, một hướng kinh doanh khác của SMCC là cung cấp dịch vụ làm giàu dữ liệu cho các doanh nghiệp khác. Dịch vụ này có thể được đổi bằng DTVN để kích thích nhu cầu trao đổi DTVN.

10.3. Cung cấp công cụ NFT các di sản số của cá nhân

Ngày 21 tháng 3 năm 2006, Jack Dorsey (@jack), sáng lập viên của Twitter, đã đăng lên mạng xã hội này dòng trạng thái đầu tiên: “just setting up my twttr”. Dòng tweet này gần

đây đã được Jack Dorsey đúc thành NFT và được bán đấu giá trên mạng. Tính đến ngày 23/3/2021, phiên bản NFT duy nhất của dòng tweet đầu tiên này đã được mua với giá 2.9 triệu USD. Nhiều nghệ sĩ, người nổi tiếng khác cũng bắt đầu bán được các vật phẩm NFT với giá cao. Chúng tôi cho rằng cùng với thời gian, kho các vật phẩm NFT có giá trị sẽ ngày một lớn lên và thu hút mọi người tham gia xây dựng tài nguyên số của bản thân. Chúng tôi đang phát triển trên SMCC một chức năng hỗ trợ người dùng Facebook bất kì đúc ra phiên bản NFT cho các bài đăng yêu thích của mình, kèm theo các chỉ số phân tích lượng tương tác mạng xã hội với mỗi bài đăng. Đây có thể là một cách để người dùng của SMCC lưu giữ các tài sản số có giá trị với bản thân mình và có thể có giá trị lớn trong tương lai.

11. Giá trị lõi của DTVN

Dữ liệu ngày nay được ví như một loại dầu mỏ mới. Giống như dầu mỏ, dữ liệu có thể sử dụng được cho rất nhiều việc. Dữ liệu là nguồn tài nguyên chính cung cấp cho các hệ thống máy móc hoạt động dựa trên trí tuệ nhân tạo, kiến thiết khả năng ra quyết định của các tổ chức, doanh nghiệp và cá nhân trong thời đại mới. Vì vậy chúng tôi tin tưởng sâu sắc vào giá trị của các quá trình thu thập, tổ chức, xử lý dữ liệu trong việc làm lợi cho xã hội. DTVN được hình thành để thúc đẩy SMCC trong việc thực hiện mạnh mẽ hơn nữa các quá trình ấy.

Có thể hình dung toàn bộ mạng lưới quét dữ liệu mà dự án Dataverse.vn hình thành nên như một hệ thống đơn nhất. Hệ thống đó là cỗ máy lớn được hợp thành từ nhiều máy tính nhỏ của nhiều người trên internet, cùng rà quét dữ liệu để đổi lấy DTVN. Cỗ máy ấy cũng có thể đáp ứng nhu cầu làm giàu dữ liệu, tổng hợp dữ liệu số lượng lớn của các tổ chức, doanh nghiệp hiện đại. Số lượng DTVN là hữu hạn nên có thể coi DTVN là cổ phần của cỗ máy ấy.

DTVN với vai trò gắn kết và khích lệ cộng đồng, có giá trị được hình thành từ nhiều yếu tố, trong đó có 6 yếu tố cơ bản sau:

Một là từ chính công sức, chi phí máy móc của các Data Miner. Để đánh đổi các chi phí ấy lấy những DTVN, các Data Miner hẳn sẽ trân quý, tiết kiệm và biết sử dụng DTVN của mình cho các mục tiêu đầu tư, tích lũy hợp lí.

Hai là từ sự hội tụ giá trị của các khối dữ liệu nhỏ khi hình thành nên dữ liệu lớn, khối dữ liệu được sử dụng làm “bản vị” cho DTVN. Giá trị của khối dữ liệu lớn mà các Data Miner cùng nhau tích tụ được cũng chính là giá trị của DTVN, sẽ ngày một nhiều lên.

Ba là từ sự nỗ lực ứng dụng công nghệ mới của SMCC để tổ chức, sắp xếp và khai phá dữ liệu, biến các dữ liệu thô thành dữ liệu tinh và cung cấp được các tri thức giá trị cho người sử dụng SMCC ở nhiều quốc gia khác nhau.

Bốn là từ sự khan hiếm tự nhiên của DTVN, khi tổng số DTVN được phát hành được giới hạn cố định là 1 tỉ coin. Điều này có thể làm giá trị của mỗi DTVN tăng tự nhiên khi cộng đồng có nhiều data miner.

Năm là từ sự liên kết DTVN với các sản phẩm dịch token. SMCC cùng những người sở hữu DTVN có thể cùng tạo ra một token mạnh, có giá trị cao và được niêm yết trên các sàn giao dịch lớn để gia tăng giá trị cũng như tính thanh khoản của DTVN.

