Preprints of the 21st IFAC World Congress (Virtual)
Berlin, Germany, July 12-17, 2020

Ironies of Automation 4.0

Tania Hancke

DLR (German Aerospace Center) Project Management Agency, Bonn, Germany,
email: tania.hancke@dlr.de

Abstract: This paper revisits a truly classic publication: Bainbridge's *Ironies of Automation* (1983) - but it also aims to make the point that the insights gained many years ago are today becoming more important than ever. As we all know, it is due to technological advances that automation is leading to increasingly complex systems which considerably raises the impact of the potential effects. Bainbridge's insights originated from manufacturing processes, but they equally apply to process control in general and to vehicle control, e.g., airplanes, road vehicles or trains. This paper shows that comparable observations can be reported and suggests a human-centered approach to overcome the problems.

Keywords: Cognitive Aspects of Automation, Human-Centered Systems, Knowledge Society

1. INTRODUCTION

We are largely designing automation systems aiming at full automation: Experiences with increasing degrees of automation have been described for many areas of application and the key insights have always been the same. They were first described by the very observant Lisanne Bainbridge as the "Ironies of Automation" (1983) in the context of process and vehicle control.

It is important at this point, to differentiate between automation of *process and vehicle control* versus automation of *mere functions or small-scale processes*, in that the predictability of the former processes tends to be considerably lower. Or to say this with Bainbridge's words: "The more advanced a control system is, so the more crucial may be the contribution of the human operator" (Bainbridge 1983, p. 775).

Here we need to differentiate between *variety* and *incidence*: systems with a high predictability show low variety (situations and states that are possible) and high frequency (occurring very often), while systems with low predictability exhibit high variety and low frequency characteristics (e.g., accidents in process control or aviation typically have a complex background and usually only ever happen once).

Thus today, nearly 40 years after Bainbridge first put her ideas forward, we may find successful automation in automatic windscreen wipers or autonomous shuttles between airport terminals, but this is not the issue here (and they still have some problems occasionally). Let us revisit Bainbridge's key findings. Here we are quoting from Bainbridge's seminal paper (1983). In the subsequent paragraphs, these quotations are printed in *Italics.*

2. BAINBRIDGE'S MAIN INSIGHTS

Bainbridge's three main insights for automated systems relate to operator intervention and monitoring, in particular:

The takeover situation: Manual takeover by the human operator tends to be **difficult** due to the lack of practical experience affecting short-term knowledge.

"If the human operator is not involved in on-line control he will not have detailed knowledge of the current state of the system. One can ask what limitations this places on the possibility for effective manual take-over, whether for stabilization or shut-down of the process, or for fault diagnosis….
The straightforward solution when shut-down is simple and low-cost is to shut down automatically. The problems arise with processes which, because of complexity, cost or other factors (e.g. an aircraft in the air) must be stabilized rather than shut-down. Should this be done manually or automatically?"

Long-term knowledge (experience): It is difficult for the operator to manage the context interpretation for necessary intervention due to the **deterioration** of his/her long-term knowledge.

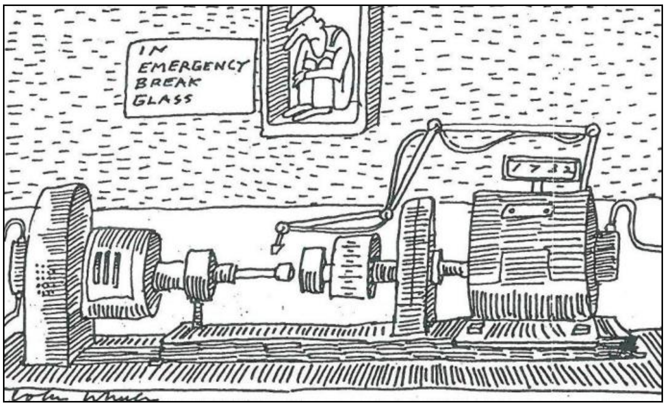
"... it can be important to maintain manual skills. ... simulator practice must be provided…. Unknown faults cannot be simulated, and system behaviour may not be known for faults which can be predicted but have not been experienced. This means that training must be concerned with general strategies rather than specific responses, for example simulations can be used to give experience with low probability events, which may be known to the trainer but not to the trainee. No one can be taught about unknown

properties of the system, but they can be taught to practise solving problems within the known information."

Marginalisation: It has shown to be problematic to **replace** the operator with a superior automated subsystem or function and to leave to the operator merely the task to monitor such a system.

