

## WHITE PAPER

## Utility Bills Made Easy

## 5 Ways Machine Learning Will Disrupt Utility Bill Management and How Energy Products are Sold

## Executive Summary

Getting utility bill data is hard. Basic frustrations persist, including inconsistent online data access and the abundance of paper bills, which are nearly half of those sent! WattzOn has built an advanced machine learning (ML) software system to make utility bill data faster, better and cheaper.

As with other industries and applications, ML is hyper-efficient at certain tasks. This creates a new set of industry economics, in which the costs to extract utility data from scans, images and PDFs is a fraction of those incurred with current industry practice.

The result is a technology advance that dramatically changes the industry and enables new business partnerships across industries.

This white paper makes some fairly audacious claims. But they are easy to validate. Just do a simple online search for "AI will change everything". Hundreds of well-reasoned articles will pop up that make the same economic arguments set out here.

Thus, the central theme is everywhere: ML technology is extremely powerful and disrupts the status quo. WattzOn is pleased to bring this advanced ML system to the challenge of getting clean, structured utility bill data at speed and scale.

**AI or ML?**

It is hard to avoid articles on Artificial Intelligence (AI) these days. But this white paper speaks of Machine Learning (ML). So what is the difference? Here is one explanation from Forbes(1): *AI is the broader concept of machines being able to carry out tasks in way that we would consider "smart". ML is the application of AI based on giving the machines data and letting them learn for themselves.* WattzOn's advanced ML system is both smart and self-learning.

**5 Ways ML Disrupts the Utility Bill Data Market**

- 1 ML is highly accurate. It beats humans
- 2 ML is self-learning. It gets better and better
- 3 ML enables automated intelligence
- 4 ML creates a barrier to entry
- 5 ML expands the market for utility bill data

**Imagine Data Liquidity for Utility Bills**

It has always been expensive and challenging to get utility bill data, so many energy markets have been starved of bill data. For example, can your company get all of its utility bills and data into one clean database? Current practice is only semi-automated, making the costs high, so many companies do without.

To create a sharp image of the future, we borrow a term from healthcare: Data Liquidity. Loosely speaking, Data Liquidity is achieved when data flows securely, quickly and freely from the original source to wherever it is needed. It enables new products and technologies, as well as intelligent choices.

The chokepoint that prevents Data Liquidity in the utility bill market is the fact that nearly half of utility bills are sent out in paper form. This occurs because of aging utility billing systems, and the fact that consumers want to review their bills on paper before paying. Data extraction at scale in this environment requires specialized solutions and human labor. The process is costly and absorbs the attention of key managers.

An ML-based software system solves this problem, and does so at such a low cost that utility data starts to pop up everywhere. With cheaper and better data in hand, energy services and energy products can be sold differently. Data Liquidity for utility bills provides a huge spillover, changing how products are defined and sold in adjacent markets as well.

---

## How Is Utility Bill Data Provided Today?

An entire industry has emerged to help utility customers and authorized third parties access and use utility bill data. While third party services offer lower costs and better data access than DIY, third parties themselves face challenges in providing data from highly varied formats.

Bill or Data Format	Share of Meters	Key Challenge
EDI	30 - 35% of commercial meters. Not in residential	Works well when available. Sometimes PDFs not included
Consolidated Bills	10 - 15%	Current technology (OCR) does not handle. Requires human data entry
PDF format	20 -30%	Requires two software functions: Login to utility site to get PDFs; Extract data from PDFs. Changing utility sites and bill formats cause problems
Paper	40 - 50%	Current practice: Scan bills, run OCR, use 100s of employees to enter or clean up data. Changing bill formats cause problems

---

## 5 Ways Machine Learning Disrupts the Industry

### #1 ML is Highly Accurate

Very simply put, ML uses every crumb of data to increase accuracy, achieving outstanding results. (2) A roundup of recent press shows that ML is beating humans in a wide range of tasks:

