

## The City University of New York Application For PSC-CUNY Research Award Program

 [Print this page](#)
[Evaluate this application](#) 

<b>Control Number:</b> 35-6931	<b>Name:</b> (Last, First) HOLOWCZAK, RICHARD		<b>Review Panel:</b> Computer Science	
<b>Rank:</b> Associate Professor	<b>Tenure:</b> Y	<b>Junior:</b> N	<b>Type of Award:</b> 1YR	
<b>Department:</b> STATISTICS	<b>College:</b> BARUCH COLLEGE		<b>Subjects:</b> Human: N Animal: N	<b>Progress Report:</b>
<b>Supplementary Materials attached:</b> N	<b>Telephone:</b> (H): (W)		<b>Email Address:</b> richard_holowczak@baruch.cuny.edu	
<b>Title of Proposed Project:</b> Streaming Multimedia and Market Data Synchronization				<b>Amount Requested:</b> \$ 7,998.00

**Brief Abstract:**

"MarketVCR" is an information system consisting of a large multimedia data warehouse containing tick-by-tick quote and trade stock market data, video, audio and news articles, a web application capable of tapping into the warehouse and a middle-tier server suite capable of streaming data to the web application. MarketVCR will allow an instructor or student, in any location, at any point in the future, to "rewind" the market back to the time of a critical market event and "play" the market with synchronized video, data and news releases just as it happened earlier in the day, week or year. A key component of the playback subsystem will rely on the ability to synchronize the playback of audio and video with the display of market data. Quote and trade data may constitute over one hundred records per minute in an actively traded financial instrument. MarketVCR must be able to scale to coordinate the display of dozens of such financial instruments simultaneously. Past work has focused on the coordination of audio and video streams with a relatively small number of discrete objects such as audio accompaniment to a slide show. Our aim is to apply current principles in this domain to augment synchronization models to accommodate streaming market data objects.

<b>Budget</b>		
Research Staff/Part Time	<b>\$7,200.00</b>	(5410)
9% of Research Staff/Part Time fringe benefits	<b>\$648.00</b>	(5940)
Manuscript Prep/Publication: Manuscript Preparation	<b>\$150.00</b>	(8040)
Total:	<b>\$7,998.00</b>	(Amount Requested)

**Budget justification in order of importance:**

The Research Assistant will participate in an intensive review of the relevant literature and implementation of algorithms using the Java programming language. To ensure continuity, the research assistant should be made available 15 hours per week over the course of 32 weeks. (32 weeks) \* (15 hours/week) \* (\$15/hour) = \$7,200 The Research Assistant should have the following skills: Proficiency with the Java programming language (applets and servlets), familiarity with working under the UNIX operating system including compiling and running programs, developing shell scripts and writing scripts in a general purpose language such as Perl or Python. Familiarity with streaming multimedia and basic web design would be useful as well. The above requirements are commensurate with the skills of a CUNY graduate student. As per the PSC-CUNY Research Award Guidelines, all reasonable attempts will be made to work with a student from the CUNY community.

**Biographical Summary**

<b>Education</b>			
<b>Institution</b>	<b>Degree</b>	<b>Year</b>	<b>Field of Study</b>
The College of New Jersey	B.S.	1990	Computer Science
Rutgers University	M.B.A.	1992	Management Information Systems
New Jersey Institute of Technology	M.S.	1997	Computer Science
Rutgers University	Ph.D.	1997	Computer Information Systems

Publications, Performances or Works

Adam, N. R., Artigas, F., Atluri, V., Chun, S., Colbert, S., Degeratu, M., Ebeid, A., Hatzivassiloglou, V., Holowczak, R., Marcolpus, O., Mazzoleni, P. and Rayner, W. "E-Government: Human-Centered Systems for Business Services." Proceedings of the First Annual Digital Government Conference (DGO 2001). Los Angeles, California. May, 2001

Liu, L. and Holowczak, R. "Using Reuters 3000 Xtra for Financial Information Education." Online Information Review: The International Journal of Digital Information Research and Use. 24(5), November, 2000. Page 371-380.

Publications, Performances or Works resulted from PSC-CUNY funding

\* Holowczak, R., Adam, N., Artigas, A. and Bora, I. "Data Warehousing for Environmental Digital Libraries." Communications of the ACM. 2003.

