

## The Riddle of the Robots<sup>1</sup>

**Jon Bing**

Professor, dr juris

Bøgata 7, NO-0655 OSLO – Norway

Phone: +47-22-675400, fax: +47-22-677478, mobile: +47-90-967659

E-mail: jon.bing@bing.no; jon.bing@jus.uio.no



### The riddle

In 1993, I made a journey to Chicago, an excursion with some colleagues and graduate students to visit the Chicago-Kent School of Law and other institutions related with my own research area, computers and law.



During the stay, we visited the Art Institute of Chicago. And there I made a remarkable discovery. In its collection I discovered a small three-dimensional figure made of cardboard.<sup>2</sup> The head is reminiscent of a cathode ray tube; the body is made up of boxes decorated with numbers, the feet are two tubes. Anybody would interpret this figure as a robot, and this is indeed the name of the artwork. The artist is Alexandra Alexandrovna Exter, and she made this small figure in 1925.

It was the date that puzzled me.

“Robot” is a rather recent word. It was used the first time by the Czech author Karel Čapek (1890-1938). Čapek was an important and prolific author, famous for his many novels and plays. Many have read *War with the Newts* (1936), which is a very funny and very serious science fiction novel ridiculing the emergence of the Nazi movement, and indicating Čapek’s strong political involvement.



His play *R.U.R.*, which is an abbreviation for “Rossum’s Universal Robots”, is also of a political nature. In this play, the main character is Rossum, an industrialist who creates artificial beings from biological material in order to have slaves in his production plants. There is a description of the process of building these beings:

*‘Spinning mills for weaving nerves and veins. Miles and miles of digestive tubes ... in the fitting-shed, all the parts are put together like motorcars. They learn to speak, write and count. They have astonishing memories. But they never think of anything new. Then they are sorted out and distributed. Fifteen thousand daily, not counting a regular percentage of defective specimens which are thrown into the stamping-mill ...’*

It is maintained that the idea of such a play came to Čapek rather suddenly, and that he immediately discussed his idea with his brother Josef, who was a cubist painter.

*“But,” the author said, “I don’t know what to call these artificial workers. I could call them Labori, but that strikes me as a bit bookish.” “Then call them Robots,” the painter muttered, brush in mouth, and went on painting. And that’s how it was. Thus was the word Robot born; let this acknowledge its true creator.<sup>3</sup>*

The word ‘robot’ is derived from the Czech word ‘robota’, which means drudgery or ‘servitude’, and ‘robotnik’, who is a ‘servant’ or a ‘serf’.

<sup>1</sup> A written version of an invited talk to ‘IRIS – Internationales Rechtsinformatik Symposium’ at the Universität Salzburg, introducing the parallel session “Science Fiction and AI”. There have been published versions of this in Norwegian, for instance “Dobbetgjengere: Om maskinmennesker og menneskemaskinger i fantastisk litteratur”, Jon Bing *Landskap med tegn*, Pax, Oslo 1998:89-106. This was also published in Yulex 2004. I want to thank Associate Professor Dr Erich Schweighofer University of Vienna, Faculty of Law, Research Center for Computers and Law for giving me the opportunity to develop this English version.

<sup>2</sup> The photograph is taken by one of the students, Viggo Elster, and has not been authorised by the Art Institute.

<sup>3</sup> <http://capek.misto.cz/english/robot.html>.

The play *R.U.R.* opened in Prague early 1921 and was extremely successful. It was produced in New York in 1922, and the English translation of the play was published in 1923.



Perhaps this is sufficient to explain what puzzled me with the figure of Alexandra Exter. For in Čapek's play, the robots are artificial human beings – they look exactly like humans, they would be what in modern science fiction jargon is known as an 'android' or a 'biomat', a machinelike man. But 'robot' is reserved for something different, for a mechanical man or a manlike machine, an artificial being built of hardware.

The word was created in 1921. Today it has a different meaning from what it originally intended. Somewhere along the line, the meaning changed from 'machinelike man' to 'manlike machine'. And that point in time would seem to be located between 1921 and 1925 – because the figure of Exter is obviously that of a 'manlike machine', a robot in the modern sense of the word.

I admit it is a little mystery. But my own preoccupation with science fiction,<sup>4</sup> artificial intelligence and robots made this an intriguing one and this essay is an explanation of the mystery.

### Prague, Čapek and the Golem

It cannot be a coincidence that Čapek wrote his play about the artificial slave workers in the city of Prague. The famous Jewish ghetto of Prague is symbolised by the Golem.