- Image recognition: Humans have a 5% error rate; ML from Microsoft and Google has less.(3)
- Understand spoken language: Humans have a 5% error rate; Google matches that; Microsoft has a lower error rate.(4)
- Playing Poker : Yes, AI beats humans at Texas Hold'em.(5)
- Playing Go : Go is a complex game, and Go tests ML's abilities to do automated predictions and optimized decision making.(6) Google's AlphaGo beats the world's best human Go players. (Elon Musk's Open AI computer beat the world's best Dota players, an even more complex game.(7))

WattzOn's advanced ML system brings similar powerful ML algorithms to utility data extraction, and similarly, provides accuracy levels exceeding the industry's current semi-automated processes.

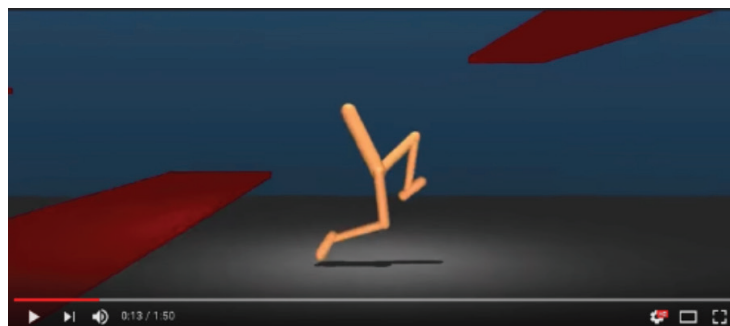
## #2 ML is Self-Learning

Self-learning software is the stuff of alarming headlines, such as "robots taking over the world". Consider this amusing video in which a stick figure teaches itself to walk. Humans set up an obstacle course and penalties for failure. The stick figure started out towards a goal and corrected itself. Self-learning is a fast-moving area of research, with recent important announcements from Google, Intel and academic researchers. (8)

Self-learning is a feature of ML with important implications for utility bill data processing. Once an ML system is set up, self-learning makes adding a new utility easier and easier. The ML system is already tuned and calibrated to many different bill types.

Self-learning accelerates the competitive advantage of an ML system, widening the cost and accuracy difference between the company that deploys ML and its competitors.

### Self-learning AI: The Stick Figure Taught Itself Every Motion



<https://youtu.be/gn4nRCC9TwQ>

## #3 ML Enables Automated Intelligence

In a now famous article, *Why Software is Eating the World*, Marc Andreessen, a leading investor and entrepreneur in Silicon Valley argued that the stock market does not recognize the power of software companies to reshape established industries, nor the ability of agile small software companies to dislodge established players. (9) Written in 2011, the article's arguments still ring true.

The utility bill data processing market is on the cusp of this very transformation. In addition to highly efficient data extraction, layers of business logic and other specialized functions will be automated, collapsing the current value stack into a hyper-efficient "digital energy management system."

### Digital Energy Management Automates Intelligence

CURRENT INDUSTRY	NEW INDUSTRY: DIGITAL ENERGY MANAGEMENT
Provide Value-Add Services (forensic accounting, resolution of bill issues, customized offers)	Provide Value-Add Services
Apply Industry-Specific Business Logic	Automated Intelligence
Apply Bill-Level Business Logic	Automated Intelligence
Clean and Normalize Data	ML System
Extract Data from scans, pdfs and images	ML System
Get Data (EDI, scans, pdfs and images)	Get Data

In today's market, everything below Value-Add Services is a necessary expense on the way to offering profitable Value-Add Services. With ML and its automated intelligence, these functions become highly automated and costs nearly disappear. This enables a profitable service for providing data and automated bill monitoring. The provision of clean, cheap, intelligent data opens up many new opportunities in Value-Add Services and a new ecosystem is launched. The fintech software market has evolved in this way, with some companies very focused on data provision and other companies using the streams of data to create highly automated and customized banking experiences.

## #4 ML Creates a Barrier to Entry

Fast forward three or four years. An industry player has deployed an ML system and Automated Intelligence. How can a new entrant compete?