\* Holowczak, R., Artigas, F., Chun, S., Cho, J.S. and Stone, H. "An Experimental Study on Content-based Image Classification for Satellite Image Databases." IEEE Transactions on Geoscience and Remote Sensing. Volume 40. Number 6. June, 2002.

\* Holowczak, R., Chun, S., Artigas, F., and Atluri, V. "Customized Geospatial Workflows for E-Government Services." The Ninth ACM International Symposium on Advances in Geographic Information Systems. November, 2001.

\* Adam, N. R., Atluri, V., Adiwijaya, I., Banerjee, S. and Holowczak, R. "A Dynamic Manifestation Approach for Providing Universal Access to Digital Library Objects." IEEE Transactions on Knowledge and Data Engineering. 13(4), July/August 2001.

\* Artigas, F., Barrett, K. and Holowczak, R. "Digital Meadowlands: A web-based environmental decision support system for management and research of an urban, estuarine watershed." Proceedings of the Integrated Decision-Making for Watershed Management Symposium, Washington, D.C. January, 2001.

\* Barrett, K., Holowczak, R. and Artigas, F. "A database of environmental documents about an Urban Estuary, with a WWW-based, geographic interface." Proceedings of the Symposium on Water Resources and the World Wide Web. Seattle, Washington, December 5, 1999

**Other Funding**





<b>PSC-CUNY Awards (over past five years)</b>				
<b>Date</b>	<b>Title</b>	<b>Project Number</b>	<b>New or Renewal</b>	<b>Amount</b>
7/1/2003	Modeling of WebGIS Applications	65387-00 34	New	3,297.00
7/1/2002	Satellite Image Classification...	64401-00 33	New	4,190.00
7/1/2001	Data Warehouse Evolution to Support	63376-00 32	New	4,142.00
7/1/2000	Introperability & aggreg data wrhse	62428-00 31	New	3,547.00
7/1/1999	Data Warehousing Models to Support Digital Libraries	61387-00 30	New	3,718.00

**External Research Grants/Award Proposals**

<b>Date</b>	<b>Title</b>	<b>Funding Source</b>	<b>Amount</b>
-------------	--------------	-----------------------	---------------

External Research Grants/Award Proposals resulted from PSC-CUNY funding			
Date	Title	Funding Source	Amount
6/1/2000	Web Based Weather Data Information System	Meadowlands Environmental Research Institute	16,691.00

Proposals Under Review or in Preparation			
Date	Title	Funding Source	Amount

Attachments
 <a href="#">35-6931_budgetnotice.pdf</a>
 <a href="#">35-6931_letter.pdf</a>
 <a href="#">35-6931_review.pdf</a>
 <a href="#">Project_Description.Holowczak.doc</a>
Progress Report
No Progress Report Available

Evaluate this application

---

The Research Foundation of  
The City University of New York  
[www.rfcuny.org](http://www.rfcuny.org)

## Project Description

### Introduction

“MarketVCR” is a combination of a large multimedia data warehouse containing tick-by-tick quote and trade stock market data, video, audio and news articles, a web application capable of tapping into the warehouse and a middle-tier server suite capable of streaming data to the web application. MarketVCR will allow an instructor or student, in any location, at any point in the future, to “rewind” the market back to the time of a market altering event and “play” the market with synchronized video, data and news releases just as it happened earlier in the day, week or year. A key component of the playback subsystem will rely on the ability to synchronize the playback of audio and video with the display of market data. Quote and trade data may constitute over one hundred records per minute in an actively traded financial instrument. MarketVCR must be able to scale to coordinate the display of dozens of such financial instruments simultaneously.

### Problem Statement

Past work has focused on the coordination of audio and video streams with relatively few discrete objects such as audio accompaniment to a slide show. Our aim is to apply current principles in this domain to augment such systems with the streaming market data components. The challenges for this work stem from the dense nature of market data, its non-deterministic arrival times and how it must be streamed to the client in VCR mode. When originally delivered directly from a market center (such as the NYSE or NASDAQ), quote and trade data are posted on the client as soon as the network and systems can deliver them. Upon playback in MarketVCR, such data must be posted on the client in step with the timing imposed by the playback of the accompanying video. **While several theoretical approaches exist to model the coordination of a relatively small number of deterministic objects, an efficient and effective approach to modeling and implementing synchronized market data has not previously been developed.**

### Significance

The significance of this work is twofold: First, the extension of existing models to account for the arrival and display of market data constitute a novel contribution to this research domain. Second, development of efficient and effective methods for implementing such models will benefit both the larger MarketVCR project as well as researchers and practitioners facing similar issues.