A Golem is an artificial being, created by Man from earth and clay. It is said to be mentioned in Psalm 139:15-16:

'My bones were not hidden from you,  
When I was being made in secret,  
Fashioned as in the depths of the earth;  
Your eyes foresaw my actions;  
In your book all are written down;  
My days were shaped, before one came to be.'

The 'depths of the earth' is a metaphor for the womb, emphasising the mysterious operations occurring there. And legend tells that at the end of the 16<sup>th</sup> century, rabbi Judah Löw followed the secret formula and fashioned in secret an artificial being from clay dug out of the banks of the river, and gave him life by placing a "Schem" – a capsule containing piece of paper with a cabbalistic word – in the mouth of the giant.

Rabbi Judah Löw, called The High Rabbi Löw because he was unusually tall, is a real person (1520-1609), his full name was Jehuda Liva ben Becalel. He was an acknowledged theological scholar, and shared his time between Prague and Krakow in Poland. In Prague, he founded the Talmud school Klause, part of the ghetto. His grave can be seen in the cemetery, and still today one may find leaves weighted down by pebbles on his gravestone, he is believed to be a powerful rabbi whose spirit can be called upon for intercession.

The ghetto was the part of town in which the Jews were permitted to live. The living space was cramped, the lodgings small, the ghetto crowded. But it was generally believed to be very rich, as the Jews were successful business people. And there was considerable tension between the Jews and the rest of the population in the city – it was rumoured, for instance, that Jews ate babies in their un-Christian rituals, especially at Easter. The Golem was created by rabbi Löw as a defender of the ghetto, and over time the Golem became the symbol of the spirit of the Ghetto, like it appears in the well known novel *The Golem* by Gustav Meyrink (1868-1932). Meyrink was another Czech author living in Prague, and his novel was published in 1915, a few years predating Čapek's *R.U.R.* Still today clay figurines and other trinkets representing the Golem are popular souvenirs sold in the Ghetto. And the tale of the Golem still continues to fascinate writers, one of the more recent demonstrations is Marge Piercy *He, She, and It* (1991).<sup>5</sup>

<sup>4</sup> I have with Tor Åge Bringsværd edited a science fiction anthology with stories about robots, *Jeg – en maskin*, Gyldendal, Oslo 1973.)

<sup>5</sup> Published in Europe as *Body of Glass*.

In the original legend of the Golem, the artificial giant was somewhat simpleminded, taking all orders literally – in one episode, he does like the sorcerer’s apprentice, when asked to fetch water, he continues to do so also even when the bowl overflows, and the house is flooded, the housekeeper having forgotten to ask him to stop.

Rumours of the giant reached the emperor Rudolph II of the Hapsburgs. His court must at this time have been a remarkable place, the emperor collected scholars and artists around him, and among them was another fascinating figure of that time, the astronomer Johannes Kepler. One may speculate on what form a dialogue between him and rabbi Löw would have taken, but we do not know whether they actually met.

It is said, however, that the court protocols of 22 February 1592 indicates that rabbi Löw was called to an audience before the emperor, with prince Bertier also present. The protocols do not tell what was said at this meeting, but leaving, rabbi Löw stated, ‘We do not need Jossele Golem anymore.’

With this two friends and relatives Jakob Katz and Jakob Sosson, rabbi Löw takes the Golem to the chamber in the tower of the Altneu synagogue, which still stands at the edge of the ghetto. Here he reads a secret cabbalist formula over the Golem, reducing him once again to clay and dust.

It has been argued, even proved, that the tale of the Golem is much older than rabbi Löw, and that there are better candidates for his part in the story. ‘Golem’ shall also have been used in Jiddish for a simpleminded person, nearly synonymous with a village idiot. This is of little concern in our context – the living tradition attributed the creation of the man of clay to rabbi Löw, and he became an obedient slave labourer.



As the novel of Meyrink demonstrates, the legend was very much alive at the time of Čapek. One of Čapek’s acquaintances was the journalist Egon Erwin Kisch (1885-1948), who at that time worked in the German language newspaper *Bohemia* in Prague, writing local stories. Kisch would go on to become a famous reporter, earning himself the nick-name ‘Der rasende Reporter’ or ‘The furious Reporter’, especially due to his reports from the Spanish Civil War and other war theatres. His interest for the alleys and tales of Prague led him to write ‘*Dem Golem auf der Spur*’,<sup>6</sup> where he traced the legend of rabbi Löw and his Golem. He entered the tower chamber of the Altneu synagogue and looked for the clay or dust remaining of the

Golem, but without success.

Therefore, when Čapek conceived his idea of the industrialist Rossum and his artificial slave labourers, the Golem provided an obvious reference for his robots. But rather than creating them from clay and mystic rituals, Čapek let them be manufactured in vats similar to the chemicals provided for the factories of his own and modern age.

