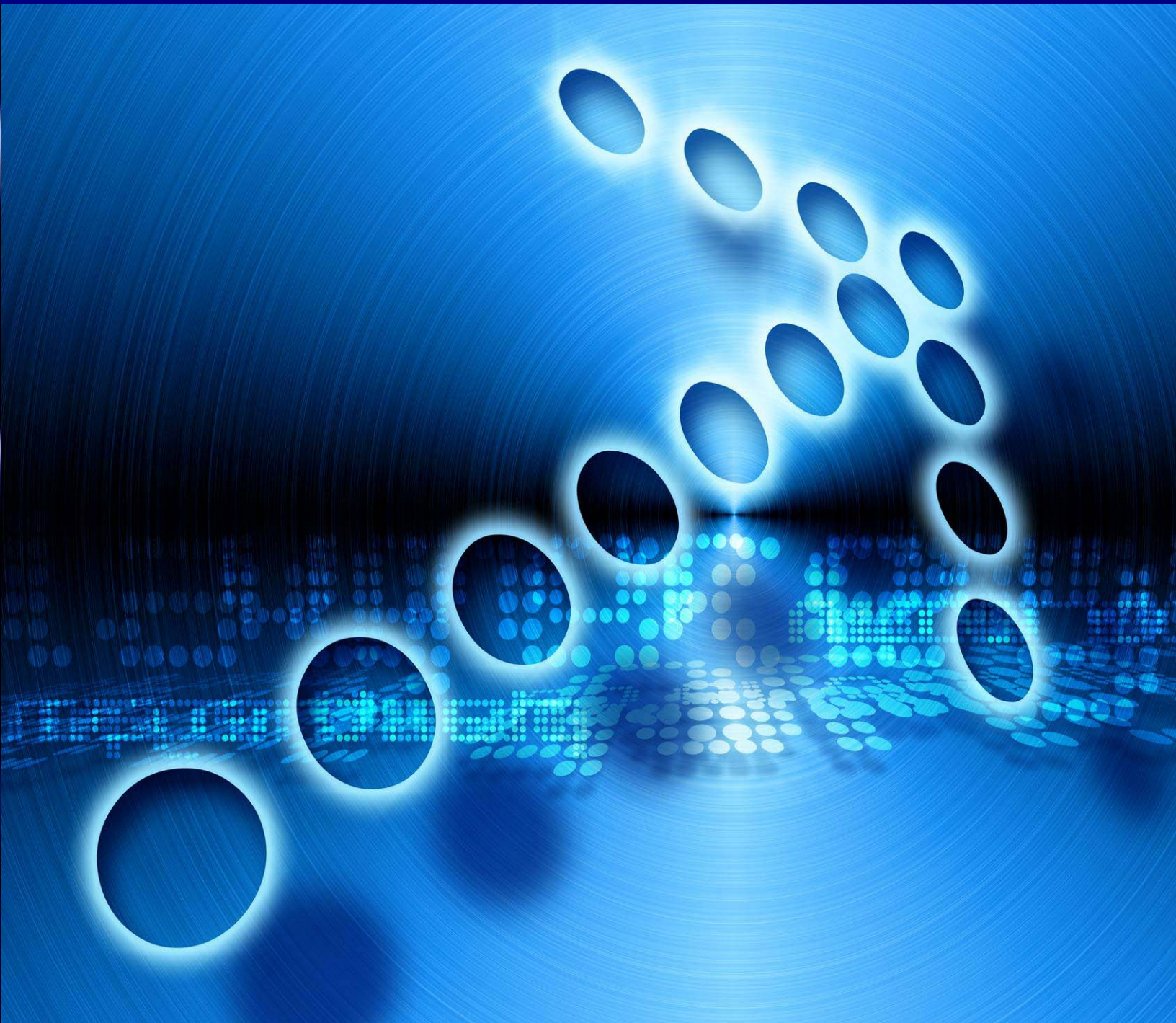# GREEN COMPUTING

Searchable PDF for Document Storage and The Concept of Green Computing

Aquaforest Limited

www.aquaforest.com

Searchable PDF for Document Storage and The Concept of Green Computing

The purpose of this white paper is to address the idea of long-term document storage as a "green IT" initiative, to highlight the advantages of adopting Green Computing and to promote the benefits of the paperless office, both commercially and environmentally.

What is Green IT?

In today's' climate, it is difficult to ignore the ongoing concerns about climate change and the environment and whereas there appear to be no easy solutions for businesses, the IT community can have a significant impact on the worldwide carbon footprint by adopting a 'greener' approach to computing.

So what is Green IT? Green computing or Green IT is environmentally sustainable IT which enables processes to run 'efficiently and effectively with minimal or no impact on the environment' [1] * Green IT strives to improve company performance and efficiency whilst offering an attractive economical saving and as such there are no real reasons as to why we shouldn't all be adopting a 'greener' approach to our processes.

This outlook is reinforced by initiatives such as the Global Warming Initiatives Inc [2] whose primary mission is to assist businesses in turning energy efficiency and environmental performance into a corporate asset while reducing global warming.