Research work in this area has a number of relevant publication outlets including several journals (Journal of Multimedia Tools and Applications and ACM Multimedia Systems Journal) as well as several high quality conferences sponsored by IEEE and ACM.

### Background

Breaking news events can have significant impact on financial markets. When the Federal Reserve chairman testifies before congress or presents open market committee rate adjustment decisions, when world leaders reveal major policy decisions, or when corporate leaders present news releases and performance statistics, financial markets around the world react within seconds often with significant impact on the pricing of financial instruments and the volume of transactions.

Live news coverage of such events coupled with live market data provide a robust and intriguing environment in which students can witness and apply theoretical notions of finance and economics in a real world context. A Federal Reserve committee decision to lower the prime

rate, for example, has implications for everything from bond pricing and equity pricing to the rates of student loans, mortgages and credit card premiums. Such real-world examples, supported by real-time financial market data and juxtaposed with live video or audio of the events, serve to reinforce in student's minds the many relationships between financial and economic concepts and provide an ideal complement to the related theory they have learned in the classroom.

Unfortunately, students not physically present at the time of the event are left out of such significant learning opportunities. In response to this need, we are in the process of constructing the "MarketVCR", a combination of a large multimedia data warehouse containing tick-by-tick quote and trade market data, video, audio and news articles, and a web application capable of tapping into the warehouse. Ultimately, the application will allow an instructor or student, in any location, at any point in the future, to "rewind" the market back to the time of the event and "play" the market with synchronized video and news releases just as it happened earlier in the day, week or year.

Cataloging video, audio, news articles and market data are relatively straightforward tasks that are currently being implemented and automated by the P.I. Streaming video and audio by themselves over the Internet are commonplace and supported by a large number of commercial software packages (e.g., Real Server/RealPlayer or Microsoft Media Server and Media Player). **The focus of this proposal is on the coordination of the playback of a video or audio stream with the display of market data.**

Market data are typically categorized into two groups: Quote data, which represent the bid and ask prices market participants are willing to accept or pay for a given quantity of a financial instrument (for simplicity here we will focus only on stocks), and Trade data that report sales transactions where a buyer and seller have agreed on a price and shares will exchange hands. In an actively traded stock such as Microsoft, 350,000 quotes will be posted and 100,000 trades will occur in an average eight-hour trading day. Each quote tuple consists of the stock symbol, date, time, bid, ask, bid size, ask size, exchange and market maker. Each trade tuple consists of the stock symbol, date, time, trade price, shares and exchange.

Most stocks do not trade as actively, however there are potentially thousands of stocks one might consider cataloging in this fashion. While this presents some challenges in terms of network throughput and storage, it is typically only researchers and students investigating market microstructure that are interested in the individual quote data. Hence MarketVCR will initially only work with the trade data that represent the "stock price" most people are familiar with.

An active stock may trade several times per second. An inactive stock may trade several times per day. This non-deterministic behavior sits in contrast to the frame based orientation of audio and video (A/V) streams that count on frames being displayed for a precise and fixed period of time. Once an A/V frame has been displayed, it can be discarded. Trade data, on the other hand, may arguably remain on display (perhaps in a graph) for an extended period of time after its arrival at the client.

These characteristics dictate novel approaches to model and implement synchronization constraints. Specifically the MarketVCR environment requires:

1. A multimedia synchronization model that incorporates the replay of real time market data.
2. A methodology to convert the synchronization model into an executable environment (e.g., SMIL code – see below) for implementation in a web-enabled client/server architecture.

## **Related Work**

A number of techniques and languages have been developed to model the coordinated playback of multimedia data objects. Dynamic Extended Finite State Machines [2], Object Composition Petri Net [3], Real Time Synchronization Model [4][6], and Multimedia Object Petri Nets [1] are examples of such approaches. Each approach has in common the notion of a media object (video, audio, text, image) represented by a place in the state space (or Petri net as the case may be) that will ultimately be enabled such that it will play or be displayed for a given duration. More recently, [7] has explored modeling of non-deterministic events permissible within the SMIL 2.0 environment that has led to the development of the Extended Real Time Synchronization Model.

