Artificial Intelligence and Augmented Intelligence for Automated Investigations for Scientific Discovery

Internet of Food Things

AI^3 Science Discovery Network+ & Internet of Food Things Network+: AI for Allergen Detection and Smart Cleaning within Food Production Workshop
17/10/2019
Society for Chemical Industry

Michelle Pauli
Michelle Pauli Ltd

09/01/2020

AI3SD-Event-Series:Report-14

AI^3 Science Discovery Network+ & Internet of Food Things Network+: AI for Allergen Detection and Smart Cleaning within Food Production Workshop
AI3SD-Event-Series:Report-14
09/01/2020
DOI: 10.5258/SOTON/P0019
Published by University of Southampton

Network: Artificial Intelligence and Augmented Intelligence for Automated Investigations for Scientific Discovery
This Network+ is EPSRC Funded under Grant No: EP/S000356/1
Principal Investigator: *Professor Jeremy Frey*
Co-Investigator: *Professor Mahesan Niranjan*
Network+ Coordinator: *Dr Samantha Kanza*

Network: Internet of Food Things
This Network+ is EPSRC Funded under Grant No: EP/R045127/1
Principal Investigator: *Professor Simon Pearson*
Co-Investigator: *Professor Jeremy Frey*
Network+ Coordinator: *Mr Steve Brewer*

Contents

1 Event Details **1**

2 Event Summary and Format **1**

3 Event Background **1**

4 Introduction **2**

5 Current Innovation **2**
- 5.1 Factory-Cleaning Robots . . . 2
- 5.2 Low-cost, High-Speed e-coli Sensors . . . 3
- 5.3 AI and Machine Learning . . . 4

6 Areas of Focus **5**
- 6.1 Ethics . . . 5
- 6.2 Data . . . 5
- 6.3 Robotic cleaning . . . 6
- 6.4 Hardware . . . 7

7 Conclusions / Next Steps **7**

1 Event Details

Title	AI for Allergen Detection and Smart Cleaning within Food Production
Organisers	AI^3 Science Discovery Network+ & Internet of Food Things Network+ & University of Nottingham
Dates	17/10/2019
Programme	Programme
No. Participants	30
Location	Society for Chemical Industry
Organisation Committee	Dr Samantha Kanza – AI^3 Science Discovery Network+, Steve Brewer – Internet of Food Things Network+, Dr Nicholas Watson – University of Nottingham
Working Group Leads	Professor Jeremy Frey – AI^3 Science Discovery Network+, Dr Nicholas Watson, Dr Joel Fishcer, Dr Peter Craigon & Roberto Santos – University of Nottingham

2 Event Summary and Format

This event was run by AI^{3}SD (Artificial Intelligence and Augmented Intelligence for Automated Investigations for Scientific Discovery) and the IoFT (Internet of Food things) Networks. It was organised in conjunction with the Food, Water and Waste Research Group run by Dr Nicholas Watson from the University of Nottingham. This workshop was centred around the usage of Artificial Intelligence in Allergen Detection and Smart Cleaning within Food Production; research areas that co-align between both AI^{3}SD & IoFT. The programme was made up of several presentations that were designed to report on the current state of affairs, and consider where we need to be going in the future. There were five main working group topics identified for this workshop, and talks were be given on the different aspects that needed to be considered with respect to allergen detection and smart cleaning. Following this, the event broke up into the working groups for more formal discussions. There were multiple sessions for the working group discussions, ensuring that there were opportunities to take part in as many group discussions as the attendees wish. The workshop will be formally recorded and the suggestions for going forward will be captured in a position paper.There was plenty of time for networking, as there was both a lunch and drinks session included as part of the day.

3 Event Background

As food allergies and intolerances are on the rise, allergen detection and awareness is becoming more critical than ever at all stages of the food production pipeline; from cleaning the factories and kitchens the food is produced in, to detecting allergens in food, right through to creating allergen free food in the future. Unsurprisingly research has turned to technological solutions to combat this issue. This research area fits perfectly into the overlapping research interests of AI^{3}SDand IoFT and was organised to explore these areas with members from both networks. This event forms part of the AI^{3}SDEvent Series, which aims to bring people together around important areas of using Artificial Intelligence for Scientific Discovery.

