Abstract
There is an increasing amount of documents in today’s businesses, phenomenon particularly accentuated in the new-centric organisations. In this paper we examine the core principles of what constitutes a basic system processing the flow of documents in such an environment. Gradually, we introduce the modules that build up the relative systems and the respective parameters including technical specifications and the human element. Furthermore, we examine the merits of the using a document workflow system, the issues that arise and potential measures to tackle with potential emerging difficulties.

Keywords: Document Workflow, CSCW, Human interaction

1. Introduction
1.1 Getting to know the concepts
One of the most important tasks in today’s business world is the management of the mass amount of documents generated throughout the course of operations. Equally, it comprises a key variable in producing cost-effective and optimal results from augmented workload (Marchertti, Tesconi & Minutoli, 2005). Handling documentation is vital considering that at many instances there are different organizational participants working in separate sections of the organization and they need to fill forms, reports etc reflecting their particular tasks and responsibilities. Considering that the process of generating documents can be concurrent, there is a requirement for efficient mechanisms to handle the flow of documents in order to guarantee the smooth operations in organizations (Alder, Nash and Noel, 2006).

Making an attempt to define workflow in an organisation, it can be viewed as a the automation of a business process, wholly or partly in which documents, information, tasks etc are passed from one participant to another for action according to a set of rules based on prescribed procedures (Krishnan, Munuga, Karlapalem, 2001). As an extension of this definition, a document workflow or DW refers to the corresponding automation and processing of the relevant document through collaborative effort with the actual document being either the objective or a part of this endeavour.

1.2 Document Workflow sequence
Figure 1 illustrates the backbone document workflows principals. At the first level we have the various authors that contributed written pieces of work, information and other data. All the documents are inputted in a central database where depending on the architectural tenets and the relevant permissions other authors or prescribed editors can then view, edit and process the documents. The outcome of the input and processing of the documents as we denote in this case is a finalized published document. Essentially, the entire process is automatised through the use of the enabling technology as per the definition of the document workflow.

At the introductory stage, the various authors input the documents in a digitized form in a PC that is linked with the rest of the system depicted in figure 1. There are several ways to input the information either directly in an electronic format or through a capture device such as a scanner, a multifunctional printer or a copier depending on the content and the expected results the author requires from the conversion process (Simske & Arnabat, 2006). Understandably there is relevant enabling software to propel the procedure and basically automate the flow of the documents (Marchertti, Tesconi & Minutoli, 2005).

The functions performed in through the software include actions such as uploading and storing the documents in a shared space where it can retrieved by other users or editors for reviewing and further processing. Examples of such technologies enabling the operation of workflows include commercial products such as Microsoft Sharepoint (Microsoft, 2006), web-enabled technologies or a hybrid mix.

In the utilisation of the technical feature, essentially there are several actions that take place on the document either from the human element that manipulates the document through the software or by the automated agent that is
programmed to process in a form the document (Marcher tti, Tesconi & Minutoli, 2005). Examples could include an author that logs into the site to retrieve a document placed there by another member of a group and transmit it to another member. Equally, an example of a software agent that performs automated functions could include splitting the document and sending bits to various members of the group and merging different documents from various authors (Marcher tti, Tesconi & Minutoli, 2005).

The whole routing of documents is performed in a structured and set manner with clearly defined rules at its stage which is the base for architectural principals when designing a relevant system (Marcher tti, Tesconi & Minutoli, 2005; Wang & Kumar, 2005). Essentially in the workflow, every person of group is responsible to carry out a prescribed task, the in sequence the software informs and passes on to the person who is responsible to carry out the task assigned to the next stage. For example, in figure 1 above an author places a document into the system for the editor to proofread and then once the editor has completed his reviewing, the software forwards it to the publishing section of the organisation.