Sáu là nhu cầu làm giàu, liên thông, cập nhật dữ liệu ngày một lớn của các doanh nghiệp, đặc biệt là các doanh nghiệp đã trang bị cho mình các hệ thống data platform. Các doanh nghiệp đó có thể tìm thấy ở Dataverse nhiều giá trị, qua đó thúc đẩy tích trữ, giao dịch DTVN thường xuyên.

Dự án khuyến khích các nhà đầu tư sớm giữ các DTVN của mình đủ lâu và tham gia sớm cả vào mạng lưới các Data Miner để được nhận thêm nhiều giá trị và tạo dựng một cộng đồng khai thác dữ liệu cùng nhau phát triển.

12. Định hướng xa

Có thể hình dung Dataverse.vn như một cỗ máy thống nhất. Cỗ máy này được xây dựng từ phần cứng của nhiều người và cùng thu thập các mẫu dữ liệu social media trên internet. Với nhiều nguồn dữ liệu social media khác nhau được tập hợp, SMCC có thể cung cấp các ứng dụng như:

- Máy tìm kiếm đa nhiệm “Make In Vietnam”
- Phát hiện sản phẩm tiềm năng, sản phẩm được ưa chuộng trên thị trường theo từng thời điểm khác nhau, phục vụ tối ưu chiến lược các mặt hàng dropshipping.

- Phân tích thông tin, phát hiện sớm thông tin tích cực, tiêu cực của thương hiệu, dịch vụ, nhân hiệu nhằm chăm sóc khách hàng, phát triển thương hiệu, nhân hiệu.
- Phân tích nhu cầu của các nhóm khách hàng tiềm năng, lựa chọn các nhóm khách hàng thích hợp với sản phẩm cần triển khai thương mại hóa trên nền các kênh digital marketing.
- Nắm bắt các vấn đề dư luận, xã hội, khảo sát dư luận khách quan dựa trên dữ liệu thay vì hỏi đáp thông thường.
- Hỗ trợ thu thập, phân tích thông tin thị trường tài chính để ra quyết định đầu tư, nhất là các lĩnh vực đầu tư crypto assets với nhiều thông tin phi chính thống.
- Tuyển dụng nhân sự bằng các phương pháp tiếp cận sử dụng marketing.
- Làm giàu dữ liệu cho các data platform của doanh nghiệp, tổ chức.
- Hỗ trợ quản trị rủi ro, xây dựng các hệ thống đánh giá chỉ số tín dụng.
- Đánh giá influencer, phát triển chiến lược affiliate marketing và social commerce.
- Phát triển hệ thống trí tuệ nhân tạo phục vụ điều phối mạng quảng cáo.

Nhiều hướng ứng dụng khác cũng có thể được đề xuất và hợp tác bởi các Data Miner trong quá trình vận hành robot, tùy vào sự phát triển của các ứng dụng mới trên internet. Nếu nhu cầu tham gia rà quét dữ liệu của cộng đồng Data Miner tăng mạnh, dự án có thể cung cấp tài nguyên của Dataverse.vn cho nhiều doanh nghiệp khác như một dịch vụ rà quét và làm giàu dữ liệu công cộng.

SMCC cũng xác định sử dụng ít nhất 20% lợi nhuận kinh doanh hằng năm của mình để đầu tư vào các DTVN đang được lưu hành nhằm phát triển giá trị của đồng token. Xa hơn nữa, SMCC hướng tới mục tiêu mở toàn bộ hệ tri thức mà mình thu thập được để phục vụ miễn phí xã hội, giống như những gì wikipedia đã và đang làm.

13. Hiện thực hóa lợi nhuận - thu hồi vốn sớm

Các nhà đầu tư sớm vào dự án Dataverse hoàn toàn có thể hiện thực hóa sớm lợi nhuận hoặc thu hồi vốn của mình bằng việc chuyển giao quyền sử dụng tài khoản SMCC Premium cho các doanh nghiệp, cá nhân có nhu cầu sử dụng. Nhờ được tặng thêm 3 năm sử dụng SMCC Premium (giá thị trường 96 triệu đồng / năm) và được hỗ trợ chuyển nhượng quyền sử dụng tài khoản nếu có nhu cầu, các nhà đầu tư bằng các quan hệ cá nhân có thể thu hồi vốn sớm, thậm chí có lãi mà vẫn giữ lại được phần DTVN của mình.

14. Take away

Trong paper này, dự án mô tả một mô hình ứng dụng công nghệ blockchain dựa trên Ethereum Blockchain Smart Contract theo chuẩn ERC20, giúp vận động tài nguyên, công sức từ nhiều người khác nhau trên thế giới để xây dựng đầu vào cho một hệ cơ sở dữ liệu lớn cấp độ toàn cầu.