4 Introduction

With food allergies on the rise[1], allergen detection and awareness are more critical than ever. They run through all stages of the food production pipeline, from cleaning the factories where food is processed to providing consumers with detailed information about what is in their food, before they eat it. Can technology help? Could artificial intelligence (AI), Internet of Things (IoT) sensors and other industry 4.0 technologies make a difference?

There is clearly a role for AI in such areas as detecting potential allergens, handling the consequences of allergen exposure, making better use of spectra and keeping track of data. IoT sensors are already playing their part in a number of food chain applications. However, these innovations also come with challenges, many of which revolve around data: getting the data, joining up the data, working with the data, getting information out and ensuring it makes a difference.

This one day Internet of Food Things network+ (IoFT) and AI3SD+ workshop brought together researchers, industry representatives, regulators and policymakers to consider the food chain from producer to consumer – and all the processes in between that ensure we know what is in our food. The day began with a look at pioneering work in three areas – the use of robots for factory cleaning, sensors for e-coli detection and the wider potential of AI in this field. Breakout groups then considered four particular areas of focus: data, robots, ethics and hardware.

5 Current Innovation

5.1 Factory-Cleaning Robots

Figure 1: Dr Nicholas Watson

Can robots clean factories more efficiently than humans? That's the challenge Dr Nicholas Watson's group at University of Nottingham has been investigating in RoboClean, a multi-disciplinary project exploring IoT-enhanced factory cleaning.

The need is clear: cleaning is essential but it is also expensive and uses large amounts of water and energy – 30% of energy in dairy processing and 35% of water in beer production are

[1]The number of food safety events relating to all allergens has more than doubled between 2014 and 2017 (Food Standards Agency, 2018).

spent on cleaning. Downtime for cleaning is increasingly affecting productivity: more frequent cleaning is needed with the rise of smaller, more diverse food varieties, such as low-gluten bread and low-salt bread, where once there might have been a long run of a single loaf variety.

Can low-cost sensors make the process more efficient? The goal is that Self-Optimising Clean-in-Place (SOCIP) will use sensors to detect when cleaning is complete, providing more information than is currently available and avoiding the unnecessary use of detergents and water if the target is clean after 10 minutes rather than the 40 minutes that the previous 'blind' system was set to run for.

"We may not be able to achieve precision in the factory in the way we can in the lab but we want actionable information rather than 100%," said Nicholas.

Furthermore, the RoboClean team is exploring the potential for robotic cleaners to work alongside humans to detect the presence of allergens on factory floors and so avoid cross-contamination. The robots use low-cost infrared sensors to detect allergens but the focus is primarily on the data – how to collect, store, use and share it.

Nicholas and his team have been working on identifying different types of foodstuffs – such as flours, nuts and spices – by taking spectral scans, using principal component analysis to reduce the data down, limiting what goes into the machine learning model and making visualisation easier. The goal is that robots can identify the materials and raise an alert if they detect something that should not be in there. Introducing robots into factories raises interesting issues of human-robot collaboration that require a multi-disciplinary approach if the introduction is to be successful.

"We cannot bring people out of the factory to teach them how to talk to robots and so we have to use the language they use and understand how they communicate, which involves fieldwork and ethnographic visits," commented Nicholas.

5.2 Low-cost, High-Speed e-coli Sensors

Figure 2: Dr Martin Peacock

Low-cost sensors are also at the heart of Zimmer and Peacock's business model. The company developed sensors for the medical space and then moved into the food sector: "Food industry needs low-cost sensors, and glucose strips are the best low-cost meters," declared Dr Martin Peacock, co-founder of Zimmer and Peacock (Z+P). He detailed how Z+P set out to move fast and create a minimum viable product (MVP) with Food Sense, a rapid tester comprising a strip (screen-printed electrode), meter and smartphone app that talks to the meter.