Not on accuracy, as the ML system will be optimized for utility bill data extraction through self-learning.

Not on costs, as the ML system will have progressed down the learning curve and be hyper efficient.

Not on utility coverage, as the ML system will have been trained so broadly that it adds the new utility at nearly zero cost.

Not on new-found intelligence, because the entrant will not control the data flow. It must buy data at market rates. Those with the data will add the intelligence features.

Of course, the industry will keep changing. But entrants will need to compete on different dimensions. The ML deployment changes the game.

## #5 ML Expands the Market for Utility Bill Data

What happens when utility bills have Data Liquidity, when bill data flows securely, quickly and freely? Here are three key areas to watch.

- **Growth in demand.** This is basic economics: ML enables a significantly lower cost structure, allowing price to drop while profits increase. The lower price expands the market for current services and also enables a profitable business model for providing data to companies that innovate in value-add services.
- **An explosion in customer-centric offers.** Utility bills are hard to understand. Software makes the data more accessible, more intelligent and more personal. Combining intelligent bill data with energy products creates new opportunities to tailor proposals, profitably manage energy services and combine products for the customer. How about a retail energy/solar/storage combo? How about a solar/EV/thermostat combo? Customers are demanding this energy future.(10)
- **Innovative partnerships to acquire customers.** Too often the cost of sales cripples business models in the cleantech industry. Liquid Utility Data opens up partnerships across industry boundaries. Utility bill management services can be a feature of enterprise accounting software. Small business customers can subscribe to bill monitoring and alerts through invoice processing companies, such as Quickbooks. Consumers can be acquired through partnerships with smart home players. These partnerships integrate the value of customized energy offers into large aggregations of potential customers.

Liquid Data for utility bills is a compelling future, with abundant data and applications that lead to customer-centric solutions everywhere. ML-based systems remove the key friction, enabling new business opportunities and innovative partnerships.

---

## When the Long Tail is the Market

Many industries have a large core of high-volume data sources, making data aggregation fairly easy. For example, the top ten credit card issuers process 90% of all credit card transactions, and 85% of U.S. bank payments originate from just 14 banks. (11) (12) But as the following table shows, utility bill data is different. (13) Competitive advantage for utility bill processing requires a solution for the long tail: Amass coverage, add utilities cheaply, keep competitors out. By design, WattzOn's ML system excels in this regard.

Utility Type	No. in US	Concentration/Fragmentation
Water	54,000	Top 400 utilities serve 46% of the market
Waste	27,000	Top 8 companies serve 50% of the market
Electric	2,300	Top 10 utilities serve 30% of the market
Natural Gas	1,300	Top 10 utilities serve 34% of the market



### What's Wrong with OCR?

Until very recently, the best technology for text data extraction from invoices, bills and receipts was Optical Character Recognition (OCR). If your company has spent the time and money to implement OCR, you're probably wondering why there is a need to update to an ML-based system.

Here are three reasons:

- **Cost to maintain an OCR system.** Regardless of the particular OCR technology deployed (full-text with search or zonal OCR), as bill presentations change, OCR breaks. Time must be spent to update the OCR software.
- **Who is doing the work?** Many OCR systems require your team, the customer, to define the zones or fields, and when there is a change in utility bill presentation, your team must find the problem and update the OCR setup. Your data analysts spend time on mundane tasks that ML automates. This is a large, hidden cost to OCR deployments. Fully-automated ML systems remove this expense.
- **Cost to add a new utility.** With its self-learning features, ML lowers the cost of adding a new utility. OCR costs stay the same. Very quickly, the ML-based has a large cost advantage.

WattzOn can incorporate your full-text OCR results, or provide an integrated OCR/ML service. Either way, less human intervention is needed to maintain software and to clean up after software breakage through human data entry and review.

### Conclusions

Here at WattzOn we're excited about our advanced machine learning system, and throughout this white paper we've tried to temper our enthusiasm and message. The economics are starkly clear, but perhaps with softer wording the reader can hear us better.