## Frankenstein, his monster and the summer of 1816

However, this was not the first (nor the last, as we have seen) time that the tale of the Golem inspired the creation of an artificial being which since then has haunted our literature.

In 1816, the British poet Lord Byron (1788-1824) fled England due to rumours of his less than appropriate behaviour with respect to his half-sister. He fled in style, in a large horse-drawn carriage containing a library and a sleeping room, and he brought with him a social companion, a young man by the name of John Polidori (1795-1821), his father an Italian translating Gothic horror novels, himself a brilliant graduate from Edinburgh medical school. This background may not be wholly irrelevant, the medical school of Edinburgh was famous for its expertise in human anatomy, and expertise gained by dissecting corpses – and there were in the 18<sup>th</sup> century a brisk trade in corpses, graves in the city being robbed and its content sold to the University. This was sufficiently common for iron cages to be put on the graves to lock out the grave robbers.

Lord Byron’s travels took him in the summer to Lac Leeman, where he rented a house, Villa Deodati (9 Chemin de Ruth, the house is still there) just outside the city of Geneva.

At the same time, another famous British poet also escaped to the Continent, Percy Bysshe Shelly (1792-1827). He had two years earlier eloped to the continent with the 17 year old Mary Godwin (1797-1851),

<sup>6</sup> Egon Erwin Kisch *Der rasende Reporter*, Aufbau Taschenbuch Verlag, Berlin 2001 (3rd edition). The original collection of essays was originally published 1925.

the daughter of the philosopher William Godwin and the suffragette Mary Wollenstonecraft. Conveniently Shelley inherited a modest income and his wife, Harriet, committed suicide by drowning herself in the Serpentine, Hyde Park, leaving the two lovers free to marry and once more travel to the Continent.

Neither of the two spoke much French, but they brought with them Mary's step-sister, Claire Clairmont (1798-1879). She was an actress, and it is believed she had an affair with Lord Byron while she was performing at Dury Lane Theatre. She was still infatuated with him, and therefore it probably was not by chance that they came to rent Maison Chapuis, very close to Villa Deodati.

It is reported that it was a wet summer, robbing the two groups of the most obvious way of passing the time, making excursions into the splendid landscape. They had visitors, one of them was Matthew Lewis (1775-1818), one of the best known Gothic authors, and his most famous book is *The Monk* (1796). He read aloud to them from the first volume of Göthe's *Faust* (1808) – Lewis translated the text from German as he read. The group also read aloud from other books, one of these was *Phantasmorgia*, a collection of Jewish folk tales which also included the legend of the Golem.

When reading about this summer and the persons staying at the shore of Lac Lehman, one is given the impression of a company of rather bored persons willing to experiment with intoxicating substances like belladonna in order to pass the time. And Lord Byron suggested that they should have a literary contest, each writing a Gothic tale. His own attempt was a less than successful attempt to write about a vampire, but which is recognised as the first literary vampire tale. It was just a fragment, Polidori came to publish his own story "The Vampyre" (1819) in such a way that it was attributed to Lord Byron, it is still disputed whether this was by design or accident.



The winner of the contest was, of course, Mary Wollenstonecraft Shelley, who published it as *Frankenstein – A Modern Prometheus* (1818). The story has become part of our shared literary mythology, repeated and retold in numerous movies; the many trivial versions may overshadow the importance of the novel. One of the major themes of the novel is the moral responsibility of man – as God has moral responsibility for his creation, man, in the same way man has a moral responsibility for his own creation, in Frankenstein's case, the monster. The moral relations between a creator and the created is a theme that still is current, biotechnology making it perhaps more acute than in the time of Mary Shelley.

In the novel, the natural scientist Victor Frankenstein creates an artificial being by stitching together parts of corpses, and inducing life in these by sending the electrical power of a bolt of lightning into its nervous system. The Italian anatomist and physician Luigi Galvani (1737-1798) was one of the first to investigate experimentally the phenomenon of what came to be named 'bioelectrogenesis'. In a series of experiments started around 1780, Galvani, working at the University of Bologna, found that the electric current delivered by a Leyden jar or a rotating static electricity generator would cause the contraction of the muscles in the leg of a frog and many other animals, either by applying the charge to the muscle or to the nerve. Percy Bysshe Shelley was extremely fascinated by modern science, and this interest was shared by Mary, who in this way gives life to an artificial being using the methods of current science.