Although figure 1 is simplistic it also serves an illustration of an important element in the flow of the documents through the organisation, the collaboration between the various parties that interact. We could consider for example author 1 and editor/author 3. The first writes a paper, digitizes and inputs in the system. Sequentially, author 3 acting as an editor logs in the database and accesses the document, reads it and edits it. Once these steps are completed, the document is inputted again in the central database where the end result is the published document. These sequential movements through the aid of the computerised system represent a simple collaboration between author 1 and author/editor 3.

The same relationship equally exists for the rest of the authors and editors in the same environment. It worthwhile noting that this collaboration takes place through the utilisation of an computerised system, basically the same technical approach used to automate the flow of documents. This type of collaboration is often termed as computer supported collaborative work (CSCW) and at is simplest it refers to designing and using computer systems to support the collaboration among people (Ellis et. al , 1991; Grudin, 1991;Schimdt & Bannon, 1992;Hutchins, 1995). The collaboration that evolves inside the routing of documents is an integral piece of the system not only as a function but also a philosophy that is embedded in every part of the document flow.

1.3 Multi-level document workflow sequence

In figure 2, we can see a graphical representation of the sequence of a document workflow within a publishing environment. At the first level, author 1 inputs a document into the system and stores it in the central database. Intelligent agent represents the agent that automates the procedures essentially the software that is utilised for this purpose. The intelligent agent as explained is the technical enabler and the various actions are precisely the automated steps within the workflow.

As such, in the illustration, author 1 inputs a document in the system through the intelligent agent which is denoted by action 1. Following, the agent sends the document to the central database of the system (action 2) and then that is forwarded to the editor for review (action 3 & 4). Continuing the sequence, the editor once he has completed the reviewing process and sends through the intelligent agent the document back to the central database (action 5&6). Finally again through the intervention of the intelligent agent, the document is dispatched to the publication section for publishing (action 7&8).

Naturally the relationship inside the operations of the system can vary as for example the editor might require alterations to the document and therefore send it back to the central database from where the system will notify the author to make the changes necessary and resubmit the document. The later would of course lead to a set of additional actions multiplying the total operations of the intelligent agents in the workflow. Additionally by placing in the picture other authors there would be a relevant increase in the layers of the workflow or even create new workflows by inputting different documents etc in the system.

With regards to the intelligent agents that essentially adjust the route of documents in the system, Marcher tti, Tesconi & Minutoli (2005) offered a purpose orient classification defining the as external and internal. External agents role can be fulfilled either by human or software and is attached to the process of a specific workflow whilst internal agent operations are embedded into the system and are useful for the execution of any workflow. Applying this categorisation in the earlier illustrations, author 1 is an external agent for this particular workflow as he inputs the original source for further processing into the system. On the other end, the intelligent agents in the illustration are the ones responsible for performing functions (e.g circulating documents, informing editor etc) necessary for the materialisation of any workflow.

Simulating figure in a real-time publishing environment (e.g a newspaper) a journalist utilizes some technological means to digitize an article he has written and place it into the main database of the newspaper. The flow of
movements, from the beginning the author commences an interaction with the system as for example using a PC connected to the intranet of the newspaper of using his home PC to access the web and send the article by e-mail the sequence of actions commences and essentially the flow of the documents inside the newspaper. As explained earlier, the intelligent agent denotes the technical facilitator both in terms of hardware and software to automate the process of document circulation. Naturally, the intelligent agent is not restricted to one computer or the use of the web as it can be a dedicated system to managing documents unique to the newspaper.

We can note the flow of the document as it progresses to the new stage once the journalist has placed it inside the central database and then the editor log on to access the article for inspection. Once the editor completes his scrutiny over the submitted piece he places his observations which could include remarks for revision by the author, corrections made on the spot by the editor or if the document is fine, the direct approval for publication. In the first instance, the article is inputted into the system and the journalist is notified to access the document, see the remarks of the editor and make the necessary adjustments. Again we can see the role of intelligent agents in enabling the flow of document from author the editor and if need be again between them, a sequence that leads to up until the publication stage. It is worthwhile pointing the auxiliary functions in a real-time workflow environment such as the communication between the parties (e.g via e-mail) and broad access for example using a mobile internet connection.