"It does not consider the integration of man and computer, nor how to maintain the effectiveness of the human operator by supporting his skills and motivation. There will always be a substantial human involvement with automated systems, because criteria other than efficiency are involved, e.g. when the cost of automating some modes of operation is not justified by the value of the product, or because the public will not accept high-risk systems with no human component. This suggests that methods of human—computer collaboration need to be more fully developed."

Thus, partial automation did and does not succeed in achieving the claims put forward in its favour.



© Colin Wheeler

In consequence, a *human-centred design* approach has been forwarded, initially for industry or manufacturing processes, to better integrate humans and technology (Cooley, 1989). This design approach aims to keep the operator not only in the loop but in control and to support him/her with appropriate tools and training (Hancke et al., 1989). For an introduction to the Human-Centered Design Approach and its basic principles please refer to the survey paper "A Discourse on AI and Society: Your calculus may be greater than his calculus. But will it pass the Sullenberger Hudson River test?" provided in this same session (O'Neill B, Stapleton L, Gill K S & Brandt D, 2020).

3. AIRCRAFT AUTOMATION

At about the same time, the early 1980s, the promises of automation opportunities reached the aviation industry. Great potentials were seen in automating flight decks of civil aircraft, facilitated by digital signal transmission and processing. Airbus, as the fast advancing European manufacturer, designed the first "fly-by-wire" (FBW) aircraft, the A320, soon to be followed by competitors. High levels of automation were claimed to improve passenger comfort and flying safety. They were correlated with savings in operational costs and, notably, savings in personnel costs as pilots would much less perform tasks of actually flying. Thus they would need lesser qualifications. As expressed in the famous words voiced by the original Fly-By-Wire architect Bernard Ziegler the goal was that

"Monkeys will be able to fly this aircraft!"

Even though these words (and the related video) are remembered well by many of us, there remains no trace of them on the internet today.

Reality turned out to be different: Bainbridge's forecasts and observations held true here and the relevant issues were mostly the same as in manufacturing processes.

Challenges of aircraft automation

Basically replicating Bainbridge's three insights, the following resulting challenges soon transpired:

Situation awareness: pilots should always be "in the loop" and aware of what the various automated subsystems are doing: technology should be assisting them by providing suitable information; one *irony* was that on these new technology aircraft pilots rightly did not trust the systems, so there has been a doubled monitoring load – monitoring the aircraft and monitoring the automated functions (Sarter & Woods, 1995). These findings relate to takeover situations and marginalisation.

Context interpretation: the availability of a variety of system modes (e.g., ways to use the autopilot) provides more flexibility but reduces pilot experience and may cause confusion (Sarter & Woods, 1995). This relates to takeover situations and longterm knowledge..

Maintenance of flying skills: the *irony* here is that while flight deck automation was introduced to widen the field of suitable applicants, it actually narrowed it down; FBW (fly-by-wire technology) ideally requires pilot candidates with even higher levels of skills – pilots who can pick up physical flying skills more quickly than candidates in the past, need less practice and can judge more quickly how automation might affect aircraft behaviour. This relates to longterm knowledge.

Automation of functions: some system functions were automated in such a way that there was no possibility for a pilot override or an over-rule. This relates to takeover situations.

Complacency: though not described by Bainbridge for manufacturing, the belief in and the reliance on automated systems is highly relevant on aircraft and in critical processes (e.g., in nuclear power plants).

Experiences with aircraft automation

A number of civil aircraft accidents were soon following the introduction of the new technology aircraft; especially hull losses of one type of aircraft illustrated these points well. The technology changes introduced by the competitor's model also led to aircraft problems. The best known examples are:

At an air show in ***Mulhouse*** (1988), the "first public demonstration of any civilian fly-by-wire aircraft" ended in a disaster as an experienced training captain crash-landed the civil jet in a forest after flying too low and too slow at the same time; this accident was attributed to a lack of situation awareness, but it may also be partially due to the fact that the captain "had total confidence in the aircraft's computer systems".

The CFIT (controlled flight into terrain) accidents at ***Bangalore*** (1990) and near ***Strasbourg*** (1992), where the aircraft hit ground because they descended quicker than they should have, are attributed to a lack of situation awareness and mode error.