Frankenstein's monster is created using electricity; in Čapek's play, the robots are created on some sort of assembly line in a chemical and industrial facility. The major difference is, perhaps, not the method of creation, but the fact that Frankenstein's monster is unique, while Rossum's robots are numerous, indicating a change of perspective from the moral responsibility of the individual to the political responsibility for a social order allowing the exploitation of the masses. In both cases, the artificial being is of flesh and blood, we have still to find the solution to the riddle of the robot.

### Excursion: Ada Augusta, Lady of Lovelace and the Analytical Machine of Charles Babbage



A note should be permitted on another curious relation between the summer of 1816 and modern information technology. While staying in Villa Deodati, Byron worked on the third Canto of

“Childe Herold’s Pilgrimage”, which the Shelley’s and Clare Clairmont brought back with them to England while Byron continued his fateful journey.

This canto was dedicated to his infant daughter, Ada Augusta (1815-1852), who was to become Lady Lovelace. Ada Augusta had a talent for mathematics, and become the collaborator of the grumpy genius Charles Babbage (1791-1871), Professor at the Lucasian Chair of Mathematics at Cambridge (and famous for never giving a lecture to the students). He designed the Difference Engine, an advanced calculator, and the Analytical Engine – the latter in principle a general, mechanical computer using stacks of cogged wheels as a short-time memory and being programmed by the same punched cards used for Jaquard’s looms.

The best description of this machine is the notes of Lady Lovelace added to a translation of an essay published by the Italian engineer and mathematician Luigi Menabrea (1809-1896) in 1842 – Menabrea himself not being a minor figure, he became in 1867 Italian premier and foreign minister. It is through Lady Lovelace’s notes we today retain the best discussion of the Analytical Machine, its construction, and the vision of how it could be programmed to solve any problem. Her insight has earned her the title of the world’s first programmer.

In this way one may trace a thin thread of relation between the modern computer and the origin of Frankenstein’s monster.

Perhaps of more relevance in our context, is one of the questions Lady Lovelace addressed – would the Analytical Machine be able to “think” in the sense that it would be able to create something novel. The question may be disturbingly modern, and is one also today haunting the research in artificial intelligence. Her answer, being known as Lady Lovelace’s response, is that as the machine only could carry out the processes it had been programmed to do; it would not be able to do anything novel. This has later been challenged, perhaps most notably by Alan Turing (1912-1954) a century later, who emphasised that the machine could be programmed to learn from experience through feedback, and in this way might improve its own programs beyond the foresight of the programmer, and therefore arrive at results or behave in a manner not foreseen.

We may want to compare Lady Lovelace’s response to the description of Rossum’s robots, which is said to not being able to “think of anything new”. I have no reason to believe that Čapek was familiar with Lady Lovelace’s notes, at his time, the work of Charles Babbage was regarded as a curiosity of no practical importance; Babbage’s restoration only came after the construction of the first electronic computer, which in principle paralleled his Analytical Machine, and this realisation also brought fame to Lady Lovelace.

### **Alexandra Alexandrovna Exter and *Aelita***



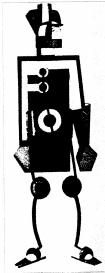
Alexandra Alexandrovna Exter (originally Grigorovich, 1882-1949) was born in Bielostock, and attended art school in Kiev. In 1916, at a time when non-figurative art was still rare, Exter created her first totally abstract paintings. In the same year, she began designing sets and costumes for a Moscow play. Her revolutionary designs won critical acclaim, and her theatrical career was launched.

In 1921 she joined the Soviet constructionists, artists who put their talents to use for the new communist state. Most abandoned easel painting, even in its most radical forms, as overly bourgeois, and turned instead to design. The general slogan of these ‘constructivists’ was ‘Art into Life’, and their goal was ‘to unite purely artistic forms with utilitarian intentions’. In their most extreme formulations, the constructivists announced, ‘Art is finished! It has no place in the human labour apparatus. Labour, technology, organisation ... that is the ideology of our time.’

For the next several decades Exter produced innovative and influential stage designs for plays, ballets, and experimental films. However, like many radical artists whose work did not fit in with Soviet ideology, Exter eventually left the country, settling permanently in Paris in 1924. There she remained an important influence through her exhibitions, her stage work, and her teaching. Exter continued experimenting and sometimes incorporated modern industrial materials such as celluloid and sheet metal into her futuristic designs.

Just before she immigrated to France, she worked with the Soviet film director Yakov Protazanov (1881-1945) in the production of *Aelita, Queen of Mars*. Protazanov was called “the old man” in Soviet film because he had started his career before the revolution, and he commanded considerable respect. The movie was – for its day – a major production which started in February 1923 and took over a year to complete. There was shot of

22,000 meters of film, which resulted in 2,841 of the final movie. The movie opened 30 September 1924 at the Ars Cinema in Moscow, and a special musical score (by Valentin Kruchinin) had been composed to accompany the silent scenes.

