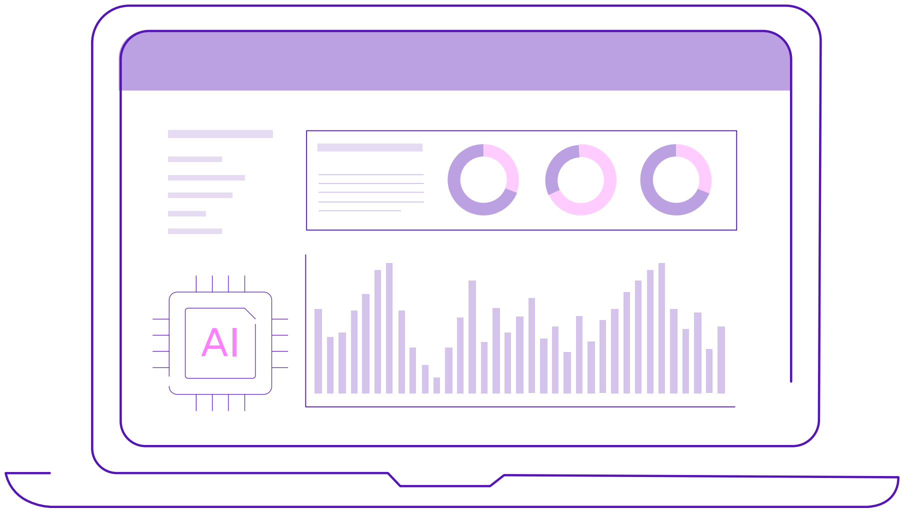
Sustainability app for banking and financial services

1. Introduction

Currently energy optimisation is often based on trial and error and the experience of the site or data centre manager. But what if you could create a dynamic AI (Artificial Intelligence) model that compares activity levels and the energy used and then starts to model the best scenarios and outcomes? The AI starts to work like a sports coach giving you recommendations to help you achieve your “sweet spot”. And what if the algorithm was aligned to existing standards like the Energy Star Rating used in the USA or the EPC rating used in the UK? But more importantly what if it was also dynamic? For example, a new central heating boiler in the UK could be A rated, but in two years, without regular maintenance, it may have dropped to a C. But if the rating was constantly reviewed every few hours or every day that would be interesting.

With our partners we’ve started to deploy algorithms that can typically identify energy savings of 5-15% within a very short period, often just a few months. Firstly, we need to identify the variables that impact energy consumption. For example, the way processes are assigned to various CPUs in data centre servers, or the configuration of lighting and HVAC in an office. Then we can start to plot an EEI (Energy Efficiency Index) to look at past activity and compare different scenarios to see which of these were the most energy efficient and importantly why. We can also compare sites and equipment to look for best practice. We now have a baseline to work to. We could do the same thing with a CEI (Carbon Efficiency Index).

Once you’ve a trusted prescriptive algorithm, by changing the parameters you can quickly create different models. For example, in an office there might be a focus on the efficiency of the HVAC, boilers, and lighting depending upon occupancy levels. Whereas for a data centre we’d be interested in which cores are busy, idle, and unallocated, and the ability to control how fast or slow a group of cores on a CPU are running to offer the end user fine grained control over where they direct their power.



2. How the sustainability app works

The app models energy efficiency (EEI score), i.e. how well you have done, this is the turquoise line on the graph above. Below this, the purple line shows how much money was left on the table on the bad days. It then shows the EEI score for each activity. For each activity it's possible to drill down and see how it could be improved and what the individual saving would have been in money and carbon.

The models are prescriptive and provide continual guidance on how to save energy and carbon. Once you trust the model, which is based on a repeatable algorithm, it can be run as a closed loop automatically fine-tuning energy usage. The data centre version has been co-developed with Intel and our AI partner QIO technologies. It looks at how data centre workloads can be adjusted to enable the determinism that's required from latency sensitive workloads whilst running a reduced power envelope

The sort of things the model could exploit are:

- P-states, or performance states – here the end user can change the frequencies the cores run at to match the demand, i.e. load on the CPU is high, run cores at their highest frequency and conversely if demand is low, reduce their frequency.
- C-states, or CPU power states – here the user can decide which applications to send to sleep, the time of day and how long for. The deeper the CPU sleeps the longer it takes to come out of it, but the energy efficiency gains are more significant. C-states are ideal when leveraged in a time of day scenario. For example, when there is a clearly defined traffic pattern over a 24-hour period.

