

G. Malcolm Dyson 1902-1978 - the first information scientist?

Martin White FRSC FBCS HonFCLIP martin.white@intranetfocus.com

Managing Director, Intranet Focus 1999-

Visiting Professor, Information School, University of Sheffield, 2003-





George Malcolm Dyson 27 January 1948 Photograph by Bassano Ltd

CC BY-NC-ND 3.0 National Portrait Gallery
George Malcolm Dyson - Person - National Portrait Gallery
(npg.org.uk)

http://www.rscicag.org/index.htm_files/CICAG%20Newsletter%20Winter%202021-22%20FINAL.pdf is a summary profile (pp14-20) focusing on his work in chemical information and contains a complete list of his publications

I hope to publish a full profile in due course but at present there is no obvious channel for doing so.

Martin.White@intranetfocus.com

In the beginning



- In 1958 Jason Farradane had the vision to establish the Institute of Information Scientists (IIS). The founding President from 1958-1961 was Dr. G Malcolm Dyson
- The IIS grew rapidly in size and impact but for a number of reasons both plateaued in 2000-2001 and in 2002 it merged with the Library Association to form the Chartered Institute for Library and Information Professionals (CILIP)
- In the process of amalgamation, the archives of the IIS disappeared. (!)
- From 2019-2021 three former IIS Presidents, Dr Sandra Ward, Professor Charles Oppenheim and myself created a history of the Institute based on secondary sources and the recollections of members. It can be found at <http://iishistory.org/> and will be published by Sheffield University later in 2022 to mark the 20th anniversary of the amalgamation.

Profiling the Presidents



- In the course of preparing the history we researched and wrote new biographical profiles of all the IIS Presidents, all of them very distinguished in their fields of endeavour and all making a significant contribution to the reputation and growth of the IIS
- When it came to G.Malcolm Dyson, the founding President, we found virtually no information on his work, and more importantly on the background of his association with the IIS at the time of its formation
- There is a brief profile in the German Wikipedia
https://de.wikipedia.org/wiki/George_Malcolm_Dyson
- This confirmed that in the early 1960s Dyson worked for Chemical Abstracts Service but sheds no light on his involvement with Farradane and the other founders of the IIS

Chemical information post WW2



- In the immediate aftermath of WW2 there was a very rapid growth in chemical research publications as an outcome of many factors, including the substantial growth in US research capabilities during the war.
- Chemical Abstracts Service was not able to cope with the demand for abstracts and indexes, resulting in long delays in publication and significant frustration by the research community
- The American Chemical Society was not in a position to fund any major investment in staff levels and the process of abstract writing and publication was largely the same as in the 1930s.

Chemical confusion

- In the field of organic chemistry there was a substantial nomenclature problem
- The compound CSCl_4 could be described correctly (according to international conventions) as
 - Perchloromethyl mecapatan
 - Thiocarbonyl tetrachloride
 - Trichloromethyl sulphur chloride
 - Tetrachloromethyl thiol
 - Trichloromethyl sulfenyl chloride
- Furthermore, there were British, French, German and US variants on these conventions
- Building indexes was therefore a nightmare!

The life of G. Malcolm Dyson 1902-1978



- Dyson was born in south London and gained a scholarship to Jesus College Oxford in 1921 to read chemistry
- He gained both a degree from Oxford and at the same time (1925) a degree from the University of London.
- He was awarded a PhD from the University of London in 1927 when he had already published his first research paper and a couple of patents
- In 1928 he published his first book The Chemistry of Chemotherapy
- From 1928 – 1938 he was a lecturer and then head of department at what is now Loughborough University
- From 1939 onwards he was a research director at two UK chemical companies

Chemical nomenclature



- In 1946 Dyson announced his development of a system of chemical nomenclature which would give every organic compound a unique alpha-numeric descriptor
- The first edition of his book *A New Notation and Enumeration System for Organic Compounds* was completed in May 1946 and published by Longmans in 1947
- This work brought him to the attention of Dr James Perry, a distinguished US chemist and librarian
- The second edition was published in 1949 and is notable for a final chapter that suggests how his notation could be used to manage collections of punched card catalogues of chemical information, the use of which had been widely adopted in the USA during WW2 and were in effect proto-typical enterprise search applications

Dyson, Perry and Luhn 1948-1958

- Perry invited Dyson to the USA in early 1948 where he arranged a meeting with Thomas Watson at IBM to see if IBM would support the development of high-speed card sorting
- Watson introduced Dyson to Luhn, and together with Perry they made substantial progress in the development of photo-electric punched card readers
- Two machines were built in 1950 and were demonstrated at the World Chemical Conclave in September 1951
- Dyson, Perry and Luhn continued to work together and this led directly to Dyson's appointment as Director of Research at Chemical Abstracts Service in 1959
- Other notation schemes were published, notably one from Wiswesser which eventually emerged as the notation of choice in the USA and de facto world-wide

Meanwhile, in the UK.....