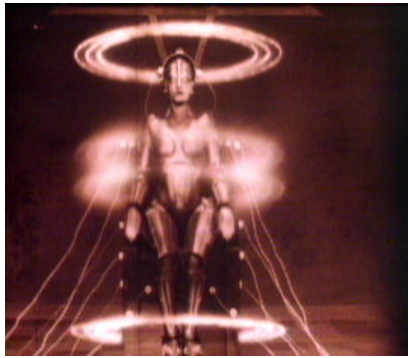


The story is based on a novel by Alexei Tolstói, a distant relative of the famous author Leo Tolstói. He had actually gone into exile, but had voluntarily returned to the Soviet Union to join the team making the film. Very little of his novel was used for the film, and Tolstói was deeply disappointed.

The plot was rather simple. The story starts in December 1921 in the chaos of Civil War in the Soviet Union, and the start of the New Economic Policy. An engineer working among the starving masses of Moscow designs a spacecraft with a friend, who is a soldier. Together they go to Mars, and find there a society divided into a ruling class and a working class. The engineer falls madly in love with the Queen of Mars, Aelita. But she is oppressing the workers, and as a good Bolshevik, the soldier starts an uprising among the slave workers of the planet.

At an international exhibition of decorative arts 1925 in Paris, Protazanov was given an award for *Aelita*, and at this time Alexandra Exter already had settled there.

Her designs for the movie were remarkable and cubistic. The sketch reproduced shows the costume for one of the oppressed workers. It is angular, making the associations to a machine not far-fetched.<sup>7</sup>



*Aelita* became an international success and was also exported to the United States. In English, it had the sub-title “Revolt of the Robots”. The play by Čapek was current at the time *Aelita* was released, and the theme of *R.U.R.* was similar to that of the film, the oppression of workers by the upper classes. In the sub-title, “robots” retained its original meaning of “machinelike men”, or slave workers used without respect for their individual rights as human beings, but rather like tools in an industrial process.

We also realise that my interpretation of the title of Alexandra Exter’s little figure in the Art Institute of Chicago was not correct. It was made in 1925, a year or so after the design of the costumes to the slave workers of Mars. It does not render a manlike machine, a “robot” in modern terminology, but rather a machinelike man, a “robot” in the sense of Čapek and *Aelita*. Its association with a machine is a political comment on the exploitation of workers like they were parts in a huge, industrial process.

We therefore still have to find how the word changed its meaning.

### Fritz Lang and Metropolis: The riddle explained

One of the finest movies ever made is Fritz Lang’s *Metropolis* and it is even included in the UNESCO list of the Memory of the World. It was produced in Germany 1925-26 by the Austrian film director Fritz Lang (1890-1976), who escaped to the United States before the Second World War and continued to make memorable movies.

The movie is based on the novel *Metropolis* (1927) written by Lang’s wife, Thea von Harbou (1888-1954). The production of the film was itself a major undertaking, new cinematographic techniques were developed, in some of the scenes, as much as 36,000 extras participated. It was produced by Universum Film AG (Ufa) in Berlin, and nearly brought the company to bankruptcy.

The film is to some extent inspired by *Aelita*, and the story has many parallels both with *Aelita* and *R.U.R.*

<sup>7</sup> I am grateful to the Arts Institute of Chicago to have provided material on Alexandria Exter and her designs for the film.

The story is set 100 years in the future (2026), in a mighty city like a Gothic version of Manhattan. There are the privileged few and the oppressed workers. The Master of Metropolis is John Masterman. His son, Freder, happens to meet a girl, Maria, in the subterranean city of the workers, where they struggle among the machines, ruthlessly exploited by the privileged. Maria encourages the workers to rise against the tyrants. The Master of Metropolis becomes concerned both with his son's infatuation with Maria, and Maria's radical agitation.

So far in the plot, the parallels to the former versions of the uprising of the workers are rather clear. But a new element is now brought into the story, the evil scientist Rotwang – a figure which still haunts science fiction. Of course, he is a genius and his right hand is replaced by prosthesis and covered in a black glove.<sup>8</sup> Rotwang is asked to make an artificial copy of Maria, which is to take her place in front of the workers and render them harmless, and at the same time provide a solution for the lovesick son.

The plan, of course, fails. But in our context, the creation of the artificial Maria, played by a magnificent Birgitte Helm (1906–1996), is important. In the laboratory of Rotwang, the real Maria is placed in an apparatus, and by her side is placed a copy of her in metal. Using electricity, the life force of the human is transferred to the machine. The laboratory does echo the operating theatre of Dr Frankenstein, where he harnesses the electrical power of lightning to bring his monster to life.