Using the Intel telemetry insights allows us to:

- slow down / power down servers that may be idle or un-allocated
- adjust internal cooling, in-rack cooling and room cooling
- change workload placement / and consolidate nodes
- place workloads with critical performance SLAs on higher frequency cores
- adjust the sleep state of cores depending on how applications are run throughout the day / week etc.

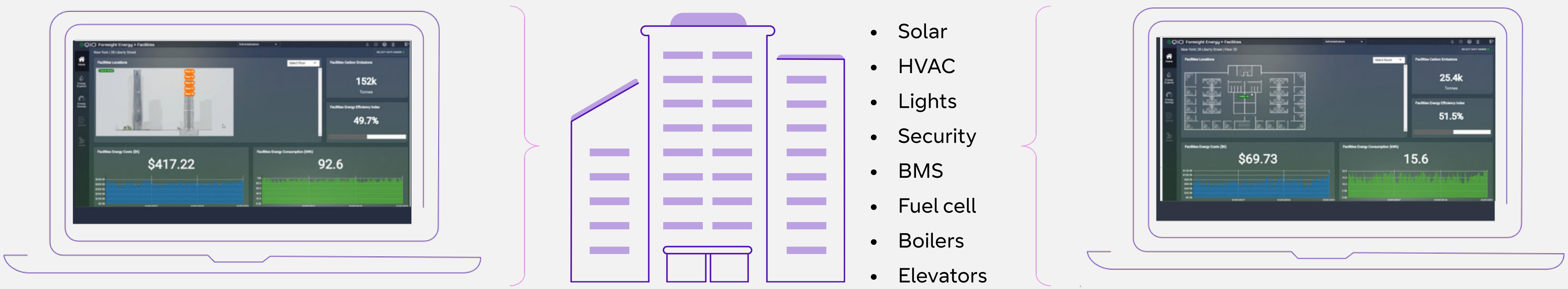
The benefits include the ability to:

- reduce power consumption leading to OpEx savings and a reduced carbon footprint
- have visibility into power consumption and carbon credits per server, node and even application
- recognise stranded servers / resources, high power and cooling scenarios, and available headroom
- identify more efficient placement of workloads to optimise power usage

Case study: Large financial company with 5m sq. ft. of commercial office space.

Problem definition: How can I optimise energy consumption hourly across all buildings?

Solution and benefits: Ingested data from multiple sources, aggregated and accurately clustered data to create predictive Energy Star rating (Energy Efficiency Index) per building based on occupancy, weather, traffic and data from HVAC systems. Modeled predictive savings of 5-8% per building.



- identify high power applications and low draw applications, with less intensive workloads assigned ‘economy class’ cores
- enable compute & network intensive workloads to achieve higher performance.

The app can also be used to optimise energy usage in office buildings (as seen in example below). Currently the buildings version of the app is being trailed in several UK hospitals.

3. Initial results in other sectors

This model has been replicated in four scenarios: marine vessels, heavy industry, construction and buildings. With the data centre version currently being tested by BT.

Other versions of the model include

- marine – certified with Lloyds Registers, 10% saving in marine diesel, or approximately $100K per vessel per annum
- heavy industry - global glassware company – 8% saving per furnace, whilst one of the largest steel works is saving $2.8M and 41,600 tonnes of carbon per annum
- NHS – Queen Elizabeth Hospital Kings Lynn – already a highly efficient site with its own energy generation 5% saving.

Finally, to demonstrate how rapidly the app can be deployed, we were recently asked to look at cement kilns for a global construction company. The client had been experimenting with data science for over two years with no success in moving from research labs to production, or in getting buy-in from the operational teams.

The challenge:

1. Large volume of data – one kiln creates 350m tags a year, at 30 second intervals, with over 300 different tag formats to contend with
2. Complexity – multiple components, e.g. pre-heater, calcifier, kiln, plus the sensors kept breaking hence lots of mis-readings
3. High degree of variability:

- energy efficiency was impacted by quality of coal
- production quality was impacted by lime mix
- plus we were trying to get consistency across three different shifts.

The outcome of the proof of value:

1. Complex datasets (actual data) from the first site and systems was ingested and the quality metrics, efficiency indexes and associated values identified in five to six weeks
2. We modelled the value per site to justify further investment and in year returns, with a blueprint created to ensure repeatability for other sites
3. Savings identified were $420,000 per annum per kiln or $840,000 per site, so the return on investment was over 8 times the outlay for proof of value.

December 2021