It is worthwhile illustrating the differences between figure 1 and 2 in both in terms of design and functionality. First of all, figure 2 describes a considerably more complex situation as we have a more detailed break-down of the various interactions between the involved parties (author, editor etc). Equally, in figure 2 we introduce the technical facilitator, the intelligent agents that automate the flow of documents. As pointed, the intelligent agents essentially is the engine to succour collaboration between the various people in the newspaper not only by being the medium to move the input back and forth but also to allow for the input to be processed at certain stages (e.g the editorial stage) while all the time, enabling communication and internal information exchange.

Further from the degree of complexity in the relationships between the interacting parties, figure 2 also depicts in more detail the functional characteristics of a document workflow sequence. As in a real-time situation there are many functions involved from the introduction of the article in a digitized format to its final publication. This particularly important in comprehending the role of the intelligent agents in the sequence and also the particular functions involved at its stage. The latter is extremely useful when having to design a relative system or even upgrade the existing one with new feature that will enable more necessary functions. Equally, by mapping down the functions in the sequence there can be the appropriate monitoring to guarantee the smooth running of the system and intervention if need be to avoid bottlenecks.

2. Workflow Management Systems

2.1 WMS in retrospect

As it was mentioned above, document workflow is an integral and essential part to the daily operation of organisations. The technical materialisation of the flow of documents can be termed as WFMS or Workflow Management System that essentially is comprised by the software used to support the automation of the business processes. Certainly, document processing is not the sole functionality of a WFMS but an important function nevertheless. (Bae and Kim, 2001). Interchangeably, we could use the term document management system which essentially embodies all the aspects pertaining to document administration and handling and thus WFMS is an integral part of it (Sprague, 1995).

With the advent of the internet and the expansion of web technologies there has been a significant pursuit of digitisation of the older paper-based workflows and a movement towards digitised document imaging, storage, distribution and overall document management system with companies reaping significant gains (LaCava, 2003; Computer Weekly, 2005; Fall 2005).

Equally, in figure 2 we introduce the technical facilitator, the intelligent agents that automate the flow of documents. As pointed, the intelligent agents essentially is the engine to succour collaboration between the various people in the newspaper not only by being the medium to move the input back and forth but also to allow for the input to be processed at certain stages (e.g the editorial stage) while all the time, enabling communication and internal information exchange.

Further from the degree of complexity in the relationships between the interacting parties, figure 2 also depicts in more detail the functional characteristics of a document workflow sequence. As in a real-time situation there are many functions involved from the introduction of the article in a digitized format to its final publication. This particularly important in comprehending the role of the intelligent agents in the sequence and also the particular functions involved at its stage. The latter is extremely useful when having to design a relative system or even upgrade the existing one with new feature that will enable more necessary functions. Equally, by mapping down the functions in the sequence there can be the appropriate monitoring to guarantee the smooth running of the system and intervention if need be to avoid bottlenecks.

2. Workflow Management Systems

2.1 WMS in retrospect

As it was mentioned above, document workflow is an integral and essential part to the daily operation of organisations. The technical materialisation of the flow of documents can be termed as WFMS or Workflow Management System that essentially is comprised by the software used to support the automation of the business processes. Certainly, document processing is not the sole functionality of a WFMS but an important function nevertheless. (Bae and Kim, 2001). Interchangeably, we could use the term document management system which essentially embodies all the aspects pertaining to document administration and handling and thus WFMS is an integral part of it (Sprague, 1995).

With the advent of the internet and the expansion of web technologies there has been a significant pursuit of digitisation of the older paper-based workflows and a movement towards digitised document imaging, storage, distribution and overall document management system with companies reaping significant gains (LaCava, 2003; Computer Weekly, 2005; Fall 2005).