Market data fails to fit within the existing categories of such multimedia data. Unlike video and audio which have distinct notions of frames of fixed time duration, or text or images that can be statically displayed for some duration, market data have no fixed "arrival" time and one may argue that once displayed, they should remain so for the duration of the presentation. Modeling each individual quote or trade report as a distinct object would prove intractable. It may be possible to model such data as non-deterministic events as discussed in [7].

Early models further assumed that a coordinated multimedia presentation would only be replayed from the beginning and perhaps paused and resumed in the middle. Incorporating full VCR-like functionality such as rewinding or fast-forwarding the presentation to different portions of the presentation as well as altering the rate of playback has recently been explored in [5] and [6]. Such approaches will need to be incorporated into MarketVCR.

The current standard client-side synchronization method of choice is the Synchronization Multimedia Integration Language (SMIL) developed by the World Wide Web Consortium [8]. SMIL is an XML based language that specifies the temporal behavior of multimedia presentations. A SMIL script containing specifications of various multimedia objects (source, appearance, and information on when display or playback should begin and end) can be interpreted by a SMIL compatible media player (such as RealOne player). Recent work [6] has explored implementing VCR-like controls using SMIL 2.0. It is our intention to make use of the SMIL language during the implementation phase to specify the coordination of MarketVCR objects.

Several sources of market data are available. Most noteworthy is the New York Stock Exchange (NYSE) Trade and Quote system (TAQ). TAQ retains quotes and trade data for all stocks listed on the NYSE. This data is available directly via download over the internet (under a license agreement) or via CD or DVD delivery. The Wharton Research Data Services provide access to the TAQ data on their systems as well as through a web based query/retrieval interface. Such interfaces follow a query/retrieval model that is generally more suited towards researchers who require a specific data set to be downloaded for their research.

## **Project Design**

The first phase of the project will consist of a more thorough literature review and the selection of the most appropriate synchronization modeling approach that will meet or comes closest to meeting the design goals of MarketVCR given the specific characteristics of market data. Extending or enhancing existing modeling approaches to accommodate market data may be necessary. This phase, currently underway, is anticipated to take no more than 3 months to complete.

In the second phase, a synchronization model of objects within MarketVCR will be developed.

Given the prior phase has been completed it is anticipated that the second phase should take no more than one month to complete.

The remainder of the project's duration will be spent implementing a proof of concept system capable of generating the appropriate SMIL 2.0 scripts given a set of input objects and the constraints imposed by the model developed in the second phase.

## References

- [1] N.R. Adam, V. Atluri, I. Adiwijaya, S. Banerjee, and R. Holowczak, "A Dynamic Manifestation Approach for Providing Universal Access to Digital Library Objects." IEEE Transactions on Knowledge and Data Engineering. 13(4), July/August 2001. pp. 705-716
- [2] C-M Huang and C. Wang, "Synchronization for interactive multimedia presentations," IEEE Multimedia, Oct.-Dec. 1998, pp.44–62.
- [3] T.D.C. Little and A. Ghafoor, "Synchronization and Storage Models for Multimedia Objects." IEEE Journal of Selected Areas in Communications, 8(3), 1990. pp. 413-427.
- [4] C. C. Yang and J. H. Huang, "A Multimedia Synchronization Model and its Implementation in Transport protocols," IEEE Journal of Selected Area in Communications, vol. 14, No. 1, Jan. 1996. pp. 212-225.
- [5] C. C. Yang, "Design of the Data Retrieving Engine for Distributed Multimedia Presentations," Proceedings, IEEE International Conference on Communications, 2001 (ICC2001), pp. 3237-3243.
- [6] C. C. Yang, C-W. Tien, and Y-C. Wang, "Supporting VCR-like Operations in SMIL2.0 Players,"  
IEEE International Conference on Multimedia and Expo. July, 2003.
- [7] C. C. Yang, C-W. Tien, and Y-C. Wang, "Modeling of the Non-deterministic Synchronization Behaviors in SMIL2.0 Documents," IEEE International Conference on Multimedia and Expo. July, 2003.
- [8] World Wide Web Consortium "Synchronized Multimedia Integration Language (SMIL 2.0)" W3C Recommendation 07 August 2001. Edited by Aaron Cohen and Thierry Michel  
<http://www.w3.org/TR/smil20/>