In ***Warsaw*** (1993), a new-technology aircraft overran the end of the runway in extraordinary weather conditions with strong wind shear, because the aircraft software was not sure the aircraft was on the ground and thus, it would not allow the braking systems to be deployed. The pilots could not override the system and therefore, they could not initiate any brakes manually, neither could they know, what the exact problem was, because the way this mechanism worked was unknown to them. At that time, incorrect pilot action was officially given as the cause, but at the same time it was stated that the pilots did not have the relevant information to make a correct decision.

In 1988, a competing aircraft model crash-landed near ***Kegworth*** *(U.K.)*, due to a combination of circumstances. However, bad design of the revised instrument panel in combination with automated and modified functions were key factors for the crash as they decreased situation awareness (Hancke, 1992, 1995).

Implications of aircraft automation

The incidents described above and the related issues raised not only pilots' concerns about the new technologies (Learmount 1991). They were discussed widely and also taken up by IFAC: e.g. at the IFAC World Congress 1993 in Sydney (Hancke & Braune, 1993) and at the IFAC Conference on Integrated Systems Engineering at Baden-Baden in 1994 (Boje & Hancke, 1994). This happened mainly through the activities and concerns of the IFAC Technical Committee Social Impact of Automation. Those discussions did find strong resonance and involvement among the many participants.

As a consequence of these discussions, the goal of *high level flight deck automation* was eventually abandoned at the time and more human-centred approaches to flight deck design were developed over many years (e.g., Billings 1996). Moreover, the rail industry has also recently begun to adopt concepts from aviation for their automation approach in locomotive cabs (Sebok et al., 2015).

While at that time the vigour to automate flight decks was halted and human-centred concepts were developed, they either have not been fully implemented in practice, or are no longer being followed. This is highlighted by fatal experiences in the recent past due to different approaches of technology development chosen in aviation. The two B737 MAX accidents, Oct 2018 in ***Indonesia*** and March 2019 in ***Ethiopia**,* represent sad illustrations of the *ironies* described above:

- An essential automated function had been introduced in the new B737 MAX, which the pilots were not aware of, nor was corresponding relevant documentation available to them – thus they could not understand how the new automation design worked and what was in fact happening concerning the aircraft flight control at a certain flight stage.
- There was clearly the lack of a possibility to overrule the relevant system[1].
- Training and/or experience and crew qualifications seem to have been an issue: both captains were young, their co-pilots had little experience or had shown some deficits in the past.

These accidents reflect all of Bainbridge's findings: taking over control was not possible because firstly the system's actions were not transparent to the pilots and secondly, even if they would have been transparent, physical takeover would not have been possible; longterm knowledge was not considered relevant, relatively unexperienced pilots were in command, while the day before an experienced guest pilot aboard remembered an unofficial fix and thus saved the plane; marginalisation is here represented by the fact that the pilots were supposed to monitor the automated functions, but major changes in their design were not considered relevant enough to be communicated to them.

Resulting in both accidents was a "loss of control", leading to a loss of the aircraft. Since then, all the aircraft of this type have been grounded. They may start flying again in 2020. Nevertheless, the reports refer to serious economic problems of the company itself as well as their partner companies and customers world-wide (Boeing, 2019).

Even before these accidents, only a few years ago, FAA research highlighted several issues that relate to Bainbridge (e.g. Learmount, 2011; Nicas & Wichter, 2019):

[1] *In fact, there was the possibility to stop the system by pulling the relevant circuit breaker, but the knowledge about this unofficial fix goes back many years - it was a common remedy in the earlier years of FBW- and only experienced pilots would have been aware of it; this obviously applied to the pilot who saved the Indonesian plane just one day before it crashed (Levin A & Suhartono H 2019).*

- Between 2001 and 2009, “Inadequate crew knowledge of automated systems” contributed to 40% of accidents and 30% of serious incidents (takeover situation/situation awareness).
- Common handling problems on flight decks include the lack to recognize when autopilot or autothrottle disconnect (takeover situation/situation awareness).
- The pilots’ assessment of both system failures and recovery from failure are difficult, and such failure modes which may *not* occur frequently, have not been anticipated by the designers (longterm knowledge).
- In ca. 60% of 46 accidents, “pilots had trouble manually flying the plane or handling the automated controls” (takeover situation/longterm knowledge).