Essentially, we have the necessities of the document route inside an organisation as expressed by the features of the document system and on the other end, the web technologies partake the role of an enabler to accomplish and supplement these features. The advantages of using a combination of web-succoured Document Management System converge to a significant degree with the requirements mentioned earlier. For example, the large available space on the internet allows the management of large amount of materials; it provides access to the user from remote places with a compatible interface and navigational mechanism. In such a fashion, the flow of documents and information becomes easier, faster and more approachable to users (Balasramania & Bashian, 1998; Aversano, Canfora, De Lucia & Gallucci, 2002; Dustdar, 2005).

Drawing a connection thus far, in generic document workflow structures we notice a level of resemblance but also some additional elements that arise. Again we have the various agents both human (author, editor etc.) and software,
the various actions that take place between the agents and all are filtered to heart, the DMS which is similar to the central database used in the introductory figures in the sense of acting as a repository of the interaction for all the agents in the workflow but further from that here it enacts, co-ordinates and processes the interactions between the parties. Further novelties of this schematic approach are the introduction of the web as a motion platform for the flow of the documents with online applications such as the link checker and the full-text indexer. Lastly, we have enriched the role of human elements introducing the role of a technical team that supports and administers the relevant interface providing to the authors the structure of product information (i.e layout, electronic format etc) using templates.

We can notice the interaction between the author, editor and the new entity introduced in this figure the legal department that also is a part in a workflow system in a newspaper for reasons of copyright check, permission to reproduce etc. In figure 2 earlier we denoted the various interactions as actions, here we have the title for each one. For example creates an article, logs it into the system and then accesses it in order modify/update it after the editor’s instructions. A new characteristic we can see here is that the author can also notify the editor about the submission status of the document and vice versa. The repository of the flow, the main database which is inspected and maintained by the technical staff again displays the interactive relationship between the contributing parties as they place the revised or edit document inside the database and from there follows the relevant communication (e.g notification to the author for changes or publication) similar to what previously seen in figure 2. The main functional difference is the use of the web in the publication stage as the supplement to the native DMS of the newspaper.

We have already pointed the technical difference with the use of text-formatting tools but is also worth stressing the difference in the architectural structure with introducing the web. Earlier, in figure 2 we only had a native Document management system with all the functions incorporated. Here we have two additional layers, the staging and the production web. Especially the staging web is an area where all the members of the flow can preview the document and from the beginning it is submitted right before publication. An important characteristic is this capacity is not incorporated in the main database but it makes use of the common workspace provided by the internet which enhances the ability of journalists or editors to collaborate more efficiently and over distances. Lastly, this feature enhances the total outcome in terms of quality as it permits to preview the finalised form of the document before it is publicised and to make if need be any modifications for the article to conform with the standards both content and in style.

The increase in demand for web-based applications in document workflow and digitations of the former paper-based systems gave a significant boost in the commercial applications of WMFS. Equally popular is the Microsoft Sharepoint server application that provides features for the collaborative management of the document workflow.

Definitely, some organisations have more enhanced needs of document management than others. For example, organisations in the news industry such as newspapers, news agency and other mass media related entities have a greater bulk of documents that can be in any sort paper-based, electronic or web-based that need to be managed effectively. As a rational consequence, the use of document workflow systems is bound to have a significant impact on the particular field and improve significantly the flow and processing of documents inside such organisations.

Essentially, the implementation of a document management system to handle the load of documents, images and data and the necessary flow among the various divisions and parties in an organisational setting. The advantages are obvious in terms of improving the collaboration within the setting and improving performance (Catton 2006; Dustdar, 2005). As mentioned earlier, in information and document intensive domains such as newspapers both conventional and online, the flow of information is an essential and integral part (Smeaton et. al 1998; Castells et. al, 2005).