- The Royal Society Conference on Chemical Information was held in London from 21 June to 2 July with 500 delegates and resulted in a conference report running to over 700 pages
- Dyson was responsible for the demonstrations and visits programmes, including two very sophisticated card readers probably developed for Bletchley Park (that is another story!)
- Farradane was also on the Conference team and it is very likely that as both Farradane and Dyson were very concerned about the lack of professional training for information specialists they had discussions about how to address the issue. Several other founders of the IIS were delegates at the Conference

And at Chemical Abstracts Service...



- Dyson worked as a consultant for CAS from 1956 before taking on the Director of Research role in 1959, reporting to Dale Baker
- Although Dyson's initial vision was for a Registry system the immediate priority was to bring CAS into the computer age
- Working with Luhn Dyson created Chemical Titles as an alerting service that had no print equivalent. This was an enormous leap of faith but paid off almost immediately
- Dyson continued to push research and development at Chemical Abstracts and with Baker was responsible for changing the policy of ACS towards not taking external funding for research. It was becoming clear that CAS needed a significant investment in technology and staff and that could not be funded by ACS member income and CAS sales.
- CAS approached the National Science Foundation and received funding of \$200,000 a year, a very significant amount in 1959/60. Gene Garfield was not amused!!

The importance of Chemical Titles



- The importance of Chemical Titles cannot be over-estimated.
 - It was the first CAS product (probably the first ever) that did not have a print precursor but was electronic from the start
 - It justified the National Science Foundation funding from both an ACS and NSF viewpoint
 - It built an important bridge between CAS and IBM
 - It was the foundation for similar CAS services, notably Chemical-Biological Titles though the Patents version did not work because patents generally had deliberately vague titles
 - It demonstrated to the chemical research industry that CAS was innovative and user-focused
 - It acted as a proof-of-concept for other organisations and publishers, notably Gene Garfield at the Institute for Scientific Information and Science Citation Index

Onwards from CAS



- The momentum generated by Dyson between 1959 and 1962 in terms of research and Tate in terms of building the computer resources continued apace after Dyson's return to the UK.
- The major breakthrough was the adoption of connection tables to build files that could be processed by a computer to both generate and search chemical structures.
- This development was based around work being undertaken by Du Pont which was brought in to CAS and developed further (but in parallel) by Morgan and then Mike Lynch, appointed by Dyson in 1961
- By the mid-1960s linear codes were on their way out though the functional group approach of the Wiswesser Line Notation (WLN) was continuing to be of value.

Rewriting history #1



- In a paper on the history of CAS the only reference to Dyson's contribution is described as
- “G. M. Dyson experimented with creating a small chemical registry file. He selected 14,000 organic fluorine compounds which he coded on to notched McBee cards using the Dyson/IUPAC linear chemical notation system.”
- Evelyn Constance Powell BS in Biochemistry, MEd in Cell Physiology and Science Education, MLS (2000) A History of Chemical Abstracts Service, 1907-1998, Science & Technology Libraries, 18:4, 93-110

Rewriting history #2



- In a paper published in 2001 (Journal of Information Science, 27 (2) 2001, pp. 119–125) Garfield noted
 - “It is significant that during the years 1956 to 1960, both with respect to citation indexing and chemical indexing, I implored Chemical Abstracts to take up ideas based upon these six points that were proposed by me and a group chaired by Max Gordon in the Philadelphia section of the ACS. However, CAS decided to continue on its traditional path.
 - Nevertheless, I maintained nothing but the friendliest relations with E.J. Crane, Charles Bernier, Dale Baker, Fred Tate, Ralph O’Dette, Gerard Platau, Jim Rush, Michael Lynch, Harry Boyle and others at CAS. I must say that I felt a great animosity towards the National Science Foundation (NSF) for preferentially supporting CAS work simply because it was non-profit.”
- There is no mention of Dyson but there is a reference to Michael Lynch who did not arrive at CAS until 1961 and would have worked directly under Dyson’s supervision.

•

Rewriting history #3



- In 1977 Baker and O'Dette published a paper (National Computer Conference 1977) looking back at the history of the development of computer-based applications at CAS, including the KWIC Chemical Titles and the Registry Number but with no acknowledgement of the role of Dyson.
- Baker, on his retirement, suggested that it was Fred Tate that masterminded the computer developments at CAS even though Baker himself had appointed Dyson. (J.Chem.Inf. Comput.Sci 1985 25, 186-191)
- Tate was in fact responsible for introducing computers in the publication process with considerable success, not in the product development and research work

Mike Lynch on Dyson's achievements



- Lynch, who worked with Dyson and Tate from 1961, gives perhaps a more balanced view.
 - “I had the good fortune to work closely with Malcolm Dyson during my apprenticeship years in this area, also with Fred Tate – characters as different as you could wish, each contributing in major ways to information science: Dyson through his vision of what the emergent technology could mean for a chemical information database spanning all of the years, and Tate’s preoccupation being with using the technology to get the show on the road, to continue to provides services which had outgrown printer’s ink and leading in time to a cumulative database.
 - Each was right in his own vision of the future, differing mainly about the means by which they were to be achieved.
 - Dyson and Tate stood above the technology of their time – the fact that they started with an 8K IBM 1401 did not limit their thinking – rather, they were confident that when the need arose the technology would arise to the occasion.”