Another study concluded that professional pilots under-estimate the degradation of their basic instrument-flying skills (Learmount 2011). Pilot training has also been critically highlighted recently as pilot students have generally been encouraged to rely on the flightpath management system rather than to practice hands-on flying. Research by the FAA suggests “that pilots concentrate on programming the automation system at the expense of monitoring the flight path”. The FAA would rather encourage pilots to “have the manual handling skills and confidence to take control of the aircraft if the automation does not perform as expected” (Nicas & Wichter, 2019). In the best case, this will produce many pilots of the calibre of a Sullenberger (2009) who used his superior knowledge and experience to find the optimal solution in view of a serious catastrophe.

Sullenberger’s “forced water landing” (2009) is a good illustration of Bainbridge’s argument: as an ex-airforce and experienced airline pilot Sullenberger assessed the situation quickly, took over control of the aircraft and made the right decision at the right time. This decision saved many lives but lost the aircraft and in the aftermath was challenged by various parties and also in court hearings. A detailed recon-struction eventually proved that Sullenberger’s decision was superior to all other options.

This case of the Hudson River landing particularly highlights the importance of Bainbridge’s *longterm knowledge* for the ability to *assess the situation* correctly and *take over control* with confidence. This knowledge derives from hands-on experience, resulting in what Cooley calls *tacit knowledge* (e.g., O’Neill B, Stapleton L, Gill K S & Brandt D, 2020) and what pilots classically refer to as “*flying by the seat of your pants*”.[2] It is called *tacit* knowledge because it is difficult to verbalize – imagine writing an instruction for someone on how to ride a bicycle.

[2] “ *'Fly by the seat of your pants' is parlance from the early days of aviation. Aircraft initially had few navigation aids and flying was accomplished by means of the pilot's judgment.” From: https://www.phrases.org.uk/meanings/fly-by-the-seat-of-your-pants.html*

4. AUTOMOBILE AUTOMATION

One specific technological development to be seen in this context is the *automation of driving functions* as in cars and trucks. The lessons learned from manufacturing and from commercial aviation imply that designing partial automation always requires special care for an adequate design of the human-machine interface. *Situation awareness* and *keeping the driver (always) in the loop* are essential to retain the skills required to be able to respond to new situations at short notice. Naturally this is of high relevance for controlling vehicles on the road. Let us, thus, consider automobile automation today.

It is generally assumed that automated driving functions will reduce complexity and thus facilitate fully autonomous vehicle travel. This is only partly true. Traffic can come in many forms: from familiar situations recurring repeatedly (like busses shuttling on short pre-defined distances, e.g., connecting airport terminals) to unique occurrences (such as, a white ***kangaroo*** escaping from a zoo and appearing on a nearby motorway)[3]. Following the differentiation previously used in this paper, traffic situations as a whole might also be categorized regarding the frequency of their occurrences (frequency) and the likelihood of their occurrence (variety). The above examples would then relate to either *low variety/high frequency* or to *high variety/low frequency* situations. The levels of driving automation as defined by the SAE J3016 norm, ranging from driver assistance (L1) to full automation (L5) are differently suited for these different contexts.

Highly or fully automated driving (L4 and L5) can be easily achieved today in ***low variety/high frequency*** traffic situations, as is amply illustrated by many autonomous vehicles on short-distance, dedicated tracks (also including rails here) with a fixed route and mostly fixed time-table, for example shuttles between aircraft terminals. With technology proceeding, the scope of such applications can even be extended, though within limits.

For ***high variety/low frequency*** situations this is different, and especially so for motorized individual traffic. Traffic situations are already complex today, at current levels of technology, because of all the different traffic participants (e.g., pedestrians, joggers, parents with prams, motorized wheelchairs, senior citizens with walking frames, bicycles, pedelecs, motorcycles, public transport in various forms, delivery vehicles of all sorts etc. and not to forget priority traffic such as ambulances), the variety of roads (including country lanes as well as 4-lane motorways) and the possible speeds of all those different participants.

[3] *This refers to the escape of a white kangaroo from a private zoo south of Essen in summer 2019 (Kangaroo 2019). The kangaroo did not actually appear on a motorway, but considering the density of motorways in the area, this is a likely scenario. The exemplary issue is, whether the AI systems currently used for visual recognition for vehicles would be able to recognize a kangaroo on a motorway in Germany, especially a white one.*

In such contexts, it will be extremely important to keep the driver in the vehicle as an *alert driver* rather than as a *passenger/part-time driver*, as fully autonomous driving (L5) is highly likely to be limited to controllable areas of application. It should also be added here, that there are no figures on how many accidents are currently avoided, because a human driver reacts appropriately and timely to an unforeseen situation and thus "saves the day".