2.2 Collaborative writing in a workflow environment

An essential part of the course of documents inside an environment and particularly in the news-intensive settings it the ability to produce collaboratively authored or edited documents. As we can derive from the majority of the illustrations, the process of inputting data, text, images etc inside the document management system further from the actual procedures it involves collaboration of the members / users aided by automisation that occurs. For example in figures 1 and 2 earlier a paper may be collectively written from scratch, or pass through a series of revisions and editing as it moves through the hierarchy of an organisation.

Support may be for synchronous collaborative writing at a distance, where two authors discuss and revise a document as they would be sitting together at the same desk, even though they are many miles apart. The style of drafting and redrafting where different people work on the document at different times requires asynchronous collaborative support. As an integral part of the document workflow, several benefits are attached with the
collaborative side of authorship. As noted in literature benefits could include (Bacon, 1990; Bogert and Butt, 1990; Cross, 1994; Forman and Katsky, 1986; Haley, 2001; Nelson and Smith, 1990).

- The production of enhanced quality in the output of the documents due to the broadened elements the users bring into flow

- Increased levels of overall motivation as members often will support one another to perform their best in their contribution.

- The participation and observation of co-writers in the early stages of the document compilation (i.e. draft stage) is possible to provide valuable comments improving thus the timeline of the overall processes.

- Less experienced and novice writers can exploit the opportunity of working with more senior colleagues thus the work relationships within the organisation will be strengthened and the final documents could have a higher acceptance rate due to the joined effort.

Looking further into the collaborative writing segment of the document workflow we can see there can be several modes of co-operation among the users. For example, referring to figure 2, the users might be working simultaneously on the same document (e.g. author 1 placing the document into the system and the editor reviewing at a synchronous mode the document). Equally, the mode can be asynchronous with the members working towards their part in the system at different time periods (Lee, Narayan & Chan, 2000). The alternating modes of collaboration through the meditation of the system further to its advantages also create some issues that need to be addressed both technical and social.

Primarily, when a document is either co-authored or edited by different members it is likely that several parts will be modified and regenerated. It is essential therefore that any CSCW system adopted needs to support consistency both by allowing access to the involved authors but also in terms of avoiding bottlenecks due to system latency or break downs. Further to system parameters the role of the users is also important. Coen and Petelin (2004) argue that it is possible when users have different writing styles, conflicting perceptions about varying issues which could make them negatively responsive to other authors in the group that generate work based on such ideas creating thus internal difficulties in the total outcome. Certainly this goes further from a system adaptation but is an important example on the role the participants and their interactions can play in the final outcome produced. Nevertheless, there are some important features within the collaborative writing environment that extend to the total functionality of the document management system.

First of all there must be a two-way view of the document, one perceiving the document as whole and one focusing on specific parts of it. In such a fashion it is possible for users to discuss about various observations or changes that are pertaining either to the total body of the text or to specific parts of it. Quite often, writers are used in specific working environments such a particular word-processors and are more at ease when using this environment. In order not to disrupt the sequence of the flow and keep the quality and a certain standard, supporting varying such environments could be considerably beneficial.

Equally important, the various participants either authors or editors should be aware or have some information on the degree of completion of the other participants appointed tasks. This is a quite important feature, as real-time information is essential for an ongoing collaborative environment as it would allow for communication to also be in alignment with the expectations from the system whilst building an overall team-spirit. At the same time we should not discard to the possibility for the system to enable co-operation at various levels of the document circle. For example, should an editor discover some omissions or need for alteration at an early stage by exchanging information with the author or intervening himself can make whatever adjustments are necessary at the current stage saving thus time in the later stages of the document flow.

2.3 Issues in collaborative author systems: Emphasizing communication

In the hybrid mix, with the utilisation of technology both internally with the native DMS and with the utilisation of the internet as pointed, collaboration becomes more enhanced and the capabilities of participants in achieving optimal result are more pronounced. As noted in earlier, the document workflow relies essentially in a combination of the technology used to automate the flow in the organisation and the people interacting with each other through the system to produce the final outcome, essentially the published document. As a consequence, in order for the DMS to function properly the human factor must be adept with the system’s features, appreciate its usefulness and maximise its potential. The later essentially assumes two important themes. First of all that the system is not complicated and as user-friendly as it is attainable without comprising it operations. The second theme involves that successful utilisation of the system by its users which is more complicated as people’s individual characteristics, perceptions and opinions are mixed together creating a rather complex setting.