The artificial Maria cannot be mistaken for an android; she is built of metal and circuitry. So in her we encounter a “robot” in the modern meaning of the word. But the word “robot” is not used to describe the artificial Maria, neither in the film nor the novel.

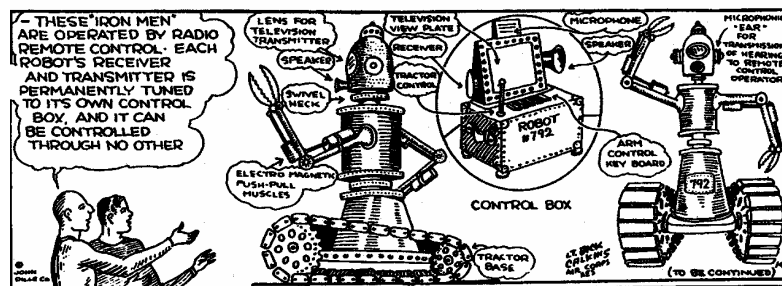
Though *Aelita* was made a couple of years prior to *Metropolis*, they were distributed to the American market about the same time. The themes of the two science fiction movies were rather similar, and both movies were spectacular by the standards of the time – though *Aelita* could not rival *Metropolis* in the force of the allegory told.

In *Aelita*, the subtitle was ‘Revolt of the robots’, and the workers were dressed in cubistic costumes associating them with mechanical beings. In the other, a humanoid clearly built of metal leads the uprising of the workers.

It seems to me that this is the explanation of how the word “robot” subtly changed its meaning, from “machinelike man” to “manlike machine”. In the wake of the popularity of the two movies, the riddle of the robots is explained.

## Robots entering popular culture: Buck Rogers in the 25<sup>th</sup> Century

In 1928, the science fiction magazine *Amazing Stories* published the novella ‘Armageddon 2418 A.D.’ by Philip Francis Nowlan (1888-1940). The story caught the interest of John Dille Company, who contracted the author to



develop a comic strip based on the same basic idea. The result was *Buck Rogers in the 25<sup>th</sup> Century*, one of the first true space operas to be created, and a forerunner for other well-known romances like that of Flash Gordon (and his companion, the scientist Dr Zarkow).

The first episode, ‘Meeting the Mongols’, is drawn by Dick Calkins and was published in 1929. In this one finds a detailed drawing of something called an ‘iron man’ – it has video cameras for eyes, speaks through a loudspeaker, the arms end in pliers, and it is moving on rolling belts like a tank. It is not autonomous, but rather operated at a distance – and would therefore perhaps not earn to be called a ‘robot’ according to the modern usage of the word. But this is exactly the term used by Nowlan as the iron man bears the brand ‘Robot #792’.

<sup>8</sup> Stanley Kubrick's *Dr Strangelove* (1964) also wears such a glove over his artificial hand as a tribute to Lang.

This is the earliest example I have found of the word definitely having changed its meaning. *Buck Rogers* became very popular as the first comic strip addressed to an adult audience. The strip ran in US newspapers 1929-1967, the last feature film was made as late as 1979, and a television series run for several seasons, ending in 1981.

The first short-story I have found using the term “robot”, is Harl Vincent ‘Rex’s *Amazing Stories* (1934), but there were several stories of intelligent or manlike machines around the same time. The robots became a theme in science fiction and the stories typically recall the original Frankenstein motif, the created intelligent machine turning against its creator.

There are a few anthologies which will give broad samples of the stories of this time: Isaac Asimov, Patricia S Warrick and Martin H Greenberg (eds) *Machines that think* (Holt, Rinehart & Winston, New York 1983) and Sam Moskowitz (ed) *The Coming of the Robots* (Collier, New York 1963).

### The modern robot: Isaac Asimov and the three laws of robotics



The image of the modern robot was forged by the extraordinary prolific and diverse Isaac Asimov (1920-1993), an American science fiction writer with a Russian background. His first short story on robots was ‘Robbie’ (*Super Science Stories* 1940). The same year, Asimov brought a proposition for a new story to one of the most influential science fiction editors ever, John W Campbell of *Astounding Stories*.