In short, tackling environmental issues has become a corporate and social responsibility as well as an economic and environmental imperative.

The Benefits of Being Green

We have established that heightened focus on environmental concerns is prompting many organizations to consider the benefits of adopting green technology which include:

- Cost
- Efficiency & Improved Performance
- Environmental Sustainability throughout the entire IT lifecycle, making it greener by addressing by addressing key areas including:
  - Green use
  - Green disposal
  - Green design
  - Green manufacturing

Interestingly enough, the greatest financial and environmental returns can be achieved by re-assessing current processes, and one area in which organizations can have an immediate impact is the long term management and storage of their business documents.

Long Term Storage of Documents

Traditionally the long-term storage of documents has been achieved through paper archiving and according to the US Environmental Protection Agency **[3]** the typical office worker in the US uses approximately 10,000 sheets of copy paper annually with the average document being copied 19 times during the course of its work cycle; this equates to an annual consumption of 85 million tons or paper each year. With each transaction costing the company both time and money, these statistics not only highlight the prevalence of paper waste but also raises the question of economic efficiency as each act of filing, replacing and retrieving paper documents has a cost implication.

However the introduction of the digital document has radically changed the way businesses are managing their work flow documents. Websites such as www.GreenPDF.com **[4]** aim to educate the public on the environmental impact of printing PDF documents. According to one of their press releases, if each PDF user printed one less document, it would save 80,000 trees & 118,000 tons of carbon dioxide from entering the atmosphere each year.

Ultimately, the paperless office may remain a hypothetical concept to many but more and more organizations are seeking to make a reduction in environmental impact by reverting to an electronic storage system. Therefore, the pertinent question is not whether we store our documents electronically but how?

PDF Files – The Ideal Format for the Paperless Office

In order to successfully implement and sustain a paperless environment, a good universal digital format is required and the choice of storage format for electronic documents is varied. Nonetheless, it is the Portable Document Format (PDF) format, created by Adobe in 1993 for the purpose of portable document exchange between systems and applications, which has emerged as the industry leader.

PDF documents have a variety of features like security, compact size, accessibility, usage restrictions and the ability to store all kinds of image formats and other data within the PDF itself. With only a free PDF Reader application, PDF files can be used on any system and all platforms with no exceptions. Moreover, PDF files are considerably smaller than other formats, which make them easier for file storage and exchanging over the Internet. With these advantages, PDF is widely considered to be the electronic equivalent of paper and comes as the best format to be used in a paperless office.

PDF was also designed to enable full 'searchability' of documents as it contains the text within the document itself. So unless the PDF is "image-only" it can always be searched - both within a PDF viewer application and via external search tools configured to index PDFs.

Furthermore, for the long term standards-based archiving, the presence of the PDF/A standard **[8]** makes PDF the clear choice.

How to Establish a Paperless Office

With the assistance of modern computers and scanners, the paperless office is not hard to establish. Scanners can be used to import paper documents into a computer for long-term storage: The scanner makes a copy of the image which is then turned into a scanned PDF or TIFF file. Whereas the files created by scanners are not electronic character-based files but unsearchable images of the original document, there are solutions such as OCR (Optical Character Recognition) technology at hand to convert these scanned documents into fully Searchable PDF files.

OCR Technology – for scanned PDFs

In order to convert a scanned document into an electronic character-based file, OCR technology (Optical Character Recognition) is required. OCR technology can interpret each character in the image within a scanned PDF, making it possible to convert these scanned images into editable and searchable formats like Text or Microsoft Word. Software such as Aquaforest's TIFF Junction Professional [5], Autobahn DX [6] or OCR SDK for .Net [7] can convert batches of scanned documents to searchable PDFs for long-term archiving and retrieval.

Summary

Concisely, it is evident that in today's climate, it is increasingly expected of companies to manage the impact that business process can have on the environment. The good news is that through modern technology, such as the Aquaforest suite of products, we can adopt greener IT processes which not only have an efficient and economic benefit to an organization, but will also fulfill the brief for the corporate and social responsibilities we all share.

For further information about adopting a greener approach to your IT, or for implementing a paperless environment with the use of OCR technology, please contact the Aquaforest team at sales@aquaforest.com

References

[1] http://en.wikipedia.org/wiki/Green_computing

[2] http://www.gwi-nc.org/services.html

[3] http://epa.gov/osw/conserve/materials/paper/faqs.htm#offices

[4] http://www.greenpdf.com/

Other useful links

TIFF Versus PDF for Document Storage
http://www.aquaforest.com/files/tiff%20versus%20pdf.pdf

Guide to Creating Searchable PDFs
http://www.aquaforest.com/files/searchable%20PDF.pdf

PDF Specification
http://www.adobe.com/devnet/pdf/pdf_reference.html

[5] TIFF Junction Professional
http://www.aquaforest.com/en/merge_tiff_junction.asp

[6] Autobahn DX
http://www.aquaforest.com/en/autobahn.asp

[7] OCR SDK for .Net
http://www.aquaforest.com/en/ocrsdk.asp

[8] BSI Legal Admissibility
http://www.bsigroup.com/en/Shop/Publication-Detail/?pid=000000000030104568