In theory, a document management system might indeed hold all the potential it promises and can significantly improve the total output both in terms of quality, quantity and time. Nevertheless, whilst the technical specifications
can be considered relatively constant, the human factor in practice can vary. In our case, in a computer-mediated collaborative document might encounter practical difficulties (Dillon, 1994). Essentially this area stresses the human side on a collaborative environment and the ability of the hardware used to support and coordinate the collaboration to actively stimulate and support the process. This is an important consideration not solely for the co-authored documents but for the entire document management system design.

In order to facilitate the smooth collaboration among the different participants in the system the architecture needs to overcome hurdles pertaining not only to technical bottlenecks but also issues such as author conflicts, real-time and active discussion on the various topics that arise in the various stages of the document’s flow. Working towards that path, it is essential to provide the opportunity and if possible the indirect stimuli for communication among authors and editors in this environment which would allow to augment to overall background for collaboration. Certainly this does not imply the disruption of the system as it is necessary for the document flow to go forth the existence of a movements with grounded rules as mentioned in the beginning. Nevertheless, the later does not preclude the existence of additional features in the system that without burdening or impeding on the procedural steps to enhance the collaborative spirit of the participants.

Table 1 recaptures the main issues in the last sections and offers a series of recommendations for future directions in the document workflow pathway.

In the same context, another important element that should be considered in the human factor is precisely its unique characteristics. People are very often afraid or sceptic about changes and new technologies (Slack et. al 1998; Bartol & Martin 1999). It could be the case in the use of collaborative author tools and for the whole of the document management system. People might not be able to comprehend the use of the system or not appreciate its usefulness or even being afraid that the automation might eventually make their work redundant. There are quite many factors influencing the behaviour of people and all are especially subtle. The definitive issue is that without the open contribution of the human element, the DMS at the very least will not achieve its anticipated targets.

3. Conclusion

It is vital not only for the system to incorporate user friendly features but also to cultivate a necessary collaborative culture in the human factor. Adopting strategies of openness, explaining thoroughly the specifications and positive outcomes the system can produce would be a outlet to nourish the proactive nature of people towards the DMS. The same could include some additional training and seminars to introduce all the technical capabilities and increase trust both to the system but also to one another. A combination of technical orientation approachable to the average user and the cultivation of a collaborative culture to succour the use of the system could provide a working solution to surpass the arising difficulties.

Acknowledgments: I wish to thank wholeheartedly Sofia Tsivartzi for all her support over the course of time without which this paper and many more would not be.

References


Table 1. Document workflow, current and forward course

<table>
<thead>
<tr>
<th>Proponents</th>
<th>Current course 1</th>
<th>Current course 2</th>
<th>Future</th>
</tr>
</thead>
<tbody>
<tr>
<td>DMS</td>
<td>Automating tasks in the flow of document</td>
<td>Allowing interactions between participants from the beginning to the end</td>
<td>Exploitation of the internet as the source for common workspace and communication</td>
</tr>
<tr>
<td>Technical Specs.</td>
<td>Allowing communication between participants in the flow, providing auxiliary functions</td>
<td>Providing enhanced effectiveness of D.M.S in total and of the people using the system</td>
<td>Security issues Making use of mobile technologies</td>
</tr>
<tr>
<td>Human Factor</td>
<td>Integral part of the D.M.S</td>
<td>Issues in regards to use of the system and communication with other members</td>
<td>Need to embrace technology and of increased motivation to become more adept in its usage.</td>
</tr>
</tbody>
</table>

Figure 1. Basic structure of Document Workflow principals
Figure 2. Document Workflow sequence