*‘My notion was to have a robot refuse to believe he had been created mechanically in a factory, but insist that men were only his servants and that robots were the peak of creation, having been created by some godlike entity. What’s more, he would prove his case by reason, and “Reason” was the title of the story.’<sup>9</sup>*

The story was published in *Astounding Stories* 1941. In our context, it may be important to emphasise that Asimov in the story, still revolves around the Frankenstein motif, the relation between the creator and the product of the creative process. The robot and its two constructors are in an artificial satellite orbiting the Earth, there being no direct evidence of the human society producing the autonomous robot, which is brought to consciousness for the first time aboard the satellite. It argues with considerable force that the two men cannot have created it:

*‘Look at you. I say this in no spirit of contempt, but look at you! The material you are made of is soft and flabby, lacking endurance and strength, depending for energy of the inefficient oxidation of organic material. Periodically you pass into a coma and the least variation in temperature, air pressure, humidity, or radiation intensity impairs your efficiency, You are *makeshift*.*

*‘I, on the other hand, am a finished product. I absorb electrical energy directly and utilize it with an almost one hundred percent efficiency. I am composed of strong metal, am continuously conscious, and can stand extremes of environment easily. These are facts which, with the self-evident proposition that no being can create another being superior to itself, smash your silly hypothesis to nothing.’*

The final addition to the modern robot was the three laws of robotics that has made Asimov famous, but which he himself attributes to Campbell from a conversation 23 December 1940:

#### THE THREE LAWS OF ROBOTICS

1. A robot may not injure a human being, or, through inaction, allow a human being to come to harm.
2. A robot must obey the orders given it by human beings except when such orders would conflict with the First Law.
3. A robot must protect its own existence as long as such protection does not conflict with First or Second Laws.

<sup>9</sup> Isaac Asimov *In Memory Yet Green: The Autobiography of Isaac Asimov* (Doubleday, New York 1979:281).



On the basis of these laws, Asimov created a series of stories. The short stories are collected in *I, Robot* (1950) and *The Rest of the Robots* (1964). The novels include *Caves of Steel* (1954) and *The Naked Sun* (1957). Towards the end of his career, Asimov started a project of intertwining his robot tales into the original *Foundation* trilogy, and it is difficult to sort out their relations. Now has also been added the movie *I, Robot* (2004) directed by Alex Proyas (1963-), and based on the characters and plots from the robot stories.

### Artificial Intelligence, replicants, cyborgs and other family members

It would be futile trying to follow in detail the image of the intelligent machine in fiction after Asimov introduced the modern robot. There are numerous and examples. One will find that in science fiction, robots lived strangely apart from proper computers. When the computers were introduced, the image was of a very large machine, typically a computer that governed society. The miniaturisation and distribution of computers that took place throughout the late 1970s, had not been foreseen in the literature.

A typical image of the intelligent computer is presented in the legendary movie by Stanley Kubrick (1928-1999) *2001: A Space Odyssey* (1968) based on a script by Arthur C Clarke (1917-), a British science fiction author with a crisp, nearly academic and matter-of-fact style. The computer is governing the space vessel journeying to Jupiter, and is called HAL (an obvious pun on IBM). The force of the image of this computer is partly its soft and emotional voice, and the way it develops some sort of paranoia, a mental disorder making it even more human. It may be maintained that in this respect, the Frankenstein motif again emerges – at least, the dialogue between man and machine is part of a discussion of what essentially makes a human being different from an intelligent machine.



This becomes one of the major themes in the work of the American science fiction author Philip K Dick (1928-1982). In his many books and stories, he has really only one major question, ‘What is reality?’ His own tongue-in-cheek answer is, ‘Reality is what refuses to go away when I stop believing in it.’ In all his books, he finds a new way of discussing the relation between reality and images of reality.

An important version of his basic question is how to see the difference between a human and an imitation of a human. The novel in which this question perhaps is most clearly in focus, is *Do Androids Dream of Electric Sheep?* (1968). “Androids” are, of course, the name given to an artificial being not constructed from metal, but rather from organic material. In the novel, androids are prohibited from visiting Earth, and have a very limited life-span. Rouge androids frequently trespass, and a corps of law enforcement officers hunts them down, administers tests to make sure they have caught an android and not a human being, and then ‘retire’ them. The protagonist of the novel is such an officer, and he becomes deeply troubled when he finds it difficult to identify androids, especially when he also falls in love with one of them and, finally, starts to believe he himself may be an android.

The novel has become even better known in the movie version, *Blade Runner* (1982), directed by Ridley Scott with Harrison Ford in the role as the protagonist. In the film, the androids are called “replicants”, emphasising the central issue of similarity between the hunter and the hunted.

In a way, this closes the ring. We started out with the robots of Čapek, being produced from organic material in chemical vats, and ends up with the replicants, sophisticated and smart look-alikes to humans. But there is still one more step to go – the fusion or union between man and machine, the cybernetic organism, the cyborg.