Mô hình của dự án cũng giúp tạo ra một nghề nghiệp mới cho nhiều người lao động trên thế giới: nghề sử dụng máy tính để rà quét dữ liệu. Chỉ với những chiếc máy tính, điện thoại đơn giản, có cấu hình nhẹ và được kết nối internet, người dân sống ở các vùng khác nhau trên khắp thế giới có thể tham gia vào Dataverse.vn để trực tiếp biến dữ liệu thành vàng.

Với bộ dữ liệu social media lớn được đóng góp bởi rất nhiều người, SMCC cũng hi vọng sẽ tạo ra được ở Việt Nam một kho tàng tri thức của nhân loại, giống như những gì Wikipedia hay Google đã làm. Và nhờ thế, DTVN sẽ là một trong những token có nhiều giá trị trong tương lai.

Mô hình này cũng giúp tạo ra một cỗ máy thể hệ mới, được hợp thành từ nhiều thiết bị tính toán nhỏ phân bố khắp nơi trên internet, có khả năng cùng thực hiện tác vụ thu thập dữ liệu lớn. Chúng tôi tin đây là một cách hữu ích và hợp pháp để các doanh nghiệp làm giàu dữ liệu của mình trong thời đại mà dữ liệu ngày càng giúp sản sinh các thông tin có ích.

Một mẫu thông tin về nỗ lực của Mị Châu lộ cho Trọng Thủy có thể khiến Triệu Đà đánh đổi được cả một quốc gia. Vậy thì một cỗ máy được tập hợp từ nhiều máy tính phân tán khắp nơi có khả năng rà quét dữ liệu trên diện rộng, theo nhu cầu của không chỉ SMCC mà có thể nhiều doanh nghiệp, tổ chức có nhu cầu cập nhật thông tin số lượng lớn khác, sẽ có giá trị đến đâu? 1 tỉ đồng DTVN đại diện cho giá trị và giống như cổ phần của cỗ máy ấy. Dự án rất mong được cùng hợp tác với các quý vị để tạo dựng nhiều giá trị mới từ hệ thống quét dữ liệu dựa trên công nghệ blockchain này.

**ĐỀ HỢP TÁC VỚI DỰ ÁN
XIN QUÝ VỊ VUI LÒNG LỰA CHỌN**

**ĐĂNG KÍ TRỞ THÀNH
NHÀ ĐẦU TƯ SỚM CỦA DTVN
<https://smcc.vn/token-gift.html>**

**CHAT VỚI INFORE ĐỂ HỎI THÊM THÔNG TIN
<https://m.me/inforetechnology>**

Các trích dẫn tham khảo

1. Trang web cung cấp dịch vụ của SMCC <https://smcc.vn>
Trang Facebook của SMCC <https://fb.com/smcc.vn>
2. Giới thiệu sơ lược công ty InfoRe <https://smcc.vn/infore.pdf>
Trang Facebook của InfoRe <https://fb.com/inforetechnology>
3. Thủ lĩnh quán quân Nhân tài Đất Việt 2016: “StartUp không phải màu hồng!”
<https://dantri.com.vn/suc-manh-so/thu-linh-quan-quan-nhan-tai-dat-viet-2016-startup-khong-phai-mau-hong-20161121080254246.htm>
4. Introduction to smart contracts
<https://ethereum.org/en/developers/docs/smart-contracts>
5. What is Tokenomics?
<https://coinmarketcap.com/alexandria/article/what-is-tokenomics>
6. Bảng giá phí dịch vụ sử dụng SMCC <https://smcc.vn/pricing.aspx>
7. Understand the ERC-20 Token Smart Contract
<https://ethereum.org/en/developers/tutorials/understand-the-erc-20-token-smart-contract/>
8. Facebook Fasttext pretrained vectors
<https://github.com/facebookresearch/fastText/blob/master/docs/pretrained-vectors.md>
9. Explained: What is coin burning in the cryptocurrency universe?
<https://economictimes.indiatimes.com/markets/cryptocurrency/explained-what-is-coin-burning-in-the-cryptocurrency-universe/articleshow/85658968.cms>
10. Tiếp thị liên kết <https://adflex.vn/tiep-thi-lien-ket-la-gi/>
11. Tiếp thị thông qua người ảnh hưởng: Nên xem họ như những sứ giả nhãn hiệu
<https://www.brandsvietnam.com/18214-Tiep-thi-thong-qua-nguoi-anh-huong-Nen-xem-ho-nhu-nhung-su-gia-nhan-hieu>
12. TP.HCM triển khai robot gọi hỏi thăm sức khỏe để lấy mẫu xét nghiệm kịp thời
<https://tuoitre.vn/tp-hcm-trien-khai-robot-goi-hoi-tham-suc-khoe-de-lay-mau-xet-nghiem-kip-thoi-20210602152636295.htm>