“The use of the chemical literature”



- At the core of Dyson's interest was the importance of being able to access research and to be able to be certain about the identity and structure of organic compounds. His work on a notational scheme was one element of this interest. His first public presentation of his scheme in 1946 was entitled “Lecture on a new notation for organic chemistry and **its application to library and indexing problems**”.
- The publication in 1951 of Dyson's book “A short guide to the chemical literature” was therefore very timely, and as far as I can determine was the first book on this subject published in the UK.
- There were two earlier books published in the USA. A second edition was published in 1958 and by a strange coincidence both editions were printed in Loughborough, the town where Dyson had lived since 1928
- It is a slim book, just 157 pages in the 2nd edition, but it is written from the wisdom of a practicing chemist and not by an information specialist.

The launch of the IIS 1958



- The inaugural meeting of the IIS was held on 23 January 1958. The meeting was to be chaired by Dyson.
- The association of Dyson with the event was of considerable significance as he was a highly regarded chemist, an innovator in the use of computers in chemical information management
- In 1952 he authored a letter in Journal of Chemical Education entitled 'Preservation and availability of chemical knowledge'. In 1958 any chemist would have recognised that a meeting chaired by Dyson was going to be a significant and popular professional event.
- In the event 125 people turned up at the meeting which for 1958 must have been a substantial majority of the putative information profession. The opening speeches were given by Farradane and Chris Hanson, who was also at the 1948 Conference
- The rest is, as they say, <http://iishistory.org/>

Dyson 1978

- Dyson died on 28 December 1978 as the result of broncopneumonia, congestive cardiac failure and issues arising from hip replacement surgery. Bertha Dyson was at his side.
- In addition to his work on chemical information Dyson actively supported Loughborough University and in the early 1930s developed some effective synthetic routes to antimony-based pharmaceutical products for leishmaniases, a class of infective parasitic diseases which are endemic in around 90 countries, affect mainly poor and marginalized populations
- The Raman effect (molecular vibrations) was discovered in early 1928 and a few months later Dyson suggested that this was the mechanism for smell recognition. It was not until 1999 that it was proved that he was going very roughly in the right direction as the physical nature of smell was finally discovered.

In retrospect



- Without Dyson's intervention CAS would have been overwhelmed by the flood of research papers
- His work on chemical notation led directly to the CAS Registry service
- From 1944 onwards his focus was on access to organic chemistry research information and the importance of training for information professionals
- Without his support the formation of the IIS may well not have generated the visibility and momentum of its formative years

The discovery process – the people



- Robin Darwell-Smith, Archivist, Jesus College, Oxford provided detailed information on Dyson's time at Jesus College from the College Archives. This profile is the first time this information has been in the public domain.
- Evan Hepler-Smith, Andrew W. Mellon Assistant Professor of History at Duke University, Durham, North Carolina, generously forwarded a substantial number of links and comments that he had brought together when he was contemplating preparing a profile of Dyson.
- Chris Walton, Senior Library Support Officer at the University Library, Loughborough University, provided information on the period that Dyson was a member of staff and also on other relationships that Dyson had with the University.
- Jonathan Bushell, Archivist (Modern Records), The Royal Society, London, provided me with the list of delegates to the 1948 conference and the Committee memberships and pointed me to a copy of the proceedings on the Abe Books site that I should have found myself.
- Chris Torrero, Clarivate Analytics, who provided me with a list of Dyson's patents
- David Allen, Librarian, Royal Society of Chemistry for tracking down an obituary of Dyson that was published in Chemistry in Britain, for highlighting the resources of the Historical Collection of the RSC and providing access at the Burlington House Library to a number of books by, and about, Dyson.
- Andrew Dalke provided me with some very useful information about the early history of cheminformatics and in particular about connection tables, and also provided me with copies of two very important articles in Chemistry & Engineering News that I had not been able to download successfully.

The discovery process – the literature

- The literature search turned up over 120 relevant papers, of which 43 were authored by Dyson, excluding 12 research papers and 36 patents
- Luhn's papers are published as IBM Reports and are not contemporaneous with his contributions to CAS, nor are technically comprehensive for reasons of technical confidentiality
- Although the CAS research is published in the Journal of Chemical Information information on the CAS/ACS relationship and ACS meetings is in Chemical and Engineering News, which has not been well indexed by Google
- Although there is an ACS archive at the Science History Institute this does not contain CAS material, and conversations with CAS suggest there is no formal accessible archive – I would love to be proved incorrect in this interpretation