The first product arose from research Martin saw in the media around a test for the hotness of Tabasco sauce. Z+P licenced and took it to market as a product that can measure the heat in chilli, the 'gingerness' of ginger and the pungency of garlic. Next up is a test for e-coli that takes 30 minutes as opposed to a three-day wait for the traditional, gold standard analytical lab test.

While the speed of these kinds of tests is a breakthrough, a challenge for allergen detection is the nature of the food being tested. While it works well for reasonably homogenous materials, such as liquids, it will be much less effective for heterogenous products. How to test for a sesame seed in a piece of bread? Martin was also keen to stress that business models are a far greater challenge than technology.

5.3 AI and Machine Learning

Figure 3: Professor Jeremy Frey

Professor Jeremy Frey, Principal Investigator of AI^{3}SD and co-investigator of IoFT, gave a whistle-stop tour of the history of AI, machine learning and the AI3SD network's preferred concept, "augmented intelligence", which brings together the strengths of both artificial and human intelligence.

Jeremy raised the particular issue of correlation v causation in food allergen scenarios: ethically we cannot intervene in some experiments in random ways, such as exposing people to allergens in a randomised trial, and so we are left with observations alone. One way around this is by using 'do' operators – what happens if I 'do' B, which provides me with more information – and is a way of tying models and data. We do need a model to unravel cause and effect. AI and machine learning can help but we must always bear in mind 'rubbish in, rubbish out'. Or, even, 'good in, rubbish model'.

"If we can join up more of the data, we can build good models on top of it. That is the ideal," said Jeremy.

6 Areas of Focus

Four working groups were tasked with exploring areas of focus – Ethics, Data, Robots and Hardware – and reported back.

6.1 Ethics

Impacts of allergen control: A highly restrictive diet that does not include known allergens narrows nutritional content and, if widespread, would increase intolerances (due to changes in gut bacteria). The cleaner we make our environments, the more we are changing our guts and tolerance. It was noted that children who grow up on dairy farms have very few allergies. We need safe exposure, but also safe foods. If we get to a point of cutting out all allergen-based foods and this has a medical impact, safe exposure needs to be considered.

Bias: there is a high potential for bias in many ways – from what data is collected, which countries are participating in research and which allergens and groups are studied. Allergen research is focused towards western countries, with richer countries tending to put more funding into it and biasing the research towards specific diets, specific groups of allergens and specific ethnic groups. There is also a strong bias towards people who have mild rather than severe food allergies. Those who have severe food allergies are most in need of fuller information but run the risk of being excluded from research due to complexity and the risks of things going wrong.

Objectivity of AI: there are ethical benefits to objective systems but people may not like these systems if they think they work against them. If a food producer gets allergens really wrong they risk killing someone and facing a corporate manslaughter charge. If a computer gets it wrong, who is liable? The engineer? The software provider? The software tester?

Augmented intelligence: we need to preserve the domain expertise of human intelligence with our nuances, and enhance it with machine-based intelligence. It has been argued that AI is no more than computationally intensive statistics – one cannot be truly intelligent until one can account for one's behaviour.

Transparency: we need transparency and communication about the actual capabilities of systems as we typically hold machines to much higher standards than humans – for example, self-driving vehicles. Greater transparency is needed around AI: it shouldn't be black boxed. If it is trained, how was it trained and how did it learn and what safeguards can we put in place to enable things to be properly transparent and tested? Greater transparency is needed in food outlets. The nutrition information in restaurants is often hidden: restaurants have a vested interest in not displaying it but it makes it harder to know what you are eating. And greater transparency is needed with the data. There is no one place where allergen incidents can be looked at. Collating that information would be an excellent whole project in itself but there are also data ethics considerations around sharing allergen data: it is sensitive information that could be used for harmful purposes but is also necessary.

Uncertainty: how do we deal with uncertainty as we progress? Research is not going into the areas that need it the most – severe food allergens.

6.2 Data

Data collection: data standards, exchange formats and use of ontologies is limited, particularly at an international level. They need to be expanded into food descriptions. A citizen science approach to data collection could have potential. Decisions would need to be made about the sort of data to be collected, how it could be pooled and what the benefit is to the citizen collecting the data. One possibility is that a user tracks their food intake against reactions such as IBS, using machine learning to track occurrences of reaction against different ingredients; provenance tracking could then feed this back into the food chain or factories.