In the 1980s, there appeared a strain of science fiction usually called “cyberpunk”. There are several prominent authors that could be counted as belonging to this school, but the exponent is William Gibson (1948-), who in his first and vibrant novel *Neuromancer* (1984) coins the word “cyberspace”, which has been absorbed by everyday language as a reference to the virtual reality of computerised networks and machines. In the work of Gibson and the other cyberpunks, the computer is directly interfaced with the neural network of man. An extra memory chip may be slotted into a socket behind the ear for easy access to a foreign language; a readout from you personal agenda is available in your field of vision, including a face recognition module to prompt you for the name of a person you have met in a cocktail party, outlining his personal information as he told you at that occasion – supplemented with what is available on the Net. And for a high, the pleasure centre of the brain is directly stimulated by an inserted probe.

In this emerging fiction, the distinction which has followed us through this essay – between “machinelike man” and “manlike machine” – is absorbed in a new synthesis, which may be an apt symbol for the post-industrial society. Frankenstein’s monster was born from the emerging understanding of the power inherent in natural sciences. The robots were an allegory to the social and political forces in the industrial society. The cyborg may be the symbol making explicit the new and ethical challenges when biotechnology, nanotechnology and information technology converges into new possibilities both for the individual and society.

### Postscript: The PIN and Zamyatin

In this essay, the robot has been considered as a powerful literary symbol, emphasising the way humans may be used without respecting their individual dignity, as machines in an industrial production system. It has occurred to me that there is one other such symbol, this is the number – it is often used as slogans, ‘Individuals are reduced to numbers in the machinery of bureaucracy,’ or similar statements.

The reason for the number having this symbolic effect has not been quite clear to me. One might think it was related to the introduction of early computerised systems, where storage capacity was scarce, and where numbers were somewhat more easily handled than letters. Some countries have introduced personal identification numbers, a unique number used typically in the communication with government systems to avoid mistakes of identity. Obviously, numbers are also used for similar purposes by countries not having a unique PIN, like the social security number in the United States, or the tax administrative number in Canada.

But clearly the symbolic value assigned to numbers predates government computer systems. Perhaps some of the negative associations are related to the use of numbers tattooed onto the skin of prisoners in German concentration camps during the Second World War. But this only must have augmented a quality of ‘the reduction to a number’ already present. Perhaps also this has a literary explanation – at least there is a powerful Utopian novel from the time of the Soviet revolution which, as far as I know, for the first time introduces numbers as a symbol of the lack of respect for human individuality.



Yevgeny Ivanovich Zamyatin (1884-1937) was a powerful literary figure in the early Soviet state. He welcomed the revolution, he criticised its repression of freedom. He was close to the leaders of the revolution, but was arrested several times, and become considered a heretic, constantly attacked in the late 1920s by Communist Party-line critics. He had to give up the leadership of the All-Russian Writers' Union. His works were banned, removed from libraries, and he was unable to publish. After writing a letter to Stalin, Zamyatin was allowed to go with his wife into exile in 1931. He settled in Paris in 1932, where he died in poverty.

He completed his only novel, *We*, in 1921. Extracts from the original text were published in an émigré journal in Prague in 1927. In Russia *We* circulated in manuscript, and was finally published in Paris.

Zamyatin's *We* (original title *My*, 1921) takes place in the One State, where all buildings, tools and machines are made of glass. People are called ‘numbers’ and all live, work and act precisely in unison, to the point that they chew their food together. Their actions are dictated by the Table of Hours, a clock system which synchronises precisely what everyone is to do and when. The people are ruled by the Benefactor and policed by the Guardians. The protagonist is D-503, a mathematician and builder of the Integral, a gigantic glass space-ship which is being constructed to go to other planets and spread the joy of the One State – here one may see the a parallel to *Aelita*.

One will recognise the theme of the novel as common to that of the two most influential utopian novels of the last century, *Brave New World* (1932) by Aldous Huxley (1894-1963) and *1984* (1949) by George Orwell (1903-1950), as well as closely related to that of the revolt of the robots in *R.U.R.*, *Aelita* and *Metropolis*. It indicates that the web of literary symbols and allegories are much more complex than we have been allowed to glimpse through the small aperture opened by Alexandra Exter’s cardboard figure in the Art Institute of Chicago. We will not pursue these tantalising threads to find new and intertwining patterns. But they remain reminders not only of the richness of the literary imagination, but also of the importance to our understanding of ourselves, and our relation to the society contained in these symbols and allegories.

I am pleased and satisfied to have solved, at least to my own satisfaction, the riddle of the word “robot”.

Copyright of *Journal of International Commercial Law & Technology* is the property of International Association of IT Lawyers and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.