

ESG data aggregation efficiency improvement through automation

Client

Europe-based ESG rating agency and proxy advisor.

Business Situation

- SGA supported the client with ESG-related data aggregation including business assessment, screening, and controversy research.
- The client requested SGA to deploy a technology team on the engagement to explore opportunities for auto-extraction of ESG data.
- The client mandated comprehensive implementation of the identified automated solution across all services.

Benefits and Outcomes of Our Engagement

- SGA was able to deliver 20% efficiency gains for the client within three months of commissioning the project.
- Auto-extraction of data was implemented with 100% precision for structured data sources, 80% precision for semi-structured data sources, and 60% precision for unstructured and textual data sources.
- The turnaround time of the data collection was improved by 50%–60% at an aggregated level.
- The accuracy of the data was enhanced by 5% for quantitative data points.

SGA Approach

- SGA deployed a team of two developers, an RPA expert, a business analyst, and an ESG research expert to explore data extraction opportunities and propose solutions.
- The team segregated sources of all the available KPIs/data points and developed three phases for the automation project –
 - ✓ Phase I: Structured data
 - ✓ Phase II: Semi-structured data
 - ✓ Phase III: Unstructured and textual data
- SGA team short listed open source automation tools to achieve the stated automation objectives.
- Key words were listed and data dictionaries were developed to enable data mapping and autoextraction.
- A Proof of Concept (PoC) for the project was developed to automate data extraction for a list of 100 issuers.
- A comprehensive solution architecture and a project plan to deploy it across the coverage universe of the client was designed.
- The project was completed in the three phases stated above and the solution was integrated with the client's technology infrastructure.
- We, further designed, an exception management process for the data that could not be completely automated.