Data storage/access: generally, datasets are often incomplete, even when data sharing is mandated or encouraged; open access data is often chaotic. The vast amounts of data being produced would be used more effectively if there was better structuring and labelling. There is a lack of knowledge about what data is available and who has the power to share it. An open food production chain would enable consumers to see provenance and dig deeper into information about companies and factories that have handled the ingredients – but such progress would require significant improvement in standards and data exchange formats. Implementation of FAIR data standards would make a difference.

Data visualisation: needs vary depending on the person viewing the data, from a simple traffic light system to percentage statistics to the full, raw data. Allergy sufferers, factory workers and policymakers will all have different needs. Designing for usability is important, interfaces need to be tailored to specific needs with a level of adaptability for different kinds of user. Dashboards and possibly voice-activated systems could be very useful in the factory environment.

Food provenance tracking: could allow deeper analysis and insight into the food production chain and allergens, but it is not a simple technical challenge to address.

6.3 Robotic cleaning

Context matters: it is crucial to understand the factory. One company may run extensive product lines that constantly produce the same goods while another may have product lines that change frequently and require the space to be reconfigured between line changes. Are wet or dry goods being produced? Some robots may need to be flexible in a dynamic environment – the more complex the factory, the more challenging it is for the robots, which do not understand the concept of space in the same way as humans.

Design matters: hygienic design of robots is an interesting challenge, along with whether the factory line can be reconfigured for robotic cleaning – can cleaning be integrated while the production processing is still running, to help with the cost of investing in robots and reduce waste?

Communication matters: the advantage of using robots is to release human workers to do other, more 'rewarding' tasks (cleaning has a low retention rate). But that needs to be communicated so that people are reassured that they are going to be skilled up rather than replaced. Human-robot communication is also an issue – voice recognition can be a useful tool but may need to be limited to a specific team rather than all factory floor workers.

Risks: trip hazards – how do you get robots to clean around people's feet when they are working? – along with water damage and security. If data is kept locally rather than uploaded to the cloud, there will be a challenge of constant manual back up.

6.4 Hardware

Concerns and responsibilities: nobody needs a sensor sensing something they can't fix, such as mould spores or nanoparticles in products. But if you collect data, whose responsibility is it to do something about it? Who looks after the sensors, cleans and calibrates them? People may become more allergic to different things in the future so how do you ensure the sensor is replaceable?

Testing and placing: some of the challenges in terms of finished product arise if it's not a homogenous product that can be put in an aqueous solution. For example, gluten can be easy to miss if you take a proportion of a product. One of the biggest issues in sampling is whether to send the entire product to a lab, where they will grind it down so it is homogenous and then test the entirety. Where should the sensors be placed? In raw ingredients somewhere in the process, and the final product?

Human error: mislabelling or mispackaging is a low probability but high impact human error. How do we look at a system that can try to reduce that? For example, scanning a 2D bar code on every pack that is reconciled with a central system to ensure that the packer is packing what they think they are packing.

7 Conclusions / Next Steps

The opportunities offered by industry 4.0 technologies, such as AI and IoT-enabled sensors, to help address allergen detection, awareness and exposure in the food chain came across clearly in this workshop. Some applications are already in production, others show great promise. Interesting challenges also emerged from the presentations and discussions. The ways in which humans and robots can interact most effectively was explored, and the ethics around such interaction was a common thread. Indeed, the point was strongly made that ethics should always be treated as an integral part of discussions across all areas rather than as a separate topic.

However, the greatest challenges were raised around data and its collection, storage, accessibility, visualisation and impact. The need to 'join up' datasets is imperative and, for that to happen, there needs to be improved standards and structure.

It was clear from the excellent discussion, which benefited greatly from bringing together people from academia, industry and policy, that these are complex issue that warrant further and continued attention. The Networks will be publishing a position paper in 2020 as the next step in addressing the challenges.