So, imagine our surprise at reading a recent report from McKinsey about the strategic impact of AI and ML in the insurance industry. (14) Here are some phrases from the Executive Summary: "Resistance to what lies ahead is futile." "Digital technology destroys value... [it] propels some companies, but for many more [it] depletes corporate earnings." "[companies] not in the game could find themselves locked out." Alarm bell wording!

Throughout the white paper we have shown the improvements in utility data extraction enabled by an ML system. But perhaps the key conclusion is at more strategic level: ML leads to harsh industry economics, including winner-take-all situations. But, the upside from an ML system is significant. We agree with the McKinsey analysts who state:

**"[ML delivers] a triple prize: Satisfied customers, lower costs, higher growth."**

### ABOUT US

WattzOn provides utility bill data for digital energy solutions in the consumer and commercial sectors. With national coverage, speedy and secure data acquisition, and an advanced machine learning system to extract line-item data from utility bills, WattzOn has the industry's leading technology platform. Our customers include market leaders in solar, smart home, storage and commercial billing markets.

## Notes

- (1) <https://www.forbes.com/sites/bernardmarr/2016/12/06/what-is-the-difference-between-artificial-intelligence-and-machine-learning/#4a00a5b42742>
- (2) A general description of using ML's predictive power for reading receipts is shown in this blog post by Dropbox: <https://blogs.dropbox.com/tech/2017/04/creating-a-modern-ocr-pipeline-using-computer-vision-and-deep-learning/>. WattzOn has used a different ML configuration, one optimized for the nature of utility bill processing. For a review of the many applications of AI that are working well, see: <https://a16z.com/2017/12/07/summit-ai-update-frank-chen/>
- (3) [https://www.eetimes.com/document.asp?doc\\_id=1325712](https://www.eetimes.com/document.asp?doc_id=1325712); <https://arxiv.org/pdf/1502.01852.pdf>
- (4) <https://www.recode.net/2017/5/31/15720118/google-understand-language-speech-equivalent-humans-code-conference-mary-meeker> <http://www.businessinsider.com/microsoft-beats-humans-at-speech-recognition-2016-10>
- (5) <http://www.sciencemag.org/news/2017/03/artificial-intelligence-goes-deep-beat-humans-poker>
- (6) <https://techcrunch.com/2017/05/23/googles-alphago-ai-beats-the-worlds-best-human-go-player/>
- (7) <https://www.engadget.com/2017/08/12/ai-beats-top-dota-2-players/>
- (8) <https://thenextweb.com/artificial-intelligence/2017/10/16/googles-ai-can-create-better-machine-learning-code-than-the-researchers-who-made-it/>  
<https://newsroom.intel.com/editorials/intels-new-self-learning-chip-promises-accelerate-artificial-intelligence/>  
<http://www.iflscience.com/technology/computer-scientists-generate-self-aware-mario-can-learn-and-feel/>
- (9) <https://a16z.com/2016/08/20/why-software-is-eating-the-world/>
- (10) <https://www.enel.com/media/news/d/2014/03/encouraging-active-demand>
- (11) <https://wallethub.com/edu/bank-market-share-by-deposits/25587/>  
<https://www.creditcards.com/credit-card-news/market-share-statistics.php>
- (12) <https://digitalfirst.com/yodlee-bank-feeds-interview/>
- (13) <https://wasterecycling.org/members/resource-library/size-report>;  
<http://www.waste360.com/financials/publicly-traded-solid-waste-companies-look-ahead>  
[https://www.globalwaterintel.com/client\\_media/uploaded/files/sample\\_water\\_market\\_USA.pdf](https://www.globalwaterintel.com/client_media/uploaded/files/sample_water_market_USA.pdf)  
<https://www.eia.gov/electricity/data/eia861/>  
<https://www.aga.org/knowledgecenter/facts-and-data/annual-statistics/annual-report-volumes-revenues-and-customers>
- (14) <https://www.mckinsey.com/industries/financial-services/our-insights/digital-insurance>