In addition, it is assumed that there will be an *interim* phase of *mixed traffic* for some years. Fleet penetration of connected automated cars will increase, and all approaches will need to take into account a considerable number of vehicles that are neither connected nor automated. The flaw in this assumption is that this phase of mixed traffic is unlikely to ever end. On the one hand, there will always be some drivers who refuse to use automated functions or who drive classic cars; and there will be pedestrians who do not want their mobile phones to be used for exchanging information with approaching cars. On the other hand, there will be a continuous process of technological improvements so that each new vehicle model will have more functions to take care of. Thus the overall fleet mix is likely to expand rather than decrease, as will the mix of other traffic participants (e.g., e-scooter as a recent example or new small-scale solutions for urban logistics in coming years).

Furthermore, the required infrastructure for automated driving is not likely to be uniform. Firstly, it will not be available all over the country. Where it is available (or will be), the systems will, secondly, need to be upgraded continuously which is a true challenge for any society. Therefore a common denominator will have to be found to regulate how this can be dealt with. Possibly cars will be developed with a priority on automated car-bound systems in order to be able to maximize benefits. This then affects the question of cost-ownership of the infrastructure installed.

The essence here is: Yes, the areas of application for autonomous driving of cars, busses, trucks and other vehicles can be extended. But this starts from the low variety/high frequency end of the scale and may slowly conquer new territory in steps. It follows from the discussion of the *ironies of automation* that *fully or even highly automated driving* for all motorized individual traffic is realistically not possible. There are many potential application cases for autonomous and automated driving, but they may not primarily be at this end of the scale. This is illustrated well by the "white kangaroo" effect and the potential "hand-over" situation it would create. Due to the *de-skilling* or *un-learning* of drivers who are no longer truly integrated into the control loop of the vehicle, these drivers are not likely to (reliably) respond appropriately and especially quickly to such a challenge. Another restriction of this vision is the effort/benefit ratio: How much effort will achieve which benefits, especially in a mixed traffic scenario? How will road infrastructure have to change and who will bear the costs? How will city layouts have to change?

5. BAINBRIDGE'S VISIBILITY TODAY

Since their original publication, Lisanne Bainbridge's observations have become even more relevant today, in all areas of application where automation has been or is being introduced. Beyond the areas which have been mentioned so far in this paper, this even includes areas such as financial trading and cloud computing (Baxter et al., 2012). Bainbridge is today one of the most visible researchers of our recent past world-wide. This view has been taken up very recently by Strauch (2018) in the following way:

"Lisanne Bainbridge's 1983 paper, Ironies of Automation, has had considerable influence on human–machine research, prescience in predicting automation-related concerns that have led to incidents and accidents, and relevance to issues that are manifested to this day. ... Bainbridge described how automation fundamentally altered the role of the human operator in system performance. Requiring the operator to oversee an automated system that could function more accurately and more reliably than he or she could, can affect system performance in the event that operator intervention is needed. The influence of the insights Bainbridge provided on the effects of automation on system performance could be seen in both research on automation and in the recognition of ironies discussed in subsequent automation-related accidents. Its inspiration to researchers, accident investigators, regulators, and managers continues to this day as automation development and its implementation continue unabated."

And further on, by the same author:

"The paper's influence can be seen in a variety of ways. At a fairly broad level, the number of works that have referenced the paper is substantial. As of early November 2016, Google Scholar listed 1800 scholarly works that had cited Ironies of Automation. By contrast, other influential works on the subject, such as Weiner and Curry's 1980 Ergonomics paper on flight-deck automation (Wiener & Curry, 1980) listed 564, … The number of citations of Bainbridge's work, large as it is, is also increasing at a considerable rate. In the two-week period from late October to early November 2016, ten additional published and presented works cited the paper… Not only does Bainbridge's paper continue to be cited in scholarly works, its influence on our understanding of the field has been substantial as well."

6. CONCLUSION

Bainbridge's insights need to be acknowledged in the design of any automated system or process, but especially so when aiming at automation in high variety/low frequency contexts. The design of automated control processes should not marginalize the human operator, pilot or driver but allow him/her to always be in or take control at any given time; to do so confidently, the design should allow for the build-up of sufficient hands-on knowledge and situation awareness through involvement. To achieve such system characteristics, a human-centered design approach is particularly suited.

REFERENCES

(All web references have been accessed in 2019)

Bainbridge, L (1983). Ironies of Automation, *Automatica,* 19, 6, 775-779.