COMPANY PROFILE

InfoRe Technology

Golden Palace, 99 Mễ Trì

Nam Từ Liêm

082 323 4576

orm@infore.vn

Giới thiệu

Công ty cổ phần công nghệ Chọn lọc thông tin (Information Retrieval Technology) được thành lập ngày 6/9/2011 bởi nhóm kỹ sư - nhà nghiên cứu chuyên về trí tuệ nhân tạo và dữ liệu lớn. Từ khi thành lập đến nay, InfoRe đã tham gia vào nhiều dự án xử lý dữ liệu và phân tích dữ liệu với nhiều đối tác lớn, góp phần giải quyết các bài toán chuyển đổi số với dữ liệu ở Việt Nam.

Kinh nghiệm

Phân tích dữ liệu, tư vấn giải pháp xử lý thông tin – Since 2012

- Khách hàng tiêu biểu: ACB, BIDV, Hanoi Telecom, HDBank, Viettel Telecom, Huawei, Reddi...
- Dịch vụ chính: Cung cấp báo cáo phân tích dữ liệu; Cung cấp công cụ phần mềm; Cho thuê chuyên viên phân tích dữ liệu.

Xử lý Dữ liệu lớn – Since 2011

- Phát triển và duy trì hệ thống lắng nghe mạng xã hội SMCC.VN, phân tích, lưu trữ và phục vụ tìm kiếm hàng triệu thông tin social media mỗi ngày.
- Xử lý dữ liệu dịch tễ theo yêu cầu của Ban chỉ đạo quốc gia về phòng, chống COVID-19
- Tham gia xử lý dữ liệu về liệt sĩ của toàn quốc.
- Khách hàng tiêu biểu: Một số cơ quan nhà nước.

Ứng dụng Trí tuệ nhân tạo – Since 2011

- Sử dụng NLP để phân tích tự động nhiều bài toán ngữ nghĩa của SMCC.VN
- Ứng dụng dữ liệu mạng xã hội để dự đoán chỉ số tín dụng, đã triển khai thử nghiệm ở một số ngân hàng lớn.
- Phát triển hệ thống nhận dạng giọng nói tiếng Việt tự động dựa trên mạng nơ-ron.
- Phát triển hệ thống nhận dạng khuôn mặt phục vụ chấm công tại các ngân hàng.
- Khách hàng tiêu biểu: Một số cơ quan nhà nước.

Nhân lực

Đội ngũ phân tích dữ liệu – 63 người

Chủ yếu là sinh viên trường Đại học Ngoại thương và Đại học Kinh tế Quốc dân, đã trải qua hơn 1 năm làm nhiệm vụ phân tích dữ liệu dịch tễ theo yêu cầu chống dịch. Ngoài ra, các nhân viên phân tích dữ liệu này được huấn luyện phương pháp sử dụng SMCC.VN để phân tích thị trường, phục vụ nhu cầu phân tích dữ liệu của khách hàng.

Đội ngũ xử lý dữ liệu – 8 người

Đội ngũ xử lý dữ liệu của InfoRe tập trung cho việc phát triển và duy trì phần mềm SMCC.VN. Ngoài ra, đội ngũ này còn phát triển một số dự án startup của InfoRe nhằm thí nghiệm các công nghệ với Trí tuệ nhân tạo và Dữ liệu lớn.

Đội ngũ tư vấn cao cấp – 7 người

Đội ngũ tư vấn cao cấp là các đồng sáng lập InfoRe, các quản trị viên công ty. Đội ngũ này hướng dẫn, hỗ trợ cho nhân sự cấp cao trong công ty liên quan tới các dự án phân tích và xử lý dữ liệu. Đội ngũ có 1 Giáo sư, 2 tiến sĩ, 3 thạc sĩ, 1 chuyên gia trong các lĩnh vực Vật lý, Xử lý ngôn ngữ, Hệ đa tác tử, Truyền thông và marketing, Dữ liệu lớn, Trí tuệ nhân tạo hiện đang sinh sống và làm việc tại Việt Nam, Nhật Bản, Mỹ, Úc và New Zealand.

Giải thưởng

Giải Nhất giải thưởng Nhân Tài Đất Việt 2016

Gương khởi nghiệp Thủ đô tiêu biểu năm 2017

Bằng khen của Bộ trưởng Bộ Y tế năm 2020

Bằng khen của Bộ trưởng Bộ Thông tin & Truyền thông năm 2020

Bằng khen của Bộ trưởng Bộ Khoa học & Công nghệ năm 2020

Bằng khen của Thủ tướng Chính phủ năm 2021