

## **Data Scientist (Python, Text Mining, Bioinformatics, Application Support)**

**Keywords:** *Text-Mining/Analytics, Ontologies, Bioinformatics, Pharma*

**Job Reference:** SCIS2018B

**Job Title:** Data Scientist

**Location:** Boston, MA

**Salary:** \$ Competitive dependent on experience

**Benefits:** Healthcare, 401K, eligible for stock options, bonus plan, 20 days' annual leave

**Hours:** 40 hours (Monday – Friday)

### **The Company**

SciBite is going places. A successful and fast-growing informatics company based in Cambridge, UK, we are transforming the way life science organizations handle text and documents. Our TERMite™ software performs rapid analysis of scientific and healthcare text to generate actionable insights for the top ten pharma, biotech and healthcare customers. In the last three years we have seen significant growth to an international organization with a presence in Japan and in the USA.

We are actively expanding our US presence, and are looking to hire a US-based Data Scientist to support our growing US team. Joining us means that as well as a highly competitive salary and fantastic benefits package, you will have the opportunity to learn and shape the cutting edge of informatics and drug discovery with some of the world's major pharma and life science companies. A dynamic, passionate and close-knit team, we're always learning new approaches and skills and will actively encourage you to engage in some of the hottest new technologies out there today.

We are now seeking an ambitious individual to provide data and text-mining support in a customer-facing role supporting exciting US-based customer projects. The successful candidate will engage with our customers on real-world use cases across pharma R&D. The role offers a chance to experience the cutting edge of big data technologies in life-sciences and gain unparalleled exposure to industry-focused data science. With a new office based in Boston, and headquarters in Cambridge UK, there will be opportunity to travel to customer sites and conferences around the world.

### **Main Duties**

- Collaborate with clients to solve key drug-discovery and other life science data analysis questions using SciBite technology, for e.g. developing scripts and pipelines to help customers achieve their goals
-

- Work with both new and existing US customers in a customer-facing role to support customer engagements, working directly with the US Sales/Tech Sales team.
- Develop use-cases, workflows and our knowledgebase to ensure we continue to build a reserve of exceptional training and “how-to” resources
- Collaborate with others in the data science team to develop new text analytics and machine-learning tools to open new avenues of data mining
- Working with partner companies to integrate and test connected solutions

### Experience and skills

Essential	<ul style="list-style-type: none"> <li>• MS plus 3 years’ experience or PhD in Bioinformatics or related field.</li> <li>• Bioinformatics, chemoinformatics or information-science background and practical experience in handling public biomedical/healthcare databases, particularly life-science text-based content</li> <li>• Comfortable in representing SciBite at customer meetings and conferences with strong interpersonal skills</li> <li>• Familiarity with major bio- and chemo-informatics public databases and handling large amounts of informatics data</li> <li>• Good Python scripting knowledge and use of RESTful web services</li> <li>• Experience in taking on new challenges outside of their comfort zone, learning new technologies as necessary</li> <li>• Motivated self-starter with ability to thrive supporting a small and growing US team.</li> <li>• Bring the right ‘Can Do’ approach to fit our company culture</li> <li>• Ability and willingness to travel up to 25%.</li> </ul>
Desirable	<ul style="list-style-type: none"> <li>• Exposure to text-analytics and text-mining and an understanding of how these can aid drug discovery and development</li> <li>• Experience in working with ontologies, taxonomies and vocabularies within a life-science context</li> <li>• Knowledge of Python data science libraries such as scikit-learn, pandas</li> <li>• Knowledge of graph database technologies, such as Neo4J</li> <li>• Knowledge of large-scale data processing and statistical data mining techniques</li> </ul>

### Application Process

To apply for this position please send a current CV and cover letter quoting the reference “SCAS2018” to [careers@scibite.com](mailto:careers@scibite.com).

Visit us at <http://scibite.com> to learn more about what we do, and how we do it.

---