Bangalore (1990): https://aviation-safety.net/database/record.php?id=19900214-2.

Baxter, G., Rooksby, J., Wang, Y., Khajeh-Hosseini, A. (2012). The ironies of automation … still going strong at 30? Proceedings of ECCE 2012 Conference, Edinburgh, 65 – 71.

Billings CE (1996). Human-Centered Aviation Automation: Principles and Guidelines. NASA Technical Memorandum 110381.

Boeing (2019). https://en.wikipedia.org/wiki/Boeing_737_MAX

Boje E., Hancke T. (1994). Summary of panel discussion: Automation in Commercial Aviation. IFAC Integrated Systems Engineering, Baden-Baden, Germany.

Cooley M. (1990). Human-centred systems, in: Rosenbrock H (ed.): *Designing human-centred technology.*

Ethiopia (2019). https://aviation-safety.net/database/record.php?id=20190310-0.,

Hancke, T. (1992). Hamilton's Principle and the Design of Human-Machine Systems, PhD Thesis. Imperial College, Univerity of London.

Hancke, T. (1995). *The Design of Human-Machine-Systems: The Example of Flight Decks*. ARMT 13, Aachen, Germany.

Hancke T., Besant C.B., Ristic M., Husband, T.M. (1989). Human-centred technology, IFAC Symposium on Skill-based Automation, Vienna, Austria.

Hancke T., Braune R. (1993). Human-centered design of human-machine systems and examples from air transport, IFAC World Congress, Sydney, Australia, 1993.

Indonesia (2018). https://aviation-safety.net/database/record.php?id=20181029-0.,

Kangaroo (2019). https://www.watson.de/wissen/tiere/943183311-feuerwehr-jagt-ein-ausgebuextes-weisses-kaenguru-in-velbert

Kegworth (1988). https://aviation-safety.net/database/record.php?id=19890108-0.

Learmount D. (1991). The Meeting Place. Flight International, Nov 27, 1991.

Learmount D. (2011). Industry sounds warnings on airline pilot skills, Flight International, 6 February 2011.

Levin A & Suhartono H (2019). https://www.bloomberg.com/news/articles/2019-03-19/how-an-extra-man-in-cockpit-saved-a-737-max-that-later-crashed

Mulhouse (1988). https://en.wikipedia.org/wiki/Air_France_Flight_296 https://aviation-safety.net/database/record.php?id=19880626-0.

Nicas J., Wichter Z (2019). As pilots rely on more automation, their skills and confidence erode. New York Times, New York print edition from 15 March 2019, Section A, 13.

O'Neill B, Stapleton L, Gill K S & Brandt D (2020). A Discourse on AI and Society: Your calculus may be greater than his calculus. But will it pass the Sullenberger Hudson River test? Paper accepted for presentation at IFAC World Congress 2020.

Rosenbrock, H.H. (1990). *Machines with a purpose.* Oxford University Press, Oxford, U.K.

SAE (Society of Automotive Engineers) J3016. Levels of driving automation, e.g. https://www.sae.org/news/2019/01/sae-updates-j3016-automated-driving-graphic

Sarter, N., Woods, D.D. (1995). How in the world did we ever get into that mode? Mode error and awareness in supervisory control. *Human Factors,* 37, 1, 5 – 19.

Sebok, A., Wickens, C., Laux, L., Jones M. (2015). Supporting human-automation interaction in the rail industry by applying lessons from aviation, *Proceedings of the Human Factors and Ergonomics Society*, 59th Annual Meeting 2015, 1661 -1665.

Strasbourg (1992). https://aviation-safety.net/database/record.php?id=19920120-0,

Strauch, L. (2018). Ironies of Automation: still unresolved after all these years. *IEEE Transactions on Human-Machine Systems,* 48, 5, 419-433.

Sullenberger, C. (2009). https://en.wikipedia.org/wiki/US_Airways_Flight_1549,

Warsaw (1993). https://en.wikipedia.org/wiki/Lufthansa_Flight_2904 https://aviation-safety.net/database/record.php?id=19930914-2&lang=de.,

Wheeler, C (1991). Copyright for cartoon granted for this publication.

Wiener, E.L., Curry, R.E. (1980). Flight-Deck Automation – Promises and Problems. https://www.semanticscholar.org/paper/FLIGHT-DECK-AUTOMATION-PROMISES-AND-PROBLEMS-Wiener-Curry/3f14be8f6e6df17b05ed74b6f2737af79fb7420e